Validity and reliability of the contingent valuation method:

A study of the Willingness to pay for insecticide-treated nets amongst the Igbos of Nigeria

by

Obinna Onwujekwe

Thesis submitted to the University of London
for the Degree of Doctor of Philosophy

Health Policy Unit
London School of Hygiene and Tropical Medicine

February 2002

ABSTRACT

Objectives: To contribute to knowledge on the reliability and validity of the contingent valuation method (CVM) and explore the role of context-specific CVM question formats in Southeast Nigeria. Other objectives were to determine the factors that will explain actual willingness to pay (WTP) for insecticide-treated nets (ITNs).

Methods: There was an extensive review of theoretical, methodological and empirical literature. A novel WTP question format that mimics price-taking behaviour in south-east Nigeria (called the structured haggling technique (SH) was developed and compared with the bidding game (BG) and binary with follow-up technique (BWFU). The comparisons were for inter-rater and test-retest reliability, content, construct and criterion validity and the study conducted in three villages in Nigeria. Stated WTP was determined using a questionnaire administered to 810 household heads, while actual WTP was evaluated by offering the ITNs for sale to all respondents after one month of the first survey.

Findings: There were considerable gaps in the literature regarding the reliability and validity of the CVM. In the empirical study, BG, BWFU and SH elicited reliable and valid estimates of WTP. The SH was the most content valid, while the BG and SH were the most construct-valid for ITNs and re-treatment respectively. The BG and SH were similarly criterion-valid while the BWFU was the least criterion-valid. All question formats were similar for tests of reliability. There were genuine reasons for divergences between the stated and actual WTP and for test and retest. Low-income status and physical accessibility were the major impediments to ITNs acquisition.

Conclusion: The CVM could be used to elicit valid and reliable WTP estimates in the study area, but it was not clearly proven that better content-valid question formats would lead to more valid and reliable estimates of WTP. It is necessary to further determine how the validity and reliability of the SH and other WTP question formats could be improved. Finally, future studies should establish the content validity of question formats in settings where they will be used, and use bigger sample sizes, along with allowing less time between the survey and administering the criterion, for comparing stated and actual WTP.

ACKNOWLEDGEMENTS

I am very grateful to my supervisors, Drs Julia-Fox Rushby and Kara Hanson for their invaluable advice and supervision. I also appreciate the efforts of my advisory committee, especially Jo Lines for his insight that talking to the traders would help determine the content validity of the question formats. I thank Chris Adams and Andrew Jones who helped to improve the econometric analyses.

I thank all the Staff of HPU and friends at Hakuna Matata place (Taviton Street) for their encouragement, co-operation and advice. Special thanks to my friends who proofread all the Chapters of this thesis: Annabel, Lesong, Maia, Monica, Nile, Salim, Seema, Stephen, Tamsin and Virginia.

My regards also go to Staff of the Health Policy Research Unit/Department of Pharmacology and Therapeutics, University of Nigeria, Enugu for all their support all through these years, especially Prof. PO Okonkwo and Dr Elvis Shu. I also thank my very diligent field workers who ensured that they collected usable data, especially the project manager, Eke UK and the field co-ordinator, Hillary Nzeka.

My warm appreciation goes to all the senior colleagues whose papers have provided such invaluable information for my work. They include Amiram Gafni, Cam Donaldson and Magnus Johannesson. I also thank Dale Whittington and colleagues for their papers on CVM in Nigeria.

I am most grateful to the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical diseases (TDR/WHO) for sponsoring my training. I wish to acknowledge the motivation from David Evans of WHO for my sojourn in Health economics, and Steven Wayling for his dedication in providing all the necessary logistic support for a timely completion of my study.

Without the support of my parents, siblings, relatives, loved ones and friends, it would have an arduous task to complete the PhD programme. I thank all of them and pray that all their good wishes will come true.

Finally, all Glory goes to God almighty.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS USED IN THE TEXT

AMREF = African Medical and Research Foundation

BWFU = Binary with follow-up

BG = Bidding game

CBA = Cost-benefit analysis

CDD = Community-directed distributors

CDTI = Community-directed treatment with ivermectin

CV = Contingent valuation

CVM = Contingent valuation method

DC = Dichotomous choice

ESG = Enugu State Government

EUT = Expected utility theory

FGD = Focus group discussion

ICC = Income consumption curve

ITNs = Insecticide-treated nets

MVCU = Malaria and Vectors Control Unit

NOAA = National Oceanic and Atmospheric Administration

OLS = Ordinary least squares

SG = Standard gamble

SH = Structured haggling technique

SID = Supplier-induced demand

TTO = Time trade-off

UNICEF = United Nations Children Fund

WTA = Willingness to accept

WTP = Willingness to pay

CHAPTER 1: INTRODUCTION

This thesis addresses the methodological issues of appropriately designing, implementing and establishing the reliability and validity of the contingent valuation method (CVM) and the different question formats for eliciting willingness to pay (WTP) in Southeast Nigeria. It is also concerned with the policy implications of the elicited WTP for Insecticide-treated nets (ITNs) programmes in the country. This is because the challenge for CVM research is eliciting valid and reliable WTP, and also using this information to aid policy formulation.

An existing gap in knowledge in CVM applications in determining WTP for goods/services is the reliability and validity of the CVM itself and comparatively of the different question formats in different areas and contexts, especially in sub-Saharan African countries like Nigeria. It is also not known whether guidelines for CVM in environmental and other fields of economics, such as one National Oceanic and Atmospheric Administration (NOAA) (1993), can be transferred to the health sector. Thus, the ongoing debates about the reliability and validity of CVM and of the different question formats provide the rationale for the thesis.

This chapter therefore, explains the importance of assessing the reliability and validity of CVM in determining the WTP for ITNs in Nigeria. Allied to this is whether guidelines and question formats developed for the western world are appropriate for developing countries, especially sub-Saharan African countries such as Nigeria. In addition, the basis for the development of a novel Nigerian context-specific WTP elicitation question-format, together with a comparison of the reliability and validity of the new method with bidding game and binary with follow-up question formats is provided.

The first section of this chapter presents the rationale for the thesis, while the aim and objectives are provided in the second section. Finally, the contents and structure of the thesis from chapters 2 to 10 are outlined.

1.1 Rationale

Determining WTP through the contingent valuation method is increasingly being used to generate information on the benefits of, and demand for, health care programmes. WTP is the maximum amount of income an individual is willing to give up to ensure that a proposed service or good is available (Phillips et al., 1997). The WTP could be for the availability of a resource to the individual (own use), for others e.g. the poor people (altruism) or needed by others or the individual in the future (option or non-use) or a combination of any of these measures.

CVM is widely accepted as a theoretically correct method to estimate the value of goods and services to consumers (Brookshire et al., 1980). The technique is also able to adapt to the special nature of health care (Golan and Shechter, 1993) which includes asymmetry of information, externalities, public goods, merit goods and the problems of choice under uncertainty. These characteristics of health care lead to market failures, thus making neo-classical and other demand models based on revealed preferences unable to correctly value consumer surplus. An informative and plausible CVM scenario acts to limit the effects of asymmetry of information and uncertainty, thus enabling respondents to make more informed and potentially more reliable and valid choices.

Nevertheless, there are questions about the validity and reliability of CVM (Mitchell and Carson, 1989; Diamond and Hausmann, 1994; Liljas and Blumenschein, 2000). The major criticism of the validity of CVM has been that stated WTP is a poor indicator of actual WTP (Diamond and Hausmann, 1994). The possible reasons for the divergence between stated and actual WTP is that people lack experience with aspects of CVM itself or lack knowledge about the good (Kealey et al., 1988).

There is even less evidence about the reliability and validity of CVM in the health care sector and developing countries where the technique has been used sparsely. Further studies are needed to explore questions regarding the reliability and validity of CVM in the health care field (Johannesson et al., 1992). There is equally the need for studies in the health care sectors of developing countries that will shed light on the validity and reliability of CVM in these contexts.

The methodological issues regarding the use of CVM to generate valid and reliable measures of consumer surplus have been explored in detail (Mitchell and Carson, 1989; Donaldson, 1993; Johannesson, 1996; Klose, 1999; Smith, 2000; and Liljas and Blumenschein, 2000). Morrison (1998), Green et al. (1998) and Alberini (1995) have tried to unravel the reasons behind the biases that reduce the validity and reliability of valuations of consumer choice and how CVM can be improved.

The debates or issues explored in the previous studies centred on whether increasing the realism of CVM would improve reliability and validity of the method, the optimal study settings, the question formats that should be used, the specification of the contingent market and the causes and how to manage biases in contingent valuation studies. Further debates have included optimal data analytic methods for determining the reliability and validity of CVM values. However, all these issues still remain unresolved, and still need to be further explored for conclusive results to be generated.

One potential approach to improving the reliability and validity of CVM is increasing the realism of the contingent market. This refers to realism in specifying the item to be valued, describing the contingent market, explaining the method of payment and selecting a questioning format for eliciting values (Reiling et al., 1990). Mitchell and Carson (1989) recommended increased realism of the CVM survey as a strategy for increasing the reliability of WTP estimates. Thus, Reiling et al. (1990) argued that if the respondents are familiar with making payments for a resource on an annual basis, a valuation exercise asking for weekly payments for the resource may not be reliable or may reduce the reliability of the estimated contingent values.

It is also important for CVM designs to take into account local settings in the study design, so that the results will be acceptable to policy makers and those who are interviewed, who might question the reliability and validity of the results that were generated using unfamiliar methodologies. It is important to use question formats that are realistic and thus, have practical relevance to the respondents otherwise they may view the whole exercise with scepticism and give invalid responses. Also, the use of question formats may make it mentally and psychologically difficult for them to state valid and reliable WTP amounts.

O'Brien and Gafni (1996) posited that to assure content validity of the WTP scenario, the valuation tasks and choices should be realistic, e.g. by choosing usual payment vehicles in the relevant market being studied. Hence, Macmillan et al. (1999) argued that lack of realism, through its influence on incentives to miss-represent payments may explain divergences between stated and actual WTP, thus translating to less valid WTP estimates in contingent valuation surveys.

Kealey et al. (1988) posited that comparisons across studies suggest that the discrepancy between hypothetical and actual CVM valuations, are smaller when the commodity is more important to the respondent and that increasing the realism of CVM scenario could induce more intention-consistent or truthful behaviour. Hence, it would be arguably easier to increase realism with private goods e.g. Insecticide treated nets (ITNs), because people would relate more to them and may also be willing to acquire them for personal use.

Greater experience with the question formats, perhaps allowing the respondents to learn the market rules better, improves the validity of the WTP elicited (Kealey et al., 1988) because it helps the respondents to reveal their true choices. Thus, it could be argued that if the respondents are trained extensively on a question format, payment vehicle or other intention-consistent attributes of the valuation, then the validity and reliability could be improved. Nevertheless, the costs of bringing the respondents to this optimal level of experience with the either all or some aspects of the contingent market may be relatively high compared to a situation where the question format had prior realism. However, if it is not possible to use realistic mechanisms in the valuation, training the respondents to understand the good, question format, payment vehicles and other attributes important for the valuation would be the next alternative for increasing the reliability and validity of CVM.

Although validity and reliability of CVM could be improved by increasing the realism of CVM, most environmental resources lack the necessary attributes for achieving intention-consistent behaviour (Kealey at al, 1988). This is because these resources are not usually concrete commodities, are not well defined and people lack market experience with them, and thus potentially less valid and reliable WTP could be elicited (Kealey at al, 1988). Some studies have tried to increase the realism of the market in

environmental economics by showing respondents pictures of the resources they would value (Loomis et al., 1996).

The consistency of the study setting has implications for the reliability and validity of WTP. Hanemann (1994) argued that most studies used unsatisfactory designs, leading to less valid and reliable results. Thus, in cases where it is not possible to ensure that the study setting is realistic, attempts should be made to ensure that the hypothetical setting would not confound the valuation of consumer preferences. For instance, Seip and Strand (1992) compared hypothetical payments from a phone survey with actual payments obtained from a mail survey, thus confounding differences in study design and implementation (Macmillan et al., 1999).

In summarising the points for and against whether the effect of more realistic CVM increases the reliability and validity of the valuation, it is argued that increased realism motivates intention-consistent behaviour, as behavioural intent resembles actual behaviour when they share similar settings (Fischoff and Furby, 1988). This is because one would have controlled for all the exogenous factors that could lead respondents to not reveal their true preferences and hence the reliability and validity of the elicited WTP is likely to be greater. However, further experimentation with different question formats, question framing methods and payment vehicles may prove worthwhile in helping to determine the reliability and validity of the CVM (Kealey et al., 1988).

Since the elicitation of valid responses is the key to the credibility of CVM, the choice of CVM question formats should hinge on what makes market sense to the respondents and mimicking their normal price-taking behaviour for either the goods being valued or for similar goods. Thus, if people usually buy goods or services by bargaining, the question format should resemble the bargaining system, and if the usual system is by fixed take-it-or-leave-it prices, the question format should be a binary choice and so on. The payment mechanism specified in the scenario for eliciting WTP should also reflect reality. Thus, insurance-based payment, out-of-pocket, general taxation mechanisms etc should be specified in areas where they are the dominant way of paying for either the goods or services under valuation or for similar goods or services.

The use of question formats that mimic usual price-taking behaviour in study areas will ensure that CVM surveys limit the abstract nature of the exercise and increase the accuracy (more reliability and validity) of measurement in a particular context. More realistic question formats would also better enable the respondents to think in the normal manner they do in everyday market situations. Plus, the respondents would better relate to the questions being asked, because the hypothetical nature of the survey has been reduced by the use of a familiar question format. Finally, the respondents may take the survey more seriously instead of viewing it sceptically, and feeling that the results will not be used for any meaningful or practical purpose.

Thus, identifying the appropriate question formats for different settings is a key research area for CVM. There are many CVM question formats for eliciting values and they have been described in detail (Mitchell and Carson, 1989; Russel et al., 1995; Diener et al., 1998; Klose, 1999; Smith 2000). There are still ongoing debates about which question formats elicit valid and reliable WTP in different settings (Mitchell and Carson, 1989; NOAA, 1993; Smith, 2000; Onwujekwe, 2001; Kealey and Turner, 1993). Unreliable or invalid questions could lead to the wrong conclusion that the CVM is not suitable in a particular situation and/or lead to erroneous policy recommendations.

The NOAA panel (NOAA, 1993) recommended the use of the dichotomous choice format because that panel felt that the technique offers a more realistic market situation and would yield more valid responses than open-ended methods. However, Smith (2000) pointing to the biases inherent in the dichotomous choice format like yea-saying, starting point-bias (in double or triple bounded formats) and statistical inefficiency, argued that there is no reason why other question formats should not be used to elicit WTP. Smith thus, recommended that research should be targeted to establishing the relative reliability and validity of different question formats in health care.

Haneman (1991) and NOAA (1993) argued that the open-ended technique does not mimic price taking behaviour in markets, in the western world, and so is not to be recommended (Loomis et al, 1997). However, the NOAA recommended dichotomous choice method does not mimic price-taking behaviour in many sub-Saharan African market situations and therefore may not be an ideal technique to be used in those areas. Balogun (1991) pointed out the cultural aspect of bargaining behaviour of Nigerians as

"different from the European or American pricing system which usually has fixed prices in open markets".

There is no conclusive theoretical or empirical justification for considering one question format to be more valid than another (Smith, 2000). Studies that have compared stated and actual WTP found that different question formats yielded similar results (Brown et al, 1996; Loomis et al, 1997; Frykblom, 1997; and Onwujekwe, 2001). Thus, the recommendation by NOAA (1993) that the dichotomous choice format should be used in preference to the open-ended method should be reconsidered, as it has not been proven to be the most reliable and valid method (Lunander, 1998; Smith, 2000).

The mixed evidence of whether realism is important for the reliability and validity of WTP elicited by different question formats could be because other intention-consistent attributes or requirements for realism within contingent valuation are not met in studies. Thus, the effect of the question format may not be apparent, if it is confounded by the absence of other factors required, to improve the reliability and validity of the CVM. Conversely, if everything else is realistic, then the effect of the question format becomes the only dimension of realism that matters and thus the comparative reliability and validity of different question formats could be better assessed.

Unless studies comparing different question formats ensure that other intention-consistent behaviour aspects of CVM e.g.: realism in goods and in other aspects, are met, there will not be conclusive evidence about the appropriateness of different question formats in different contexts. Researchers also need to control for possible causes of divergence between the stated and the actual WTP prior to comparing different question formats.

Refining the design and implementation of contingent valuation question formats should thus, continue to be a research priority as there is potential to decrease bias in contingent valuation from this source (Kealey et al., 1988). Smith (2000) argued that the knowledge about the appropriate question format for a context (good, study area, study setting, etc.) would be better known by determining the criterion validity of WTP estimates. Until such studies are undertaken, researchers will not know which is more valid and would be relying on opinions and drawing their own conclusions (Smith, 2000).

The CVM question format is one tool that is unlikely to be transferred wholesale successfully from one context to another because of price-taking differences and hence should always be indigenous to the area or context where it will be used (Onwujekwe, 2001). Although the application of CVM technique in health care is growing rapidly (Diener et al., 1998; Klose, 1999; Olsen and Smith, 2001), there is still no consensus on the best question format to use in this sector (Onwujekwe, 2001). However, for improving validity and reliability, the key principle behind the choice of question format for a particular study should be to use the one that best accounts for the local setting and the realism of the contingent valuation method, for improved validity (Onwujekwe, 2001).

My argument is that the best rationale for using any CV question format is that it mimics the market price-taking behaviour of the area or context in which it will be used. Context in this instance refers to the culture of the people, the country, nature of goods (e.g. ITNs) and bargaining strategies for either the goods in question (e.g. ITNs) or similar goods (e.g. untreated nets) and services. Clarke (2000) sought to know why a respondent's WTP declines when s/he is asked a follow-up question and the explanation was that the follow-up question comes as something of a surprise. The surprise at being asked a different amount in the follow-up question might induce a negative reaction (Clarke, 2000). However, a different scenario occurs in Nigeria, where respondents expect a follow-up question preparatory to bargaining. This example illustrates the cultural contexts that need to be considered when using different WTP question formats in different areas and contexts.

Thus, urgent work is needed to determine the content validity of different question formats in different settings and for different products. In particular, such studies are needed in sub-Saharan Africa where CVM is still in its infancy and where market behaviours are different from the Western world which mostly uses the dichotomous choice method or its variants to elicit WTP. Qualitative research methods like focus group discussions with the general populace, observations of market behaviour and interviews with actual practitioners of market price elicitation (traders) could be used to determine the content validity of the CVM scenario and different question formats. It would then be possible to determine whether better content validity of a question format translates to better construct and criterion validity, together with better estimates of reliability, ultimately translating to lower occurrence of hypothetical bias.

The ideal CVM question format in Nigeria for goods like Insecticide-treated nets would mimic the haggling technique normally used in the markets for the purchase of goods and services. In Nigeria, and in most of sub-Saharan African, people haggle for goods and services before deciding on the final amount to pay. The bidding game and the binary-with-follow up techniques to some extent resemble the haggling process in rural markets since the respondents are given at least one more opportunity to either increase or decrease their bargain prices. However, the bidding game and binary-with follow-up methods are still limited by the bargaining options they offer to the respondents.

Thus, a novel WTP question format that more closely mimicked the price bargaining mechanisms found in day-to-day markets in Southeast Nigeria was designed. The novel technique is called the "Structured Haggling Technique". However, concerns with regard to the development of any new methodology are whether the methodology measures the constructs the researcher desires to measure and how reliable the measurements are (Loomis et al, 1989).

The novel technique, together with the bidding game and the binary with follow-up technique was subjected to tests of content, construct and criterion validity, along with inter-rater and test retest reliability. This was to see whether more realistic question formats that mimic the price-taking behaviour in the market improve the reliability and validity of CVM. Such a strategy of holistically examining multiple tests of validity and reliability within one study setting is likely to provide better information on the reliability and validity of CVM.

This comparison would inform the CVM field on the performance of the three question formats and on the overall reliability and validity of the contingent valuation method. The study involved theoretical and empirical literature review, together with determining the stated and actual WTP for ITNs and stated ITNs re-treatment amongst the "Igbos" of rural Southeast Nigeria.

In order to objectively determine the reliability and validity of the three question formats, appropriate data analytical methods need to be used. Further methodological debates in CVM research concern the best ways for determining construct and criterion validity, together with estimating the reliability of the elicited WTP from different question formats and, by implication, of the CVM itself.

Tests of theoretical constructs through the use of econometric modelling of WTP is the most common test of construct validity, with convergent and discriminant validity being the others. However, the choice of the best econometric model to use in examining the theoretical constructs for a particular data set is still a focus of debate in CVM. If the wrong model is used, one could erroneously conclude that the elicited WTP was either valid or invalid.

Donaldson et al. (1998) posited that the way the data is gathered should determine the model that should be used, with regression diagnostic tests performed to ensure that the right model was used. This is so that robust conclusion can be made about the construct validity of the elicited WTP, as different models potentially lead to different results. For instance, in modelling WTP values that were limited at zero, Donaldson et al (1998) found that the Heckman selection model was the best when compared to the Tobit and ordinary least squares (OLS) models. However, for the same type of dependent variable, Gyldmark and Morrison (2001) found that log OLS was better than the Heckman selection model and Dalmau-Matarrodona (2001) found that the Double-Hurdle model was better than the Tobit model.

Comparisons of stated and actual WTP data would inform the extent to which CVM is criterion valid and the relative criterion validity of different CVM question formats. The few studies that have compared the criterion validity of different question formats found no differences in criterion validity of WTP generated by different question formats. The earlier similar findings have led Loomis et al. (1997) to question whether the question format mattered. However, Smith (2000) argued that to resolve the issue of the appropriate question format to be used requires empirical research that compares all types of survey questions with real choices, since this would be a better assessment than using hypothetical data only.

Criterion validity is the strongest validity test because it compares whether stated WTP will correctly predict actual behaviour. It is also devoid of the assumptions and the influences of the data analyst that occur in tests of construct validity. Although people are sceptical of the accuracy of responses to hypothetical questions (Dickie et al., 1987), very few studies in health care and in other sectors have attempted to determine criterion validity. This is possibly because most WTP applications have valued resources that are public goods or there was no gold standard for investigating criterion validity.

The results of the comparative criterion validity of different question formats that found no differences between different question formats cast doubt on the NOAA (1993) recommendation that the dichotomous choice format should be used in CVM applications. Nevertheless, most of the studies were exploratory, using very small sample sizes, inappropriate study designs and data analytic. Thus, the criterion validity of CVM has yet to be proven.

Most of the studies that have determined criterion validity found divergences between actual and stated WTP, and this divergence has been described as "hypothetical bias" (Cummings et al., 1986). Divergences in WTP could be because "hypothetical demand data are subject to payment bias which arise because hypothetical situations may not provide sufficient incentive for respondents to reveal their true preferences" (Dickie et al., 1987).

However, an important question is, "if the stated and actual WTP differed, can we conclude that the differences are unambiguously attributable to differences in real and hypothetical payments?" (Neil et al., 1994). Unfortunately, none of the previous studies that had explored convergent validity could answer this question since none explored reasons for the divergences.

One reason for a divergence between stated and actual WTP may be that the factors influencing WTP may change over time. Tastes, preferences and other major factors affecting demand like income, changes in relative prices and prices of substitutes are dynamic and these changes could lead to divergences between stated and actual WTP. Thus, the divergences would be reflecting real life changes. Without taking these into account, criterion validity may indicate lower test scores and may lead to the erroneous conclusion that CVM or the question formats used were not valid due to hypothetical bias. Therefore, tests of criterion validity leading to conclusions about the presence of hypothetical bias require that reasons for the divergences should also be explored. This is a better way of informing on the criterion validity of CVM and of different question formats. Furthermore, Dickie et al. (1987) argued that comparisons of stated and actual WTP data would be of value in assessing the relative usefulness of hypothetical data as well as for indicating ways to improve demand-revealing data collection methods.

There is a dire paucity of data, on the reliability of the CVM in all sectors of economics. Thus, the reliability of CVM and of the different question formats has not been proven and it (like validity) is a necessary condition for accurate contingent valuation results (Reiling et al., 1990). Test-retest is the most commonly applied test. However, as with tests of criterion validity, changes in economic factors and other determinants of demand could lead to lower test-retest reliability results than expected. Hence, test-retest reliability should also be accompanied by the exploration of factors that could lead to low reliability scores.

It is also necessary to explore reliability tests that could be used in CVM where the effect between time intervals could be limited. The possibilities are inter-rater and intra-rater reliability. Better and more holistic information on the reliability of CVM and of different question formats could be obtained if these other reliability tests together with test-retest are used in the same study. Additional use of other tests of reliability would be feasible in cases where studies are not designed for the retest.

Dickie et al. (1987) in exploring some of the weaknesses of CVM found interviewer effects on the WTP estimates because people differ in their natural abilities and salesmanship. Hence, Dickie et al. advocated that future studies should have the same respondent interviewed by more than one person at different times; a test of inter-rater reliability. Therefore, there are opportunities for all WTP studies to determine the reliability of their valuations, as this information is a major weak link in the debates about the accuracy of CVM. If feasible, studies could determine more than one type of reliability and use the results as a basis for better informing the level of reliability of the estimates.

Eliciting reliable and valid WTP for ITNs is important, as the policy makers need to be know how people value goods and services, which could inform policy decisions on the level of government intervention in delivering and financing goods and services, such as ITNs. Furthermore, the analysis of elicited WTP would also present the distributional implications of using the WTP amounts for implementing cost-recovery efforts, as the groups that would be excluded from benefiting would be known. The government for instance could, based on the information, distribute the nets free of charge to poor and high-risk groups while charging others.

The information about how people value the nets through the measurement of consumer surplus could also be used to inform cost-benefit computations of whether the society feels that the ITNs are worthwhile interventions. However, if the results indicate that people do not value the nets or that the potential net-benefit is minimal, the government might want to stimulate interest by conducting health education campaigns. The factors that cause divergences between actual and stated WTP could also inform the government on where to place emphasis in stimulating demand for the nets and in developing financing and distribution mechanisms for the ITNs programme.

1.2 Aims and Objectives

1.2.1 Overall aim

The thesis aims to contribute to knowledge on the reliability and validity of the contingent valuation method and within this context, explore the role of appropriate context-specific question formats for eliciting willingness to pay in Southeast Nigeria. It also aims to explore the policy implications of the results for enhancing the coverage of ITNs in Nigeria.

A context-specific question format is compared to two existing ones. Thus, the study generates information to fill key methodological and empirical gaps in knowledge about CVM, especially as applied in Nigeria. It is the first to compare three question formats and to determine inter-rater reliability and content validity of the question formats. Also, the steps in the design and validation of the novel structured haggling technique will be adding new information to the field.

The thesis generates needed information on the reliability of the CVM and different question formats, especially "as there continues to be a view that the reliability of the WTP method has not been proven" (Klose, 1999). It also contributes to knowledge about the use of validity tests in CVM, especially criterion validity, and undertakes a comparative content, construct and criterion validity of three question formats.

1.2.2 Specific objectives

1 To review the evidence on the theoretical and empirical CVM related-literature in order to understand the gaps in knowledge and develop the framework for the empirical part of the thesis.
2 To inform on the reliability and validity of the CVM amongst the Igbos of Nigeria.
3 To develop an indigenous CVM question format that will mimic the usual haggling price-taking behaviour among the Igbo people of Nigeria.
4 To determine the reliability (inter-rater and test-retest) and validity (content, construct and criterion) of the novel structured haggling technique, the bidding game and the binary with follow-up methods; also exploring why test-retest reliability could be lower than expected.
5 To develop and use methodological frameworks for determining content, construct and criterion validity.

6 To determine whether divergences between stated and actual WTP always means the presence of hypothetical bias.
7 To explore the policy and methodological implications of the findings for the CVM and ITNs programmes.

1.3 The organisation of the thesis

The thesis is organised into 10 chapters. The contents of the remaining chapters are summarised in the following paragraphs.

Chapter 2 provides the theoretical underpinnings of the contingent valuation method. The chapter thus starts by reviewing the Kaldor-Hicks formulations on WTP and WTA followed by the principles and measurement of consumer surplus. The special characteristics that may make standard demand models inappropriate to value consumer surplus in health-care, together with the alternative demand models that have been developed in response to these features and why even these models may fail to fully value consumer surplus are then detailed.

Chapter 3 details the methodological framework for operationalising CVM, together with the framework for determining the reliability and validity of the method. The components of CVM are described, together with descriptions of the question formats. Then the tests of reliability and validity are presented together with the type of biases they could detect. Finally, a framework describing the concept of the empirical part of the thesis is described.

In Chapter 4, a critical review of the empirical literature on WTP is presented. The critique particularly focuses on the question formats and approaches to testing reliability and validity. It begins by considering approaches to testing content, construct and criterion validity, before delving into reliability. In each part, the contexts of the studies, the question formats and tests used were reviewed. Specifically, in the case of construct validity the type of models used were examined and in criterion validity the findings of divergences between the stated and actual WTP were explored to see whether the studies provided an insight into interpreting these divergences. Studies of reliability were explored to appraise the causes of lower than expected test-retest scores. The literature review sets the agenda for future research and locates the thesis within the CVM literature.

The purpose of Chapter 5 is to present the methods used in the empirical part of the study. The objectives of the empirical part are first articulated. Then the overall framework of the study is presented. This framework includes information on the study area, background information on malaria and ITNs, the nature of the goods to be used and the price-taking mechanism among Igbos of Nigeria. The study outline is then presented after which the three major arms of the study are described in three sections respectively. The first section deals with qualitative research for establishing and improving the content validity of the different question formats and the CV scenario. The next two sections present the methods used for the two major surveys for determining inter-rater reliability and construct validity (1st survey) and test-retest reliability, criterion validity and causes of divergences between actual and stated WTP (2nd survey). The basic design of the structured haggling technique, together with the rationale for choosing the bidding game and the binary with-follow-up for comparisons with the structured haggling technique are also presented.

Chapter 6 presents a comparative analysis of content and construct validity of the different question formats. In the case of content validity, the results from the observations of trading activities and price-taking for untreated mosquito nets (no ITNs yet in the markets) are presented, along with from interviews with sellers of ordinary nets and focus group discussions (FGDs) with consumers. The steps leading to the final design of the structured haggling technique is detailed along with the steps taken to improve the content validity of the WTP scenario. The information on response rates, together with testing for similarity in demographic factors among respondents in the question formats, are presented in the next section of chapter 6. This is followed by alternative econometric models that were used to determine the construct validity of WTP for the ITNs and their re-treatment.

Chapter 7 focuses on the results of the comparative criterion validity of the different question formats. The results presented in the chapter were derived from the first survey that elicited stated WTP and the second survey that accompanied the sale of ITNs. The levels of convergence between stated and actual WTP analysed using phi correlation coefficient and positive predictive validity tests, are presented. The causes of variation and direction of divergences between stated and actual purchase decisions are presented using descriptive and non-parametric statistical methods. Finally, the factors that could explain actual WTP are also modelled econometrically.

Chapter 8 compares inter-rater and test retest reliability across the three question format. The results of correlation coefficients for the different estimates of WTP are presented. In order to provide a clearer understanding of the test-retest reliability coefficients, the causes of variation between the levels of WTP elicited in the first survey and in the retest are presented. Furthermore, the reliability of responses regarding a number of socio-economic and demographic factors is also assessed.

The results are drawn together and discussed as a whole in Chapter 9, and integrated along the themes represented by the specific objectives of the thesis. Thus, the key questions debated are the comparative reliability and validity of the novel contingent valuation technique compared to the bidding game and the binary with follow-up methods. Others were the extent to which the thesis filled the gaps in knowledge and whether CVM was reliable and or valid in the thesis, together with whether divergences between stated and actual WTP always means the presence of hypothetical bias. The implications of the results for ITNs markets and resource allocation are also discussed. Finally, the problems and limitations of the thesis are presented together with suggestions for improvement in future studies.

Chapter 10 draws together the principal policy and methodological conclusions from chapters 2 to 9. These conclusions are made with respect to the fulfilling of the aims and objectives of the thesis. Suggestions for future research in both CVM itself and the use of context-specific approaches are also provided.

CHAPTER 2: THEORETICAL FRAMEWORK OF THE THESIS

This chapter reviews the theory underlying the use of the contingent valuation method (CVM) to determine willingness to pay (WTP) for goods and services. The measurement of consumer surplus using welfare economics principles is the major theoretical underpinning of CVM. Thus, the review will establish the link between WTP determined through CVM to the notion of measuring consumer surplus within welfare economics.

Gafni (1998) distinguishes two alternative uses for WTP questions. The first is to establish the social value of a good or service using cost-benefit analysis (CBA), which relies on the normative foundations of welfare economics. The second is to investigate the relationship between price and demand in order to inform pricing and marketing decisions, which is based on positive consumer theory.

Nevertheless, the boundary suggested by Gafni is not as clear as it may first appear. This is because, in practice, governments or organisations may want to use elicited WTP to both value societal benefits of a health programme and to determine the price to promote equity.

As "WTP is inevitably associated with ability to pay" (Donaldson et al, 1997a), the examination of WTP across socio-economic groups could be used to ensure distributive justice in resource allocation, thereby actually maximising the societal benefits of interventions. This is especially applicable in situations where subsidies or exemptions from user fees are to be implemented.

Therefore, most uses of CVM to estimate WTP for health-care interventions (whether for CBA or to understand demand), have some basis in welfare economics. This is distinguishable from the market research by private companies for profit maximisation, where they sometimes apply quick methods to elicit WTP as will be described in section 2.5 of this chapter.

Policy decisions based on the measurement of the welfare consequences of health programmes is rooted in the Pareto principle; that a programme is worth investing in if it makes at least one person better off without making anyone worse off. However, since the majority of policies produce both gainers and losers, the Pareto principle was modified in the form of the potential Pareto-improvement compensation criterion (Golan and Shechter, 1993; Donaldson, 1993).

The proposition in welfare economics is that individuals should be considered the best source of information on their welfare (Drummond et al., 1997). Welfare economics uses the concept of consumer surplus, measured as the area under the demand curve and above the market price, to measure the value of a resource to an individual. CVM is a method for estimating consumer surplus through the measurements of WTP and or willingness to accept (WTA). WTP and WTA are both derived from the Kaldor-Hicks compensation criterion that has Paretian welfare economics as the theoretical basis (Johannesson, 1996).

The chapter thus starts by reviewing the Kaldor-Hicks formulations on WTP and WTA. This leads to expositions on demand theories in order to provide a basis for understanding and measuring consumer surplus. The principle of consumer surplus is presented before moving onto the issues involved in its measurement. The CVM is used in the health-care sector because health care has a number of special characteristics that may make standard demand models inappropriate to value consumer surplus. Hence, in the last part of this chapter, I review the special characteristics of health-care, together with the alternative demand models that have been developed in response to these features and why even these models may fail to fully value consumer surplus.

2.1 Paretian welfare economics

Welfare economics makes two crucial assumptions regarding the consumer (Culyer, 1991); The first, a normative judgement, is that an individual's own interpretation of his welfare is the one that counts. The second, a non-normative (or positive) judgement is that choices reveal preferences (if people are not behaving strategically). The first fundamental value judgement made in welfare economics is the Pareto principle (Johannesson, 1996).

The first theorem of welfare economics states that, a competitive general market equilibrium is under certain assumptions Pareto-optimal (Boadway and Bruce, 1984). The second theorem of welfare economics states that under certain assumptions it is possible to attain any Pareto-optimal situation as a result of a competitive general market equilibrium given any distribution of income (Boadway and Bruce, 1984). There are a number of different possible Pareto optimal situations in an economy and which point is reached in a perfect economy will depend on the initial distribution of income (Johannesson, 1996).

The basic Pareto principle that a change should take place if it makes at least one person better off whilst no one is made worse off, is very limiting (Donaldson 1993). This is because, the majority of policies produce both gainers and losers, thus making the Pareto principle of little practical use (Johannesson, 1996).

There exist a number of cases where the market fails to achieve an efficient outcome leading to market failure. This provides an argument for public interventions in the market (Johannesson, 1996). According to Johannesson, interventions in markets that have failed lead to changes where some individuals gain and some other individuals lose, and in those cases the Pareto principle cannot be used to determine whether a change should be carried out or not.

The theoretical solution to this problem is the Kaldor-Hicks criterion, otherwise known as the potential Pareto-improvement principle (Kaldor, 1939; Hicks, 1941, 1943) or compensation principle (Donaldson, 1993; Johannesson, 1996). The potential Pareto criterion is that if WTP of the gainers is greater than compensation demanded by the losers of a policy change, the change should go ahead, whether or not the compensation is paid; consistent with the economic principle of efficiency (Donaldson, 1993).

Kaldor defined the consumers' value in terms of the amount gainers from a policy change are willing to pay the losers and still be better off than without the change. However, Hicks defined consumers' value in terms of the amount the losers are willing to accept from the gainers and still be better of with the change. These two definitions were brought together in formulating the Kaldor-Hicks criterion.

The Kaldor-Hicks criterion states that a proposed policy change is desirable on social welfare grounds if everyone can potentially be made better off, meaning that a project is acceptable if the gainers compensate the losers (Golan and Shecheter, 1993). Golan and Shecheter posited that the possibility of compensation does not imply actuality, but instead indicates the existence of a possible net gain to society with the distributional decisions left to policy makers.

The disadvantage of the Kaldor-Hicks criterion is that it ignores considerations of distributive justice and can lead to decisions which favour the rich at the expense of the poor simply because the rich have a greater ability to pay. This is because, the criterion basically asks, "are the gainers from a change able to compensate the losers and still remain better off?" (Donaldson, 1993). Because of this, Donaldson noted that a welfare criterion was introduced to the Kaldor-Hicks criterion by Scitvosky (1941) and Little (1949, 1957), allowing for change to take place as long as direct redistribution cannot take place. Distributional or equity weights could be used to improve the equity-aspect of resource-allocation based on estimates of Hicksian surplus (Donaldson, 1999). However, Sassi et al (2001) queried the usefulness of such weights for resource allocation since their sources are arbitrary.

An illustration of Pareto principles and Pareto improvements is provided in figure 2.1 (Drummond et al, 1997).

Figure 2.1 Box showing Pareto principles and Pareto improvements

Key assumptions
Social welfare is made up from the welfare (or utilities) of each individual member of society.
Individuals are the best judges of their own welfare (consumer sovereignty).

Principles
Actual Pareto improvement: A policy that makes one or more persons better off and makes no person any worse off.
Potential Pareto Improvement. (Kaldor-Hicks criterion): A policy that creates gainers and losers in welfare, but if the gainers could compensate the losers and remain better off themselves after the change then society as a whole has benefited. Because compensation does not actually have to be paid, this principle raises some equity issues about who gains and who loses.

Unless a project or program is proposed primarily for the purpose of redistribution, the potential Pareto-improvement is the proper criterion for cost-benefit analysis (CBA) and, therefore, Hicksian compensating measures are the proper measures of value (Brookshire et al, 1980). O'Brien and Viramontes (1994) posited that the practical extension of the theory underlying CBA is the estimation of the individual's maximum willingness to pay (WTP) to secure implementation of a program or the minimum individuals would need to be compensated (willingness to accept (WTA)) to forgo a programme.

The foundation of welfare economics is neo-classical demand theory. This will be reviewed in the next section. This will be followed by a discussion of the notion of consumer surplus that forms the theoretical underpinning of the contingent valuation method.

2.2 Theories of demand

The demand for medical care, as for any other good or service, is concerned with a willingness and ability to pay (Fuchs, 1968). As noted by Dupuit, (1844), "political economy has to take as the measure of utility of an object the maximum sacrifice which each consumer would be willing to make in order to acquire the object ...the only real utility is that which people are willing to pay for" (Donaldson, 1993).

Demand under certainty is also termed the rational model of choice (Arrow, 1984). It forms the basis of the neo-classical demand theory. Samuelson's revealed preference theory and Hick's indifference curve theory, are extensions to Marshall's neo-classical demand theory (McGuire et al, 1988).

The different demand theories could all be used to derive the demand curve, which "is observable in the real world, and though derived from a fairly abstract utility theory, it offers an observable model of consumer behaviour" (Phelps, 1992). The demand theories are functions of relative prices but those of the neo-classical demand theorem are derived from a constant level of money income and those of indifference curve analysis for a constant level of utility (Dinwiddy and Teal, 1996)

Neo-classical demand theory is based on the cardinal utility theory. This assumes an economically rational consumer who aims to maximise utility. The principle of marginal utility is used to understand demand and there are five important postulates (axioms)[1] of the theory (Culyer, 1985). Using marginal utility and price changes, the demand curve could be derived from the neo-classical demand theorem.

[1] They are: an economic good yields utility; the more of a good an individual consumes, the greater his or her utility; the marginal utility of a good diminishes as more is consumed; the marginal utility of money income is constant; and individuals maximise utility subject to their money income.

The drawbacks of Marshall's neo-classical demand theory include the assumption of certainty; that utility and hence marginal utility of a good depends only upon the amount of that good and is independent of other goods[2] and; that money income is held constant (Culyer, 1985). As the author argued, many of Marshall's strongest assumptions[3] are not necessary to derive a demand theorem, and consequently Marshall's demand theorem is no longer generally accepted as a satisfactory account of the demand phenomenon due to its internal contradictions, its applicability to trivial goods and its strong axioms. The drawbacks of this demand theory led to the development of alternative demand theories.

The revealed preference theory[4] was developed by Paul Samuelson based on the need to overcome the subjectivity of the neo-classical demand theorem. The theory is based solely on observable and measurable phenomena, namely the bundles actually bought by a consumer and the prices and money incomes at which they were bought (Gravelle and Rees, 1981). The revealed preference theory is more general and satisfactory than Marshall's theory and has two axioms[5] (Culyer, 1985). It also introduced the notion of using the budget line to understand demand. However, though the theorem is popular, it has some drawbacks that limit its use.

One drawback of revealed preference theory is the axioms that require strong ordering of goods. It proposes that individuals are always able to express a preference for a good over another. It does not allow for indifference between two bundles (Culyer, 1985). There are also many problems with the transitivity axiom, as it may not hold in many cases. Furthermore, revealed preference theory assumes that utility can be measured ordinally. It is, therefore, not strictly true that the theory does without utility measurement. However, it uses a much weaker measurement than Marshall's theory (Culyer, 1985).

---

[2] This implicitly assumes that goods are neither complements nor substitutes (Culyer, 1985).
[3] The measurability of utility on a ratio scale and the independent additivity of component utilities to derive the utility of the total.
[4] The theorem states that if the demand for a good increases when the income alone rises, then quantity demanded must definitely fall (rise) when the price alone rises (falls) (Culyer, 1985).
[5] a. Comparison axiom: there is a comparability between bundles of goods such that either B1 P B2 OR B2 P B1 (B1, B2 & B3 are bundles of goods having different amounts of, say, goods X and Z in them. The relation P stands for 'chosen rather than').
b. Transitivity axiom: If B1 P B2 AND B2 P B3, then B1 P B3.

Indifference curve theory is the most widely used theory of choice in economics and was developed by Sir John Hicks (Culyer, 1985). Like the two other theories already described, it is axiomatic[6]. However, it assumes that the behaviour of individuals is governed by consistent pattern of preferences with respect to varying bundles of goods (Arrow, 1984).

The superimposition of an indifference map on a diagram showing an individual's budget line is used to derive demand from the indifference curve. The diagram is used to find the equilibrium condition of a consumer, where utility is maximised. If consumers act according to the model portrayed by the indifference curves, then, we infer the standard "demand curve" (Phelps, 1992).

The convexity axiom of the indifference curve implies that there is some amount of one good that an individual will sacrifice to obtain more of another. The maximum amount voluntarily sacrificed for a little more of another good, divided by the extra amount of good, is the marginal rate of substitution (Culyer, 1985). This means that one is indifferent between the pre and post sacrifice situations. Thus, the slope of the indifference curve is the marginal rate of substitution.

Critics of the indifference curve analysis argue that the method may be satisfactory if only two or three commodities are involved, but that the procedure will be useless if there are several hundred commodities. However, proponents feel there is no real benefit in considering more than two commodities because in principle, all properties of multi-commodity indifference maps are deducible from those of two-commodity maps (Arrow, 1984).

---

[6] The 3 axioms of the indifference curve theory (Culyer, 1985)

a. Comparison axiom: There is comparability between bundles of goods such that either B1 P B2, B2 P B1 or B1 I B2 (and B2 I B1, of course).
b. Transitivity axiom: if B1 P B2 and B2 P B3, then B1 P B3; and if B1 I B2 and B2 I B3, then B1 I B3. Likewise, if B1 I B2 and B2 P B3, then B1 P B3; and if B1 P B2 and B2 I B3, then B1 P B3.
c. Convexity axiom: The marginal rate of substitution of one economic good for another diminishes the more an individual consumes of the one, utility remaining the same.

Norris (1952) in comparing Marshall's neo-classical and Hicks' indifference curve demand theorems noted that: apart from giving us an orthodox demand schedule by means of the indifference curve approach, Hicks was able to tell us how much of a commodity will be purchased if its price remains unchanged. As the individuals' income increases, this leads to an increase in the purchase of normal goods. Norris concluded that there was no counterpart to this Income consumption curve (ICC) in Marshall's approach, for he explicitly assumed that the marginal utility of the monetary unit remains unchanged.

The three demand theories are concerned with choice under certainty. Expected utility theory (EUT) however deals with understanding peoples' demand under uncertainty or risks (Hellinger, 1989; Moatti et al., 1995). The basic principles of EUT model the behaviour of a rational economic person faced with choice under uncertainty over which he, as the consumer is sovereign (McGuire et al., 1988).

Although, EUT addresses uncertainty in the demand for health care, there is a range of other characteristics that are not addressed, limiting its usefulness. For instance asymmetry of information means the consumer cannot judge what the effect of health care consumption on his health status will be, and thus decision making is done through an agent (McGuire et al, 1988). Hence, in terms of a spectrum which stretches from commodities where EUT fits very well to where it fits badly, health care is at the latter end (McGuire et al, 1988). Thus, the theories of demand under certainty are perhaps more useful for understanding consumers' valuations of health-care goods and services.

2.3 Consumer surplus

The gain to the consumer is traditionally measured by consumer surplus. This is a concept first developed by a French engineer, Jacques Dupuit, in 1844 in an attempt to quantify the benefits from a proposed bridge (Dinwiddy and Teal, 1996). As Dinwiddy and Teal noted, Alfred Marshall in 1920 defined consumer surplus as "the excess of the price, which the consumer would be willing to pay rather than go without the thing, over that which he actually pays". Demand curves based on revealed preferences may not reveal the exact value of the goods to the consumers, as each consumer has a surplus whenever the price of the good was less than the maximum amount the consumer is prepared to pay. This is because some of them would have been prepared to pay more than the price offered in the market.

Since the demand curve could be derived from both the Marshallian neo-classical theory and indifference curve analysis, the consumer surplus can be measured from either (Dinwiddy and Teal, 1996). The concept of consumer surplus is pivotal to the study of welfare economics and is most commonly used to analyse the welfare impacts of price changes (Brookshire et al., 1980). All things being equal, higher prices will result in a loss of welfare or "utility" to the consumer and vice versa if prices drop. (Dinwiddy and Teal, 1996).

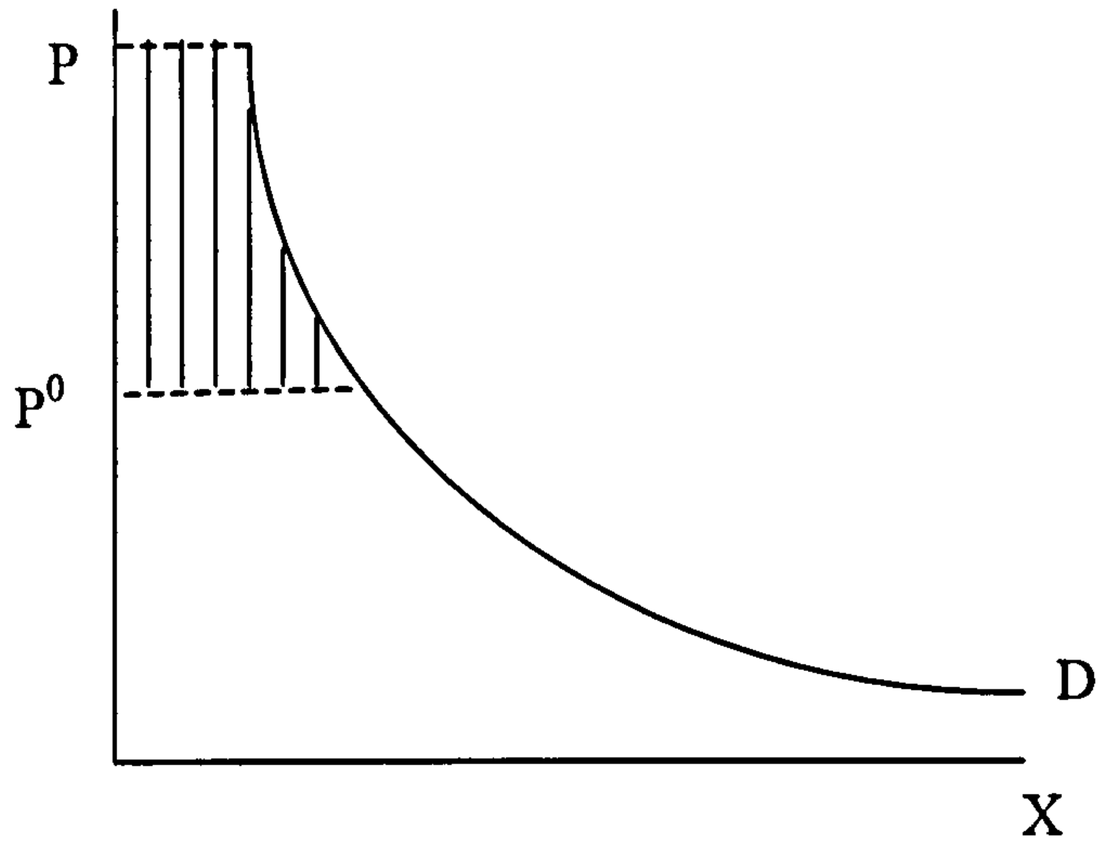
Consumer surplus is correctly measured by the Hicksian indifference curve, rather than through the Marshallian method (Mishan,1971, 1976; Brookshire et al.,1980; Donaldson, 1993). Marshall used the concept of willingness to pay in a static context[7] and the formulation was later refined by Kaldor (1939) and Hicks (1941;1943) to allow it to apply to contexts associated with change, in the price or quantity of a good (Donaldson, 1993).

Demand curves can also be used to describe the incremental (marginal) value consumers attach to additional consumption of health care given a level of consumption (Phelps et al., 1996). This is the willingness to pay interpretation of a demand curve. Rather than implying that the quantity demanded depends upon price, it involves determining the incremental value of consuming more health-care. This is equivalent to the consumer's willingness to pay for additional health-care.

Diagrammatically, consumer surplus is the area below the demand curve but above the market price. In Figure 2.2, P^0 is the market price and the shaded area from P to P^0 is the consumer surplus, thus for all quantities to the left of the demand curve, the consumers gain welfare (Dinwiddy and Teal, 1996). However, the total value to consumers of using a certain amount of medical care is the area under the demand curve (Phelps 1992). Total consumer surplus may be difficult to measure, but most policy applications are concerned with change in consumer surplus: with change in price and the value of the change a measure of the consumer's willingness to pay (Dinwiddy and Teal, 1996;).

[7] The value of a good represented by the maximum amount the consumer would be willing to pay for a given quantity of a good relative to a situation in which s/he was not able to buy the good at all - at a particular moment in time with everything else held constant.

Figure 2.2 Illustrating consumer surplus



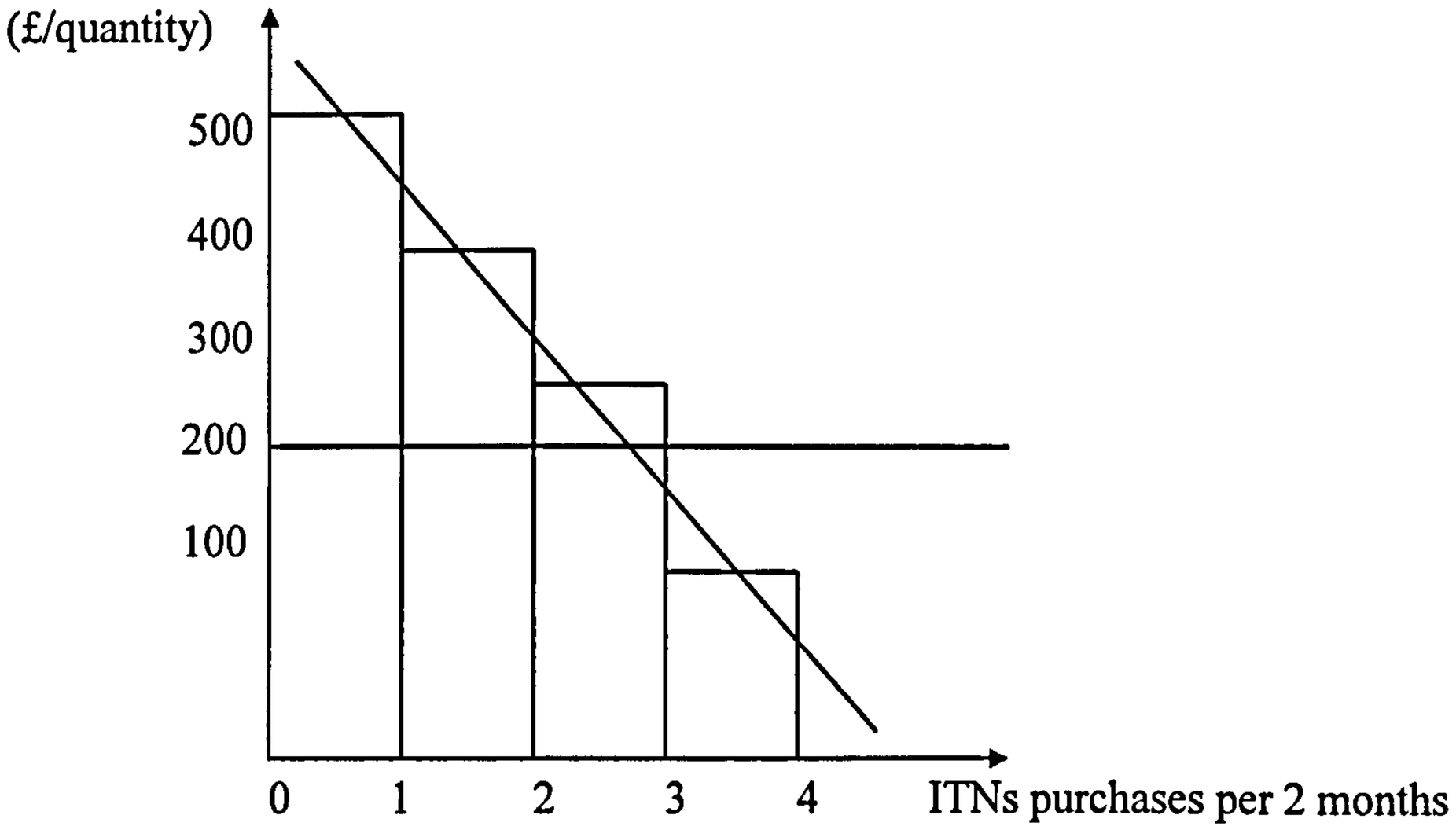
Illustrating consumer surplus using the demand for ITNs: Suppose the first ITN created £500 in value to the consumer (protection from mosquito bites and insects and prevention of malaria) (Table 2.1 and Figure 2.3). However, if the nets cost £200, the consumer would get £300 in consumer surplus out of that purchase. A second net purchased after 2 months at £200 might create a further £400 in value, and creating £200 in consumer surplus. A third purchase after 4 months would give a value of £300, creating £100 in consumer surplus, and a fourth net purchased after 6 months would create only £170 in value but would cost £200, and here the cost exceeds the value. The consumer would not buy at the fourth time, since it would cost more that it is valued.

Table 2.1 Illustrating consumer surplus for ITNs

Quantity purchased	Incremental value of ITNs	Net value of the purchase worth £200	Total net value (consumer surplus)
1	£500	£300	£300
2	£400	£200	£500
3	£300	£100	£600
4	£170	£-30	£570

From this discussion, a rational consumer will continue to expand the amount consumed until the marginal value received just equals the marginal cost of the services (Phelps, 1992). In the above example, equality was not achieved but the consumer would stop after the third purchase in order to maximise total consumer surplus.

Figure 2.3 Graph illustrating consumer surplus for purchase of ITNs



A smooth line through the mid-points of tops of the bars forms the demand curve, and adding up such curves across many individuals would smooth things out even more.

2.3.1 Theories and measurement of the consumer surplus

The Kaldor-Hicks (or Hicksian) measure of consumer surplus involves utility being held constant, because if actually asked to pay his or her maximum WTP, a person would pay no more than, and just up to, that amount which maintains her/him at her/his original level of utility (Donaldson, 1993). This is called 'holding real income constant'[8] as distinct from the Marshallian measure which holds money income constant. Using the Kaldor-Hicks (or Hicksian) measure, to keep real income constant, some of a person's money income is taken away (in case of WTP for a gain) or his/her money income added to (in case of receiving compensation for a loss) (Donaldson, 1993).

The effect of price changes in the Hicksian indifference curve analysis is further decomposed into the substitution and income effects to provide four measures of consumer's surplus called equivalent surplus, equivalent variation, compensating surplus and compensating variation. The surpluses differ from the variations in that the latter are calculated after the consumer has made optimising adjustments in his consumption set, while the former do not permit such adjustments (Brookshire et al., 1980).

The Hicksian compensating and equivalent measures differ with respect to the reference level of welfare. In Hicksian equivalent measures, the initial welfare level is different from the reference level, while in compensating measures, the initial situation is the reference level (Brookshire et al., 1980; Golan and Shechter, 1993).

The Hicksian compensating measure is defined as the amount of compensation, paid or received, which would keep the consumer at his initial welfare level if the change did take place (Brookshire et al, 1980). This compensating measure is operationalised as willingness to pay for the change to go ahead or willingness to accept compensation for the change to go ahead (Mitchell and Carson, 1989). Conversely, the equivalent measure is defined as the amount of compensation, paid or received, which would bring the consumer to his subsequent welfare level if the change did not take place.

Thus, when comparing two alternative levels of provision of a good, there are four relevant Hicksian value measures. These are WTP (compensating variation) to obtain the preferred level; WTP (equivalent variation) to avoid the less preferred level; WTA (compensating variation) to accept the less preferred level; and WTA (equivalent variation) to forgo a promised increment to the preferred level (Brookshire et al., 1980).

The Hicksian compensating measures of consumer surplus are consistent with the potential Pareto-improvement criterion, while Hicksian equivalent measures of consumer surplus are not (Brookshire et al., 1980). As Brookshire et al. Explained "the compensating measure, by using the initial level of welfare level as the reference level, measures the welfare impact of changes as if the individual had a right to his initial level of welfare (that is, as if he had the choice of keeping what he has or voluntarily trading for changes). The equivalent measure, by using the subsequent welfare level as the reference level, treats the individual as if he had only a right to his subsequent level of welfare (that is, as if he must accept his subsequent level, or seek to trade his way back to his initial situation)". In most cases, the perspective of health care policy is from the current (or original) level of utility and this requires estimates of compensating variation (Donaldson, 1993). Figure 2.4 illustrates comparisons between compensating and equivalent variation.

---

[8] Holding constant the amount of utility from the goods a person can purchase with his/her money income.

Figure 2.4 Graphs illustrating compensating and equivalent variations (Dinwiddy and Teal, 1996).

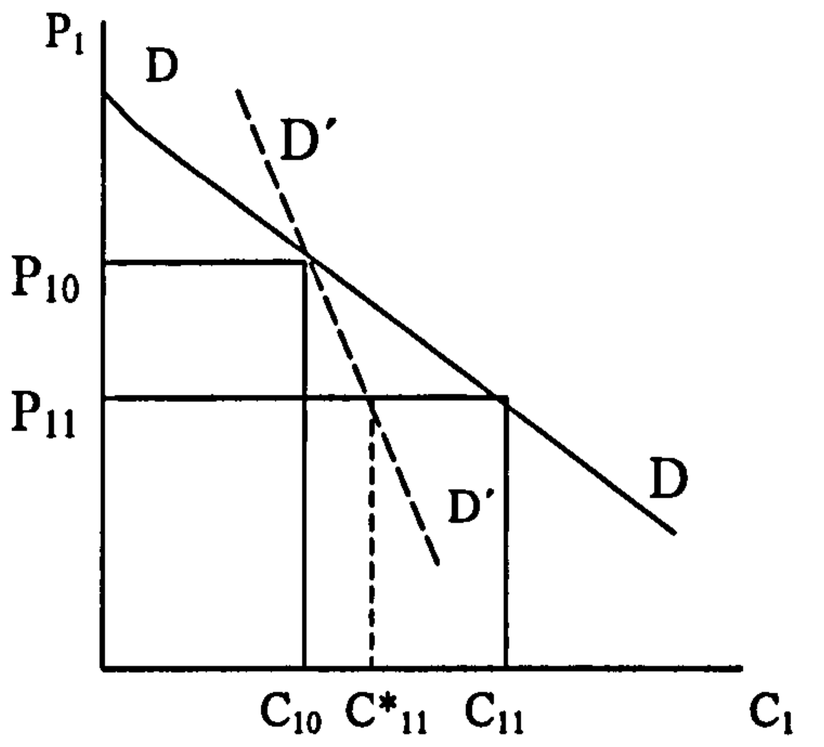


Figure 2.4a CV with the demand curve

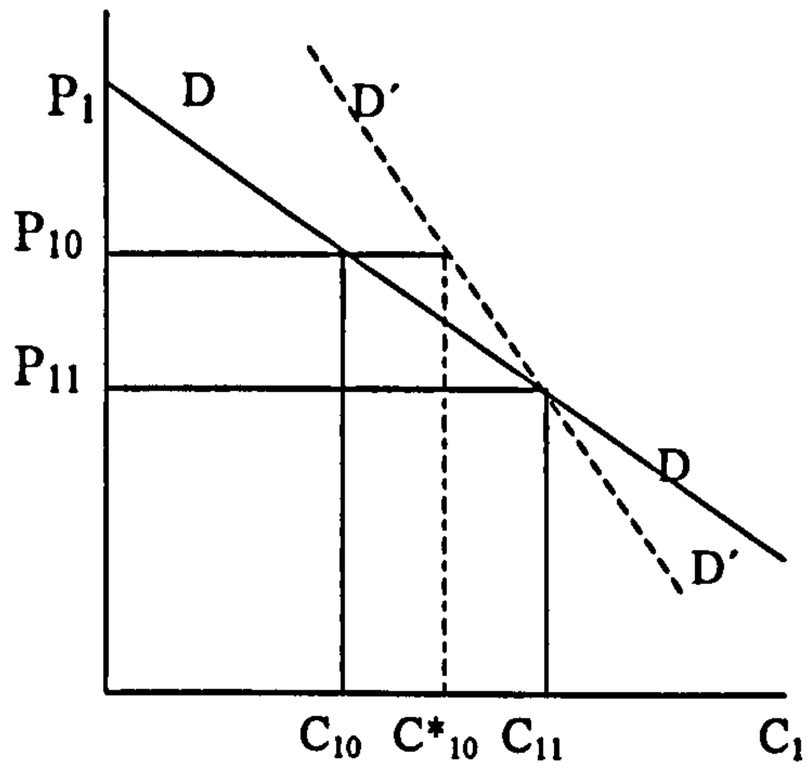


Figure 2.4b EV with the demand curve

In both Figures, the Marshallian demand curve is labelled DD, while, the Hicksian or compensated demand curves are represented by D'D' (Dinwiddy and Teal, 1996). For a price reduction from P_{10}, a consumer would naturally demand more of this good, C_1, substituting it for other goods whose prices have been held constant (Donaldson, 1993). While the Marshallian demand curve demonstrates the effect of a drop in price on the consumption of C_1 when both income and substitution effects are included, D'D' shows only the substitution effects of the price change at a constant level of utility (Dinwiddy and Teal 1996). The inclusion of both income and substitution effects accounts for the larger increase in demand demonstrated by the Marshallian curve in Figure 2.4a.

In Figure 2.4a, the move from C_{10} to C^*_{11} is the pure substitution effect and from C^*_{11} to C_{11} is the income effect. In Figure 2.4b, the consumer's income is hypothetically increased at the original set of prices until a higher level of utility is achieved. The increase from C^*_{10} to C_{11} is identified as the substitution effect. Thus, the area under the compensated demand curve between the two prices in Figure 2.4a is the compensating variation measure of surplus, while the corresponding area under the compensated demand curve between the two prices in Figure 2.4b is the equivalent variation measure of consumer surplus (Dinwiddy and Teal, 1996).

Operationally, compensating and equivalent variation can be measured using the willingness to pay (WTP) and willingness to accept (WTA) approaches. WTP is used to determine the maximum amount individuals would be willing to pay to acquire a good or service. On the other hand WTA asks the individuals to state the minimum amount they would have to be paid to willingly accept the loss or reduction of a good or service.

In the absence of income effects, WTP and WTA should be equal, but both have consistently been shown to differ. Estimates of WTP were always less than estimates of WTA for the same change that can hardly be explained by income effects (Johannesson, 1993; Donaldson, 1993, O'Brien et al., 1998; Gyldmark and Morrison, 2001). The practical explanation is the lack of perceived budget constraints on WTA (Mishan, 1988; Donaldson et al., 1997a). According to Donaldson (1993), other explanations for the WTP-WTA disparity are that the "WTP-WTA comparison exercises may be mis-specified, those in possession of a good, protest, at having to give it up (Gregory, 1986), some people protest at being bought-off (Rowe et al., 1980)."

Other causes of the WTP-WTA disparity from Gyldmark and Morrison (2001) include that respondents cannot give a precise WTA, but simply a "ball park" figure (Duborg et al, 1994). Another was loss aversion whereby people place a higher value on goods once they own them, hence their buying price is lower than their selling price (Kahneman and Tversky, 1979). People also require larger sums in compensation when there are few or no substitutes for the good in question (Hanemann, 1991). Hanemann also argued that the disparity arises from movements along different indifference curves or different segments of the same indifference curve leading to different marginal rates of substitution.

The WTP-WTA disparity has led to a recommendation from most researchers in the area to elicit WTP in contingent valuation studies (NOAA, 1993; Johannesson, 1993). In the discussion that follows, the focus will be mainly in the area of contingent valuation using WTP with compensating variation. WTP is the amount of money that, if paid by the consumer, leads to the same level of utility as would be the case without the service, but with the income available for other services (Phillips et al., 1997).

The WTP technique is increasingly being used in health care to inform policy decisions on consumers' valuation of health care goods, stemming from the failure of demand models to explain the demand for health care goods. The failure of the conventional theories of demand arises because of the peculiarities of health care and the inherent market failures in that sector. Hence, the next section presents a review of the demand for health care, particularly investigating the special characteristics of health care and some models that had been developed to better understand demand for health care.

2.4 Demand for health care

The demand for health care is derived from the demand for health. The demand function models the relationship between the quantity and consumer's willingness and ability to pay for a product over a given period of time. This demand is influenced by the needs (for the goods or services), expected utility on consumption, alternatives, tastes, quality and price of the good itself, complements and substitutes (Jacobs, 1997). The demand function may also be extended beyond the relationship of price and quantity to include the various possible income levels of consumers.

In the conventional demand theory consumers are assumed to be well informed about goods and services in a world of certainty. They are thus in the best position to make decisions on the consumption patterns that will best maximise their utility. However, demand for health care has many peculiarities that limit the use of conventional demand theories.

2.4.1 Special characteristics of health care goods and market failures

Health care has several characteristics that make it "different" from other goods (Culyer, 1991). These characteristics will prevent an efficient allocation in a competitive market (Cullis and West, 1991). The unique attributes of health-care include asymmetry of information, uncertainty, externalities, merit good[9] (Cullis and West, 1991), public goods[10] and questions about consumer rationality[11] (Culyer, 1991).

---

[9] The concept of merit goods introduced by Musgrove (1959), covers those private goods and services that are considered to be of such importance that they are financed and provided by the public sector in addition to the quantities produced and purchased in the market.

[10] They refer to goods and services that do not have rival and exclusive consumption patterns.

[11] Three arguments which are potential impediments to the optimisation of welfare in open markets for health care are:

a. Many consumers, though sick, do not desire treatment and may even be ignorant of their sickness. This violates a fundamental necessary condition for the attainment of an optimum through open markets.

These attributes limit the application of the conventional demand theories in analysing demand for health care services.

There are some consequences of these special attributes of health care. A consequence of externalities is the public intervention or provision of certain health care goods such as vector control activities and immunisations against communicable diseases. In these cases, private decisions do not take into account the social benefits or adverse effects. Asymmetry of information between consumers and providers can lead to the existence of supplier-induced demand (SID), while the presence of uncertainty has contributed to the development of health insurance markets.

In the case of asymmetry of information and uncertainty, the consumer may be ignorant about the current and future health states, as well as the range and effectiveness of treatments available (Mooney 1992). The consumer may also be uncertain about the cost of illness and whether s/he would have enough resources to cover the cost. The presence of these factors and the consequent dis-utility that may occur if wrong decisions are made create the need for an agency relationship in demand for health care. The consumers now depend on third parties to make help their consumption decisions. However, the supplier of health care in a market regime can increase the demand for health care by gaining information on the factors that influence the consumer's willingness to pay for relief from a particular disorder (Cullis and West, 1991).

Although consumers often transact in markets with incomplete information, the informational asymmetry in health care takes a rather special form in that suppliers know much more than most buyers about the characteristics of the commodities (Evans 1984). Physicians can, in many instances, provide information that will influence patients demand curve, but this information may be inaccurate (Jacobs, 1997). There is some inter-dependence of the consumers' and supplier's utility functions (Mooney, 1994).

---

b. The mentally sick fit oddly into a "consumers" sovereignty model.
c. Patients requiring emergency treatment are frequently not in a position to reveal their preferences.

The origins of SID can be traced to contributions by Shain and Roemer (1959) and Roemer (1961); with the postulation of "Roemer's law" which states that a "bed built is a bed filled" (Folland et al., 1997). This referred to the very close correlation between the availability of short-term general hospital beds and rates of utilisation. However, one alternative explanation is that the increase in utilisation may be a market response whereby patients in need of hospitalisation now have their demand satisfied. It was also argued that "increase in demand led to increase in beds – the post hoc (false cause) fallacy" (Folland et al., 1997).

Thus, the supplier becomes an agent for the consumer in demanding the good. The agency arrangement may then put the provider in a situation to exploit the consumer (McGuire et al, 1998). However, the problem in some cases may be that even when given the correct information by the provider, the consumer is unable to grasp the meaning of the information given and store it as knowledge.

Financing strategies like health insurance has been developed to limit the effect of uncertainty and adverse effects of supplier-induced demand. However, health insurance markets developed to counter the uncertainty in health care have also been dogged by their peculiar failures, with the two major ones being moral hazard and adverse selection. Moral hazard is controlled by the introduction of co-payments, while strict regulation could minimise the extent of adverse selection. Nevertheless, SID could also be present in health insurance markets depending on the peculiarities of that market.

These attributes of health care also apply to Insecticide-treated nets (ITNs) programmes and thus, there are potentially many causes of market failure in such programmes. One could be due to positive externality, where the use of ITNs by majority of people in a community can protect the few people that do not use the ITNs, the so-called mass effect. Another is potentially negative externality where mosquitoes are driven from the homes of those using ITNs to those without, leading to increased incidence of mosquito bites in the latter group. However, there is not conclusive evidence about these potential positive and negative externalities of ITNs.

Asymmetry or lack of information about the ITNs can lead to consumers making uninformed purchase and consumption of the ITNs that may not reflect their real choices if they had adequate information. Thus, people may not buy the nets if they wrongly perceive them to be harmful or not beneficial. Conversely, others may buy and use the nets for the wrong reasons and perceived benefits. These factors are compounded with uncertainty about the benefits or otherwise of ITNs.

In order to gain reliable and valid knowledge about the consumers' demand for ITNs and other health goods, a method that could limit the effect of the attributes of health care will be needed to provide a better valuation of consumer choice than the standard revealed preference methods. One such method is the contingent valuation method, where a scenario is the tool deployed to limit the effects of the special characteristics of health care.

The major causes of market failure in health care and for goods like Insecticide-treated nets that the contingent valuation method could address are uncertainty and asymmetry of information. This is because a well-specified contingent valuation scenario presented to the respondents would to a large extent reduce the informational gap they have about the good. The scenario also addresses some of the uncertainty in the peoples' demand since the benefits and adverse effects of consuming the good would be made explicit. However, the scenario may not adequately dispel uncertainty regarding the probabilistic nature of demand, as people may not be sure that they will, for example fall ill and need the health care intervention. Nonetheless, if the scenario gives the correct unambiguous information and the respondents believe in it, there should be elicitation of WTP values that reflect the respondents' demand function for the good.

However, externality could become a cause of market failure in future when the mass effect is proved or if many people have purchased ITNs, thereby decreasing the mosquito population and subsequently incidence of malaria. Hence, the WTP of the remaining people without ITNs could be lowered because they feel less risk of contracting malaria and having nuisance mosquitoes. However, an externality could become a cause of market failure in the future if a mass effect is proved or if many people have purchased ITNs, thereby decreasing the mosquito population and subsequently incidence of malaria. Hence, the WTP of the remaining people without ITNs could be lowered because they feel less risk of contracting malaria and having

nuisance mosquitoes. A brief description of health-care demand models will next be presented.

2.4.2 Models of health care demand

The peculiarities of health and health care have led to use of some specific models for the analysis of demand for health care services. Some of these models were developed for the analysis of demand for other types of goods and services outside health-care. They include utility-maximising models (Akin et al., 1985; Heller, 1982, Mwabu, 1986, Gertler et al., 1987), and household production models (Goldman and Grossman, 1978; Phelps and Newhouse, 1974; Mwabu et al., 1986, Akin et al., 1995). However, Grossman (1972) specifically developed the human capital models for analysis of demand for health-care.

The utility maximisation models are limited in explaining the total value of the good to the consumers, because they use health expenditures based on revealed preferences rather than consumers surplus to explain demand. Health care expenditures are a minimum amount or "lower bound estimate of the amount that the respondents are willing to pay for health care (Weaver et al., 1996). In other words, in some cases the respondents may have been willing to pay much more, but we only know that they were charged (Weaver et al., 1996)". Hence, the models will not determine the factors that explain the consumers' value for the good.

The contingent valuation technique (CVM) is used to operationalise the Kaldor-Hicks measures to determine the maximum value to the consumers of the goods and services. The CVM was developed in environmental economics and is now the most commonly used method of measuring environmental benefits (Johannesson et al., 1992). The use of CVM in the health sector is increasing (Diener et al., 1998; Klose, 1999; Smith, 2000). The CVM makes no restriction on what should be valued and is therefore a good model for determining the value and demand for health-care goods (Gafni, A; personal communication, 2001). Additionally, by modelling the socio-economic and other characteristics of the consumers using contingent valuation, the factors that affect demand can be determined.

2.5 Alternative methods for determining WTP

There are alternative methods to the CVM that could be used to determine peoples' valuation of goods and services. These alternative methods are conjoint analysis and price acceptance test, and both methods were developed as marketing tools for price-setting from market research. However, both methods have some disadvantages that limit their usefulness in valuing consumer surplus.

2.5.1 Conjoint analysis

While CVM techniques were initially developed in environmental economics to value environmental amenities, stated-preference methods like conjoint analysis evolved as market-research tools for evaluating consumer behaviour and predicting sales of new products (Johnson Reed et al., 1998). Conjoint analysis involves presenting individuals with hypothetical scenarios which contain different levels of the attributes of a resource and their preferences elicited (Ryan, 1993). By including a cost attribute, the monetary valuation of resources is achieved in conjoint analysis. However, conjoint analysis relies on indirect methods to determine the maximum WTP for different attributes of a resource and cannot be used to fully measure the consumer surplus. However, if the objective is to derive the relative weights of the arguments in the individual's utility function, then both methods are equally useful (Ryan, 1983). The formats for eliciting stated-preferences include the rated-pair, discrete-choice and ranking formats.

2.5.2 Price-acceptance test

There are many methods for determine the price people are willing to pay for goods in market research (Kracmar, 1971). These methods include method of price acceptance, experimentation with the proposed prices on a small scale to give a valid demonstration of how the price would work on the market and the evaluation of revealed preferences after the price and modifying the price in light of the observations. The price-acceptance method though a quick method does have many resemblance to CVM valuations in that a scenario is presented to the respondents before eliciting their WTP. The difference between the test and CVM is that the test can only be used for goods that are already in existence. The processes involved in the delivery and payment for the good are also not explained in the price-acceptance test, meaning the value elicited from the respondents may not be their valid consumer surplus. The description of the price-acceptance test as provided by Kracmar (1971) is summarised in the next two paragraphs.

In the price-acceptance test, the respondents should be well informed about the product, its features, and its advantages. Whenever possible, a sample of the product should be handed to each respondent during the test. In the case of large and bulky products, all features should be described; pictures of the inside and the outside, in colour, should be shown; and, if possible, a dummy model should be produced.

The respondents should be divided into two groups (groups I and II). Group I will have different prices indicated from the low end continuing upward, with the investigator naming prices and registering the number of respondents that consider each individual "price acceptance," or regard an individual price as "still reasonable," until reaching a "too expensive" response. The prices and the number of respondents, in percentages, per each price can be plotted on a diagram (the maximum price curve). Then prices are indicated to other group (group II), starting from the upper end and continuing downward, with the investigator again registering the number of respondents considering the price acceptable, until those who do not believe the product to be good enough at such low price are reached. The results are again presented graphically (the minimum price curve). When both curves are projected on a single graph, the price that appeals to the largest number of respondents will be revealed.

2.6 Conclusion

Revealed-preference information has proven to be an insufficient basis for obtaining policy-relevant valuations of goods and services. This is because although markets exist for some aspects of the health care, the special characteristics of health care and the existence of market failures, together with "market participants that are unrepresentative of the population of policy interest often obscure essential supply and demand relationships" (Johnson et al., 1998).

Although there are numerous alternative methods for placing money values on health improvements, the CVM is the preferred method. Willingness to pay determined from CVM has been used in the context of CBA and market research in the health care sector. The contingent valuation method is grounded in welfare economics, with the measurement of consumer surplus being the major theoretical underpinning. To empirically ascertain the Hicksian surplus or willingness to pay, the CVM is most often employed (Golan and Shechter, 1993).

CHAPTER 3: THE CONTINGENT VALUATION METHOD

This chapter reviews the methodological and operational issues concerning the contingent valuation method (CVM). Thus, it shows how the method could be used to elicit consumer preferences, its advantages and disadvantages, together with ways for determining whether estimates of WTP are reliable and valid.

The method is called contingent valuation because the respondent is being asked to consider the contingency of a market existing for the thing being valued. Thus, it involves respondents evaluating, in monetary terms, goods or services with benefits that may not be directly measurable (Morrison and Gyldmark, 1992). Where market failures exist or the prices charged for services do not reflect real market situations, the CVM technique through WTP can be used to determine the maximum value to the consumers of the goods and services. According to proponents of CVM, asking people directly has the potential to inform about the nature, depth, and economic significance of these values (Mathiyazhagan, 1998).

CVM is one of the most commonly used direct techniques for eliciting willingness to pay (Bala et al., 1998) and other direct methods include conjoint analysis and market research using price-acceptance tests. All direct methods use data gathered through a survey of individuals by asking respondents directly about how much they would be willing to pay, or willing to accept, for having, or not having, a particular attribute or good (Asgary and Willis, 1997). Indirect methods such as the hedonic price method and hedonic travel cost method in contrast, use data on observed behaviour to assess the value that individuals are willing to pay for a particular attribute (Asgary and Willis, 1997).

CVM has an advantage over many of its alternatives, since it is believed to be able to estimate a total economic value, rather than just components of the total value (Frykblom, 1997). This is because the CVM values the processes as well as all the dimensions of the outcomes produced by the good or service. The processes and outcomes in this case are the welfare gains from the use, non-use and altruistic dimensions of the valuation. CVM was first applied in determining the value of outdoor recreation in the United States (Davis, 1963).

The methodological issues in implementing CVM are first reviewed in order to provide a basis for understanding the operational aspects of WTP elicitation and in determining the reliability and validity of estimates of WTP. The question formats are then reviewed together with the underlying reasons for the use of different formats in different settings to achieve more reliable and valid estimates of WTP. The next section of the chapter reviews the biases associated with CVM, while the following two sections, detail the forms and tests of reliability and validity that are applicable in operationalising the CVM.

The insights gained from the review are used in the penultimate part of the chapter to develop the study framework, which hinges on strategies for improving the reliability and validity of WTP estimates from the CVM. The conclusions presented in the last section include implications for using the CVM in measuring WTP for ITNs and other goods in Nigeria.

3.1 Operationalising the CVM

There are three components of value that the CVM measures (Mitchell and Carson, 1989); (1) Use value: This is the valuation of the respondents' willingness to pay for the good because s/he will directly consume the good. (2) Non-use or existence or option value: This is the valuation of the respondent's willingness to pay for the good not that s/he will directly consume it at present, but wants to reassured that it exists in case s/he will need to consume it in the future. (3) Altruism: There is where the respondent is willing to pay for others to consume or benefit from the good or service in question.

The valuation of consumer surplus is the theoretical basis for these CVM values, because an individual might gain full value from consuming a resource just from the use value alone and or for non-use or altruistic reasons. The specification of a CVM scenario could be used to inform the respondents to value all three, or alternatively decomposed scenarios could also be used and the full value derived by combining WTP values. Decomposed scenarios could lead to higher valuations of WTP than single scenarios because the budget constraints may not be clear to the respondents or because of the possible bias called embedding effect (reviewed later in this chapter). However, most CVM applications have considered only the use value of the resource.

Three guidelines given by Mitchell and Carson (1989) to improve the reliability and validity of CVM are: the WTP questions must be clear and unambiguous; the respondents should be familiar with the commodity to be valued; and the respondents should have had prior valuation and choice experience with respect to consumption levels of the commodity, thus increasing the likelihood that they will have well formed values for the commodity.

Any CVM survey design involves three survey elements (Gonzalez-Caban and Loomis, 1997; Asgary and Willis, 1997). These are (1) portrayal of the resource to be valued; (2) description of the particular financial mechanism to be used to pay for the resource; and (3) the question format used to elicit the respondent's amount of WTP. The various methodological and operational issues associated with using CVM to determine WTP based on the three survey elements are presented in the following sub-sections.

3.1.1 Scenario

The scenario represents the key to a good contingent valuation survey and should provide details about the hypothetical market, including conditions for provision of the good, the frequency of payments and who will have access (Neumann and Johannesson, 1994). To enhance the reliability of a CV study, Mitchell and Carson (1989) recommended that the key scenario elements must be understandable, meaningful, and plausible. The resource to be valued is usually described in the scenario. It has been advocated that the more realistic the scenario is, the more likely that reliable and valid the contingent values are derived (Kealey et al., 1988; Reiling et al., 1990; Loomis et al., 1996). However, more familiarity with the good/services (which could be achieved by training the respondents) could equally increase the reliability and validity of the measurements (Kealey et al., 1988).

The detailed portrayal of the resource to be valued, the processes for the delivery and use of the resource, the benefits, the probabilities of gaining or losing from the consumption of the resource are provided in the scenario. In addition, the category of CVM, the payment vehicle and details about the delivery of the resource are provided in the scenario. The scenario should, if possible, include showing the respondents a sample of the commodity they are valuing.

The correct estimation of the consumer surplus from the CVM depends on how the scenario is constructed. Wording the scenario in such a way that makes it clear that all aspects of value namely, non-use, use and altruism are being valued (if the good/service has the three attributes) could simultaneously lead to determination of the full value of the good to the respondent. However, since some goods may have only one two of the CVM measures, it is also possible to determine the maximum value of each of the aspects of WTP, with the most commonly elicited being use values. For instance, a scenario for ITNs that portrays it as a private good would elicit the full consumer surplus for individual use of the net.

The scenario thus helps to limit the effect of asymmetry of information and uncertainty that would predispose to the elicitation of biased contingent valuation estimates. This is because the consumer is empowered with adequate information about the good/service, the processes and the outcomes in order to make an informed decision. In health care, information about the benefits, adverse effects and the potential cost implications of either consuming or not consuming the good/service will reduce the level of uncertainty the consumer may have.

3.1.2 Payment context

WTP can be measured in either a point-of-consumption or an insurance-based context (Gafni, 1991; Bala et al., 1998). These are also described as either the ex ante insurance perspective (where it is uncertain whether the event will occur) or the ex post perspective (where the event is certain to occur). In a point of consumption question, the respondents are asked what they would be willing to pay for treatment assuming they have the disease (Bala et al., 1998). In the insurance-based context, "given that they were at risk of the disease in the future, what is the maximum insurance premium they would be willing to pay to cover the treatment for the disease" (O'Brien et al., 1998).

Gafni (1991) recommended that the CVM question should be framed in the insurance context and used to determine ex-ante WTP in order to resemble how people pay for health care services. In insurance context, the probabilities of the different health states and outcomes from using the health good or service are presented to the respondents, who are then expected to make more informed valuations of their choices. Bala et al. (1998) argued that the insurance-based question is appropriate when the objective is to measure the societal WTP for a treatment, because people are been asked their WTP in

case they would need the treatment in future. However, those recommendations would apply where the payment for the health-care good under valuation is subject to an insurance cover. In the case of ITNs, where the fee-for-service mechanism, is the usual payment system used in a setting like in Nigeria, ex-post WTP would be the appropriate mechanism for framing the WTP payments.

3.1.3 Approaches to CVM surveys

Many different survey approaches have been used for eliciting WTP. They include; telephone, mail, interviewer-administered questionnaires, computer-based, group elicitation and self-administered questionnaires. NOAA (1993) recommended face-to-face interviews because it limits non-responses and also ensures that the hypothetical market is thoroughly described to the respondents.

It is also common to elicit WTP from those who already are, or could be, users of the service being evaluated (Olson and Donaldson, 1998). This is presumably because these people would give more valid and reliable estimates of WTP, as they would have better information and would pay more attention to the CVM scenario. Conversely, eliciting WTP from non-users or people who are not going to use the good or service could require a lot of information for them to understand what is being valued. Also, they may lack interest in the scenario and so would not give deep thoughts in making their choices; leading to the elicitation of less reliable and valid estimates of WTP from the non-users and potential non-users.

3.2 Contingent valuation question formats

The various question formats (or elicitation methods) for eliciting WTP currently include open-ended, the bidding game, payment cards (or categorical scales), dichotomous choice (binary, close-ended, take-it-or-leave-it) and the dichotomous with follow-through question formats. There are also variants of all these question formats. Smith (2000), Liljas and Blumenschein (2000), Klose (1999), Russell et al. (1995) and Mitchell and Carson (1989) described some of these formats in detail. These question formats would be discussed in detail next.

Each of the major CVM question formats will be reviewed in this section. Their advantages, disadvantages together with the special types of bias they might cause are described. The formats will be presented in the order of their period of introduction in CVM. Thus, the review starts with the open-ended technique and ends with the dichotomous with follow-up questions. The last part presents the arguments for realism in the choice of question format.

3.2.1 Open-ended technique

For this technique, respondents are asked their maximum WTP amount as an open-ended question. An example of its use was Donaldson et al. (1997b) in comparing the validity of the method and the payment scale (card) approach in eliciting WTP for bone mineral density screening. The open-ended question asked, after describing the resource to the respondents, was "What is the most you would pay to have the scan?"

An advantage of open ended question format is that it is easily amenable to comparison between stated with actual WTP because one could easily relate the maximum amount that was stated to the actual amount that the respondent paid. Also, if the good in question being valued is normally bought using an auction-like or donation-like question format, then it should be the appropriate question format. Finally, it does not require a large sample size for tests of construct validity, unlike the dichotomous choice method that requires a large sample size for the different bid vectors.

Concerns with open-ended questions centre around their face validity, in that they are hypothetical and do not reflect the way people behave in a real market (Ryan et al., 1999). It can be argued that in everyday life people are rarely asked to place monetary value on a good or service – they are given a price tag (Mitchell and Carson, 1989; NOAA, 1993). A further problem with this technique is that it is cognitively challenging for respondents to just come up with an amount for a good or service without guidance (O'Brien and Gafni, 1996). Thus, very wide WTP ranges will result. It also leads to large numbers of non-responses (O'Brien and Gafni, 1996). One may overcome this problem by guiding respondents by presenting them with possible answers (Klose, 1999).

3.2.2 Bidding game technique

The bidding game was developed by Davis (1963), and operates by the respondent being presented with an amount and being asked whether s/he would be willing to pay that amount. Depending on the response, s/he is bid up or down using a bidding iteration. Thus, if the respondent was willing to pay the initial amount, the amount is raised and the WTP question is repeated. If the respondent was not willing to pay the initial amount, the amount is lowered and the exercise continued. Respondents are bid up or down in increments or decrements until the designed number of iterations is reached. The elicited WTP amount could be continuous if the final question is open-ended. However, if the final question is binary, the maximum WTP amount is assumed to lie in the range between the last two bids.

O'Brien and Viramontes (1994) used the bidding game to determine the WTP for treatment of chronic lung disease. Their bidding iterations started at $10, $25, $50, $75 and $100 for different sub-samples so that they could test for starting-point bias. However, using $50 starting-point as illustrative of the bidding game, O'Brien and Viramontes first asked a respondent whether s/he was willing to pay $50. If the answer was yes, the amount was increased to $100 and if still yes, the respondent was then asked an open-ended question on his/her maximum WTP. However, if the respondent said no to the $100 bid, the amount was further lowered to $75 and the iteration stopped at this stage with the WTP being inferred from the interval between the highest and lowest stated WTP. However, if the original response to the starting-point of $50 was no, the amount was lowered to $10, either a yes or no response, would be used to infer the maximum WTP.

An advantage of the bidding game is that it requires only yes/no responses and thus has more market realism than single open-ended questions (O'Brien and Viramontes, 1994). Its other advantages are that it requires small sample sizes and also guides the respondents using the bid amounts and stimulates thought processes as the exercise continues. It also resembles, to an extent, the usual price-taking behaviour of Nigerians.

Its disadvantage is that it may introduce starting-point bias, i.e. maximum WTP can be influenced by the first bid (Klose, 1999). The problem of starting-point bias can be avoided if the respondent is asked directly for the maximum WTP (i.e. open-ended question) without payment cards or a bidding game (Johannesson, 1996). However,

Johannesson posited that such a question is perceived as very difficult to answer, as it is unrealistic and this leads to problems of non-response. One means of overcoming the problem, if it exists, is to use different starting bids for different sub-samples, although the sample size for the study, as a result, would need to increase.

3.2.3 Payment card (scale) technique

The payment card method was developed by Mitchell and Carson (1981;1984) to help overcome starting point bias (Ryan et al., 1999). Here, cards with different amounts or a card with different amounts, are/is shown to the respondent, following the presentation of the scenario. Respondents then indicate their maximum WTP out of the range of given amounts. This can be done either directly by marking the amount closest to their WTP or by indicating the maximum amount they would pay and the minimum amount they would not (Klose, 1999). Thus, in the first case, an open-ended WTP amount is elicited, while in the latter case, the interval between the maximum and minimum WTP is taken as the WTP. The respondent is not bound to choose any of the amounts on the payment scale and may volunteer a figure outside the list depending on the study design.

The payment scale format has been used by Donaldson et al. (1997b) in comparing its validity with that of the open-ended approach. The payment scale question posed, after describing the resource, was "What is the most you would pay to have the scan? Put a tick to the amounts that you are sure you would pay and the options given were £0, £5, £10, £15 and £20."

The advantage of the payment card format is that it could be used to start a bidding game in order to eliminate starting-point bias. Also, it could be used as a dichotomous choice technique whereby the respondents indicate whether they are willing to pay each of the amounts presented.

However, payment cards may lead to 'range bias', i.e. the maximum WTP could be influenced by the range of presented amounts (Klose, 1999) or 'mid-point bias', where respondents tend to choose the middle amounts presented as their maximum WTP (Ryan et al., 1999). Another disadvantage of the technique is that it could be mentally challenging for respondents to pick out an amount from an array of prices.

3.2.4 Dichotomous choice format

Hanemann (1984) developed the conceptual and theoretical arguments for using this question format to estimate welfare benefits (Ryan et al., 1999). In this technique, a price for the product under valuation is presented to the respondent, who is expected to either accept that he would pay the price or not. The elicited WTP amounts are binary yes or no answers to the different bids. By varying the bid across sub-samples, it is possible to calculate the proportion of people who are willing to pay as a function of the bid (Smith, 2000).

An example of the use of this question format was Ryan et al. (1999) in comparing the validity of the method with that of the payment card method, in determining WTP for the treatment of cancer. The question Ryan et al., asked after presenting the scenario, was "Would your household be willing to contribute £100 each year for this expansion in the number of cancer treatments?" 1 = Yes; 2 = No; 3 = Don't know. This question was presented to different sub-samples with different 20 amounts ranging from £2.50 to £450.

An advantage of this question format is that it could easily be used to plot a demand curve (Smith 2000). This method has been said to be "closest to real market decisions of individuals" (Klose, 1999). Other advantages are that no one respondent can unduly influence the overall WTP (since different amounts are presented to different sub-samples) and it avoids both starting-point and range bias, since each respondent is presented with only one bid (Smith, 2000).

A major disadvantage with the question format is that it statistically inefficient, requiring substantially larger samples for the same level of precision when compared with the open-ended technique (Loomis et al, 1997). This is because each respondent provides only a limited information on whether they accept or reject a bid and hence, the survey requires different sub-samples to be offered different bids. (Smith, 2000). An additional problem is how to design the bid vectors (proportion of the different amounts that would be presented to specific proportion of the respondents). This is because with the dichotomous choice format, the estimated mean is sensitive to bid design (Cooper and Loomis, 1992). Also, sophisticated statistical analyses are needed to determine the mean and median WTP, though the non-parametric methods introduced by Kristrom (1990) has helped to simplify some of these analyses.

The dichotomous choice format also introduces a peculiar type of bias called "yea saying", whereby the posited bid is accepted as a cue for what is a reasonable payment (Kanninen 1995; Mitchell and Carson 1989). Furthermore, the dichotomous choice format results in symbolic votes for the good in question, not because the respondent would pay the posited price for it, but rather to register their support for providing the public programme. It thus leads to an overestimating of WTP (Brown et al., 1996). It was also argued that the dichotomous choice format could introduce starting-point bias since the respondent is confronted with an amount of money and asked whether or not they would be willing to pay (Ryan et al., 1999). The problems with the dichotomous choice format led to the development of dichotomous choice with-follow-up technique.

3.2.5 Dichotomous choice with-follow up technique

Smith (2000) describes this question format as a truncation of the bidding game. There could be single or multiple follow-up questions. In the single follow-up, the single-bounded dichotomous choice question is followed by a second dichotomous choice question (a form of bidding truncated at two bids) and even additional follow-up questions after the first follow-up question (Smith, 2000). The last follow-up question could be either open-ended or closed-ended.

Onwujekwe (2001) used the dichotomous choice with a single open-ended follow-up question to determine WTP for family-size insecticide-treated nets. The question posed was "Are you willing to pay 400 Naira for a net?" The yes or no answer was followed up by asking what is the maximum amount that you are willing to pay.

The advantage of this question format is that the follow-up questions reduce the range within which the respondents' true WTPs lie and thus reduces the need for large sample sizes (Cameron and Quiggin, 1994; Smith, 2000). Its disadvantage is that it might lead to starting-point bias (Herriges and Shogren, 1996).

3.2.6 Mixed techniques

Many contingent valuation studies combine these question formats in order to elicit more reliable and valid WTP. For instance, to eliminate starting-point bias, some studies first ask an open-ended question and then use the resulting answer as the starting-point. The payment card is also used to help develop appropriate bid vectors for the dichotomous choice technique.

3.2.7 What should determine the choice of question format in different settings for more reliable and valid CVM estimates?

There is still no consensus on the best question format to use in the health sector (Onwujekwe, 2001). All CVM question formats should ideally yield similar results, as there is no conceptual basis for expecting them to differ (O'Brien and Gafni, 1996). However, CVM questions should be constructed so that respondents view their answers as accurate representations of how they would behave if confronted with an actual market for the good (Mitchell and Carson, 1989). Thus, an argument is to use question formats that mimic price-taking the respondents are used to. This use of realistic question formats should increase the validity and reliability of the CVM.

It could thus be argued that different question formats would perform differently in different areas and contexts depending on the extent to which they mimic price-taking behaviour. Those that better mimic price-taking in any particular context would generate more valid WTP estimates than others that do not.

Lending weight to the necessity for question formats to mimic price-taking was the criticisms of the open-ended question format (NOAA 1993, Mitchell and Carson, 1989). The NOAA panel felt that the use of the open-ended format does not lead to reliable and valid WTP measures, because it lacks realism, since in everyday life people are rarely asked to place monetary value on a good or service – they are given a price tag.

Smith (2000) in arguing in favour of realistic question formats stated that the dichotomous choice technique more closely resembles consumer choice in actual market situation where price is given, which contributes to greater realism and respondent understanding and therefore yields more valid WTP values and higher response rates.

Supporting the argument for realism in CV elicitation is the stand taken by Gyldmark and Morrison (2001) that in order to limit bias in CVM, proper understanding of the scenario and the questions is best obtained by making the questions as relevant and familiar to the respondents as possible.

It has also been argued that the trade-off in any contingent valuation survey is between making the market sufficiently realistic to elicit a "true" valuation and not making it so realistic as to provide incentives for strategic behaviour (Smith, 2000). The dichotomous choice approach will meet this trade-off more effectively than will other approaches because the market situation is more realistic. The non dichotomous choice approaches present the respondent with a highly unusual and unfamiliar valuation mechanism, causing an inability to respond in a valid way (Smith, 2000; Jones-Lee, 1990).

Based on the criterion of mimicking price taking behaviour, the NOAA expert panel recommended the dichotomous choice approach for eliciting the non-use value of environmental resources, "because it resembles actual market decisions and real world referenda" (NOAA, 1993). The dichotomous choice system now dominates the CV of non-market goods and services (Herriges and Shogren; 1996).

However, there are also many contexts where the close-ended probably does not mimic market decisions, and so the blanket use of the technique may actually run against the principle of using an question format that mimic normal price-taking in order to improve the validity of WTP responses. Also, studies that have compared stated and actual WTP found that actual WTP was roughly the same regardless of which question format was used (Brown et al, 1996; and Loomis et al., 1997). The inference from these studies could be that the question format does not really matter.

3.3 Biases associated with the contingent valuation method

To enhance the validity and reliability of a CV study, the potential biases in valuations must be controlled for (Golan and Shechter, 1993). Mitchell and Carson (1989) identified three principal sources of these biases in CV studies. One is that the scenario contains strong incentives for respondents to misrepresent their true willingness to pay amounts. Another is that the scenario contains implied value cues that help determine WTP amounts, and the third is there is mis-specification or misperception of the scenario.

The two major forms of biases are hypothetical and strategic bias (Mitchell and Carson, 1989). Hypothetical bias arises due to the hypothetical nature of the market, and thus respondents may unwittingly not give correct responses. In relation to strategic bias, the respondents, due to their vested interests, set out to deliberately present valuations that do not reflect their true preferences.

Hypothetical bias may arise for two reasons (Whittington et al., 1990). First, the respondent may not understand or correctly perceive the characteristics of the good being described by the interviewer. Second, the individual may not take contingent valuation questions seriously and will simply respond by giving whatever answer occurs first to them (Whittington et al., 1990). Where this type of hypothetical bias is prevalent, bids will presumably be randomly distributed and not be systematically related to household characteristics and other factors suggested by economic theory (Whittington et al., 1990).

However, Cummings et al. (1995) argued that if a subject perceives that his expected utility is affected by the possibility of the good actually being provided, he has no incentive to misrepresent. That is, responses obtained with the hypothetical method would not differ substantially from those obtained from real WTP, but such a presumption cannot be made regarding the hypothetical method (Cummings et al., 1995).

The existence of biases in contingent valuation, their possible causes, together with solutions have been investigated in many studies; Liljas and Blumenschein (2000) provide a detailed review of those studies. While some studies compared the validity and reliability of different question formats some investigated the inherent biases that exist with particular question formats.

Many of the biases that arise in CVM could be related to the market failure in the health care. In health care, the major bias would be hypothetical bias because the information provided is either inadequate or is not understood by the consumer. Some people may also behave strategically because they are uncertain of the benefits and other issues with the good in question; its provision and payment vehicle.

The types of hypothetical bias include payment vehicle bias, scope effect, question order effect, embedding effect, yea-saying and warm glow. On the other hand strategic biases include free-riding and overestimation. Starting-point bias could either be classified as hypothetical, strategic or not a bias depending on the context of its occurrence (Onwujekwe and Nwagbo, 2001). These biases will be explained fully and the link established between them and issues of reliability and validity.

3.3.1 Hypothetical bias

Providing inadequate information and a mis-specified scenario are leading causes of hypothetical bias. Use of the wrong question format, especially one that does not mimic how respondents normally purchased goods and services, can also lead to this bias since the respondents will view the entire exercise as being hypothetical. The various forms of hypothetical bias are described in the following sections.

Payment vehicle bias

If the mode of payment in the CVM scenario affects one's stated valuation, there is a payment vehicle bias (Mitchell and Carson, 1989). If respondents are averse to any of the methods of financing health care, then one might expect their WTP for scenarios financed through such means to be reduced (Ryan, 1993). Payment vehicle bias can be potentially reduced, by using payment methods that resemble how respondents usually paid for the goods and services or their close substitutes. Pre-testing of the scenario could be used to ensure that the appropriate payment vehicle is used.

Scope effect

This refers to whether the WTP is sensitive to the size of the good/service being valued. If for instance, respondents value the probability of a more successful treatment for an illness similarly with the probability from a less successful treatment, it means that the valuation was not sensitive to the scope of the resource. Insensitivity to scope invalidates WTP valuation, and question formats that fail to detect scope effects are sub-optimal. Kartman et al. (1996) detected the existence of scope effects, but Kartman et al (1997) did not. A possible solution to scope effect is the eliciting of values from only the current or potential consumers of the good/service that will more critically evaluate the scenario and give more valid and reliable responses. Pre-testing of the scenario can also signal whether scope effects will be present and inform the wording the scenario.

Question order effect

This refers to how the ordering of question affects the WTP valuation. It is commonly a problem where multiple scenarios are presented or many goods are being valued. Here, a high value may be given to the first good on the list and less to subsequent others, although the latter goods may be worth more. This could arise from the respondents attaching a higher value to the first good and then realising their budget constraints, attach lower values to subsequent goods. Changing the order the goods are presented to different sub-samples of the respondents and observing whether this has any effect on the valuation is a way of testing whether question order effect is present. Better specification of the scenarios and instructions to the respondents are possible means of limiting the occurrence of this type of hypothetical bias.

Embedding effect

Embedding refers to the solicitation of willingness to pay for a good that is valued as a component of a larger good (Brown et al., 1995). For instance, people may state lower WTP for the re-treatment of ITNs if valued as part of a broad ITNs programme offering both ITNs and re-treatment services. Embedding significantly lowers the amount that respondents say they would pay for it compared with a non-embedded valuation. Kahneman and Knetsch (1992) were the first to demonstrate the existence of the embedding effect and actually introduced the terminology. It has been recommended that better information and description of the goods will eliminate embedding effect. Brown et al. however concluded that the amount of information had little impact on the effect of embedding on WTP. Thus, the method for decreasing this type of bias is an unresolved research topic.

Warm glow

This refers to a situation where the respondents derive an inner happiness by stating a WTP amount that does not reflect their real WTP because of the hypothetical nature of the survey. This is common in valuing public goods where the feeling is that the higher the stated WTP, the higher the support of the provision of the service. This has also been termed the purchase of moral satisfaction by Kahneman and Knetsch (1992). Kahneman and Knetsch stated that contingent valuation responses (in this instance) reflect the willingness to pay for the moral satisfaction of contributing to public goods, not the economic value of these goods. Warm glow can be reduced by better

introduction of the purpose of the survey to the respondents and impressing on them the necessity to state the value of the good to them.

Time-to-think effect (Stability of responses)

Whittington et al. (1991) argued that though most contingent valuation surveys assume that respondents can provide rapid answers to valuation questions, respondents however, may need time to think about valuation questions, before they can give reasoned answers that take full account of their budget constraint. They also noted that respondents may need the opportunity to consult with other family members before reaching a decision that is binding for the household unit. They found that giving respondents time-to-think improved the validity of the contingent valuation responses.

3.3.2 Strategic bias

Strategic bias may arise when a respondent thinks he may influence an investment or policy decision, by not answering the interviewer's question truthfully (Whittington et al, 1990). Thus, this bias results from strategic behaviour of the respondents. This may lead CV responses to overestimate WTP because respondents do not face the final consequences of actually paying the hypothetical WTP and so do not consider their budget constraints in stating their WTP (NOAA, 1993). Conversely, in relation to free-riding, the respondent feels that the good will be provided even s/he does not pay or just pays a little. This will lead to low WTP amounts. Though, it is difficult to counter this, some strategies could be used to reduce the occurrence of strategic bias. They include, better introduction of the reasons for the CVM survey, using community leaders (in a Nigerian setting) to inform people to be truthful in responding to the questions and informing people (in the scenario) that they might be required to pay the amounts they have stated.

3.3.3 Starting-point bias

Starting-point bias means that the respondents are influenced by the amount used to start the bidding and may think that it represents the value of the good or service in question. Thus, the final WTP amounts they give will cluster around the starting bid. If the respondent's bid reflect his or her "true" value of the good or service, then it should not matter what initial amount (or starting point) the enumerator uses to begin the bidding game (Whittington et al, 1992). The problem of starting-point bias can be avoided if the respondent is asked directly for the maximum WTP without payment

cards or a bidding game (Johannesson, 1996). However this can be very difficult to answer (Johannesson, 1996), and may lead to non-response. Nonetheless, in real markets with differential pricing for the same good, people may still buy the goods in places where the prices are highest and this is not a reflection of any bias. Thus, it is uncertain whether starting point bias is really a bias as such, rather than a reflection of actual phenomena that occur in real market places with differential pricing for the same good.

3.3.4 Link between CVM biases with reliability and validity tests

A contingent valuation study is reliable if unsystematic bias is eliminated and valid if systematic bias is eliminated (Mitchell and Carson, 1989). Hence, the assessment of reliability and validity of WTP estimates are recommended for contingent valuation studies so that the credibility of the valuations are determined (NOAA, 1993) and ways of improving reliability and validity known. Table 3.1 matches the different tests of reliability and validity with the types of biases they can detect, method of testing for reliability and validity and reducing the biases.

Table 3.1 Validity and reliability tests versus the CVM biases they could detect

Validity	Type of bias	Method for testing	Method of reducing bias
Content	*Hypothetical *Strategic bias	Pre-testing of scenario and question formats	*Use context-specific questioning *Informative scenarios *Well trained interviewers
Construct	*Hypothetical *Strategic	*Econometric modelling *Other parametric tests *Non-parametric tests	*Use context-specific questioning *Informative scenarios *Well trained interviewers *Careful selection of bid vectors *Careful selection of starting-points *Calibration of responses
Criterion	*Hypothetical *Strategic	*Divergence between stated and actual WTP	*Better surveys *Better questioning *Calibration of responses
Test-retest and inter-rater reliability	*Strategic *Stability of responses	*Correlation coefficient *Non-parametric tests *Econometric modelling	*Use context-specific questioning *Informative scenarios *Well trained interviewers *Careful data management

Table 3.1 shows that it may be possible to generally detect different types of biases associated with the CVM using various tests of reliability and validity. However, unless aided by further qualitative studies like debriefing the respondents, it may be difficult for the tests to distinguish between different forms of bias. Nonetheless, if studies are established for the detection of specific biases, for instance scope effect or free-riding, then it is possible to determine the extent the tests detected the biases. The different forms and tests of reliability and validity tests could be used to determine the presence of the various CVM hypothetical and strategic biases are reviewed in section 3.4 (Reliability) and 3.5 (Validity).

3.4 Forms and tests of reliability

Reliability concerns the extent to which an experiment, test, or any measuring procedure yields the same results on repeated trials (Carmines and Zeller, 1979). A reliable measure is one that is relatively free from "measurement error"(Fink, 1993) and also "produces the same results, particularly when applied to the same subjects at different time periods" (Bowling, 1991).

Four kinds of reliability are usually distinguished: test-retest reliability (stability), alternative-form reliability (equivalence), internal-consistency (homogeneity) and inter- and intra-rater reliability (Fink, 1993). Bowling (1991) adds multiple form and split-half reliability. In multiple form reliability, the correlation of two tests that measure the same attributed is determined, whereas split-half reliability is used to assess the internal consistency where items on an instrument are divided into two equivalent parts and correlation between the scores on each part are computed (Bowling, 1991).

There is no guideline on the acceptable criteria for assessing the reliability of tests. However, Munro (1997) classified the strengths of the correlation coefficient (r) as: little, if any (0 – 25); low (0.26 - 0.49); moderate (0.50 - 0.69); high (0.70 - 0.89); and very high (0.90 - 1.00). Furthermore, reliability tests using tests of means should demonstrate that the different data sets are not statistically significantly different from each other in order to assure the reliability of the tests.

The common forms of reliability together with the types of bias they could detect are described in more detail in the following sections. However in Table 3.1, there is a presentation of the types of bias that could be detected by the two forms of reliability investigated in this thesis together with methods for testing the reliability and reducing the biases.

3.4.1 Test-retest reliability

Here the same test is given to the same people after a period of time (Carmines and Zeller, 1979). The results are compared, usually by correlation (Bowling, 1991) and the retest reliability coefficient will be 1.00 if one obtains exactly the same results on the two administrations of the test (Carmines and Zeller, 1979). However, the correlation across time will be less than perfect because of many reasons including the instability of measures taken at multiple points in time (Carmines and Zeller, 1979), intra and inter observer variability.

A major conceptual difficulty in establishing test-retest reliability is in determining how much time should lapse between the first and second administrations (Fink, 1993). If too much time elapses, external events might influence responses for the second administration; if too little time passes, the respondents may remember and simply repeat their answers from the first administration (Fink, 1993). There is still no consensus on the time lapse in test-retest reliability for the determination of reliable estimates.

Parametric and non-parametric comparisons of the means and medians of elicited WTP in the two time periods are also used to determine test-retest reliability, with statistical indifference between the measures at the two time periods signifying that the estimates are reliable. Some people also estimate identical econometric models in the different time period and determining whether the two models statistically significantly differed as a test of reliability (Kealey et al., 1988; and Carson et al., 1997). Test-retest reliability is the commonest form used in CVM applications.

Test-retest reliability could be arguably used to determine whether strategic bias (free-riding and overestimation) and stability of WTP responses (time-to-think effect) are present in a valuation. This is because, the CVM questions are administered at different time periods, people that were not honest in their previous valuations could have forgotten their answer and thus give a different response. However, genuine changes could have taken place in the intervening period to cause the variation in WTP and these changes need to be controlled for before one could definitely conclude that bias was either present or absent.

3.4.2 Alternative-form method

This refers to the extent to which two assessments measure the same level of difficulty (Fink, 1993). It is similar to the test-retest method in that it requires two testing situations with the same people (Carmines and Zeller, 1979). However, as Carmines and Zeller noted, it differs from the retest method in that the same test is not given on the second testing but an alternative form of the same test is administered. This test is not commonly used in WTP applications.

3.4.3 Internal consistency method

Internal consistency method refers to the extent to which all the items or questions assess the same skill, characteristic, or quality (Fink, 1993). They do not require either the splitting or repeating of items (Carmines and Zeller, 1979). They require only a single test and provide a unique estimate of reliability for the given test administration. According to Carmines and Zeller (1979), the most popular of these reliability estimates is given by Cronbach's coefficient alpha. It is basically the average of all the correlation between each item and the total score (Fink, 1993). This test is not commonly used in WTP applications.

3.4.4 Inter-rater reliability

This refers to the extent to which the data generated by two or more individuals match (Fink, 1993). It is operationalised by having different interviewers to administer the test on exactly the same respondents. Ideally, both interviewers should administer the test on the same day. Inter-rater reliability is enhanced by training data collectors, providing them with a guide for recording their observations, monitoring the quality of data collection over time to see that people are not "burning out", and offering a chance to discuss difficult issues or problems (Fink, 1993). Inter-rater reliability could be used to

test for the presence of all forms of strategic bias and stability of responses, due to the reasons argued in case of test-retest reliability.

3.4.5 Intra-rater reliability

In this test, the same subject is rated by the same observer at different times, usually on the same day or within a few days, but not as long as in test-retest reliability. It may pose reliability problems because "each observer may apply slightly different standards from day to day" (Streiner and Norman, 1995). The resulting reliability is called an intra-observer reliability coefficient, since it measures variation that occurs within an observer as a result of multiple exposures to the same stimulus (Streiner and Norman, 1995). Intra-rater reliability can be enhanced by training, monitoring and continuous education. Intra-rater reliability could be used to test for the presence of all forms of strategic bias and stability of responses.

3.5 Forms and tests of validity

The validity of measurement refers to the degree to which a measurement measures what it intends to measure (Klose, 1999). Validity tests are divided into content, criterion and construct validity. Construct validation is an ongoing process, of learning more about the construct, making new predictions, and then testing them, as opposed to content and criterion validity that can often be established with one or two results (Streiner and Norman, 1995). Criterion validity, where actual WTP is compared with stated WTP is one that carries much weight among economists (Loomis et al, 1997). Criterion validity is the ideal test for the validity of WTP responses (Gyldmark and Morrison, 2001). Table 3.1 matches the three different measures of validity investigated in this thesis with the types of CVM biases they can assess, together with the methods of testing and reducing the biases.

3.5.1 Content validity

This relates to the content of the survey instrument and associated materials presented to respondents (Asgary and Willis, 1997). Thus, it may refer to whether the contingent valuation scenario represents the object of measurement adequately (Klose, 1999) and the extent to which a measure thoroughly and appropriately assesses the skills or characteristics it is intended to measure (Fink, 1993).

Carmines and Zeller (1979) argued that content validity is not so much a specific type of validity as it is a goal to be achieved in order to obtain valid measurement of any type – namely, that the empirical measure covers the domain of content of the theoretical concept. Additional methods for ensuring content validity are determining that the components of the scale/item cover all aspects of the attribute to be measured and that the content of the variable matches the name which it has been given (Bowling, 1991).

An important aspect of content validity in CVM is that the hypothetical contingent valuation scenario presents a valuation task and choices that are realistic, e.g. by choosing payment vehicles that are usual in the health care system (O'Brien and Gafni, 1996). This latter point could be used to argue for the necessity of using WTP question formats that resemble the bargaining procedures for buying similar goods under valuation in the market. However, "because of the complexity of establishing content validity in certain contexts, the literature is often consulted either for a model or for a conceptual framework" (Fink, 1993).

There is no universal criterion for determining that a measure is content-valid (Carmines and Zeller, 1979). Thus, content validity rests mainly on appeals to reason regarding the adequacy with which important content has been sampled and on the adequacy with which the content has been cast in the form of test items (Nunnally, 1978). This limitation has possibly prevented content validation from becoming fully accepted for assessing the validity of social science measures (Carmines and Zeller, 1979).

Various research methods could be used to either test or improve the content validity of different CVM question formats. Qualitative and quantitative research methods could be used and elements of pre and pilot testing of the survey instruments are usually used to ensure that the study tools are content-valid.

Content validity could possibly be used to inform about the potential for free-riding and over-estimation of WTP. It could also be used to test whether hypothetical bias is present, especially with respect to the scenario elements and the question format. This is because prior qualitative and or quantitative studies with a sub-group of the respondents and an understanding of the reasons for their answers by thorough debriefing could provide the evidence on whether these biases would be present. Methods of increasing

Thus, such analyses have two main purposes (Whittington et al., 1991). First, if the WTP bids are correlated with the variables suggested by economic theory, this increases our confidence that the WTP bids indeed reveal information about the respondents' preferences and are not simply random numbers. Second, models of the determinants of respondents' WTP bids can be used to predict how changes in socio-economic characteristics will affect the demand for the good or service offered (Whittington et al., 1990).

Convergent validity

Convergent validity measures how well WTP is correlated with other measures that are intended to measure the same object, such as QALYs for preferences over health (Klose, 1999), or the time trade-off technique or standard gamble (O'Brien and Viramontes, 1994). Additionally, a determination of whether two different CVM question formats yield similar estimates is a test of convergent validity (Desvousges et al., 1993).

Another approach to judge how reasonable WTP estimates are, could be to compare them to cost of illness estimates (Whittington et al., 1996). The expected illness costs, in theory, should serve as a lower bound to the individual's willingness to pay (Harrington and Portney, 1987; Berger et al., 1987) for two reasons (Whittington et al., 1996): First, the cost of illness approach ignores the value of pain and discomfort suffered by the individual, as well as any resources spent to prevent illness; Second, the expected cost of illness assumes that the individual is risk-neutral. In a specific disease example, risk-averse individuals would be willing to pay more for the malaria vaccine than risk neutral individuals (Whittington et al., 1996).

Establishing Construct Validity

This involves three distinct steps (Carmines and Zeller, 1979): (1) The theoretical relationships between the concepts themselves must be specified. (2). The empirical relationship between the measures of the concepts must be examined. (3) The empirical evidence must be interpreted in terms of how it classifies the construct validity of a particular measure.

---

[13] For example, the WTP measure is correlated with other measures that are intended to measure the same object such as QALYs for preferences over health (Klose, 1999).

the content-validity of the CVM include using informative scenarios, context-specific questioning and well-trained interviewers.

3.5.2 Face Validity

In testing for face validity, one asks whether the research tool seems to ask all the needed questions together with the use of appropriate language and language level to do so (Fink, 1993). Other face validity questions include whether the items appear to be measuring the variables they claim to measure and whether the meaning and relevance of the indicator is self-evident (Bowling, 1991). Some people argued that face validity is a form of content validity (Bowling, 1991), while others maintained that they differ because face validity (unlike content validity), does not rely on established theory for support (Fink, 1993).

3.5.3 Construct validity

Construct validity refers to whether the measurement corresponds to theoretical concepts (Klose, 1999). Unlike other types of validity testing, testing for construct validity involves assessing both theory and method (Bowling, 1991; Streiner and Norman, 1995).

Thus, hypothetical constructs are used to understand peoples' observable behaviours (Streiner and Norman, 1995). Construct validity must be investigated whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured (Cronbach and Meehl, 1955). It further sub-divided in the literature into 'theoretical/internal[12] and 'convergent validity'[13] and the tests could be used to detect the presence of all CVM biases.

Theoretical validity

One test of validity in a contingent valuation study is to assess whether hypothesised theoretical relationships between the elicited WTP and its explanatory variables are supported by the data (Mitchell and Carson, 1989). If CVM results are valid, the estimated parameters should normally be in accordance with prior expectations (Asgary and Willis, 1997).

---

[12] For example, WTP for a health effect rises with the scope of the health effect;

However, there is no one single experiment that can equivocally 'prove' a construct (Streiner and Norman, 1995). One must be able to state several theoretically derived hypotheses involving the particular concept and construct validity is then established through a pattern of consistent findings (Carmines and Zeller, 1979).

If the evidence relevant to construct validity is negative according to the prior hypothesised theoretical expectations, one could conclude that: the measure lacks construct validity; the theoretical framework used to generate the empirical predictions is incorrect; the method used to test the theoretically derived hypotheses is inappropriate; it was due to lack of construct validity or the unreliability of some other variable(s) in the analysis (Carmines and Zeller, 1979); or the theory is wrong and the scale is useless (Streiner and Norman (1995).

Carmines and Zeller (1979), however noted that there is no fool proof procedure for determining which one (or more) of these interpretations of negative evidence is correct in any given instance, and that it is the total configuration of empirical evidence that lends credence to one interpretation over another.

3.5.4 Criterion validity

Criterion (or criterion related) validity is the correlation of a scale with some other measure of the trait under study, ideally, a "gold standard" (Streiner and Norman, 1995). The operational indicator of the degree of correspondence between the test and the criterion is usually estimated by the size of their correlation, through a validity coefficient (Carmines and Zeller, 1979).

Criterion validity compares CVM values either with identical markets in which the same good is bought and sold, or with an experimental market which creates or simulates a market in which the good is actually bought and sold (Asgary and Willis, 1997).

Criterion validity is the gold standard of determining the validity of the CVM (Loomis et al., 1997). One of the longstanding criticisms of the method is that stated WTP may be a poor indicator of actual WTP. Thus, criterion validity is regarded as offering a more definitive test of CVM results (Asgary and Willis, 1997).

There is no single criterion-related coefficient for determining criterion validity (Carmines and Zeller, 1979). Instead, there are many coefficients as there are criteria. Streiner and Norman (1995) identified two forms of coefficient depending on whether the criterion (outcome) measure is dichotomous or continuous. If dichotomous, the phi coefficient is used, while if continuous, Pearsons correlation is used.

Carmines and Zeller (1979) noted that the test would not be useful unless it correlates significantly with the criterion; and similarly, the higher the correlation, the more valid is this test for this particular criterion. Nunally (1978) argued that even a modest correlation (e.g. a correlation of 0.30) between test and criterion can prove quite useful for selection purposes. The test could refer to a questionnaire survey to determine stated WTP for ITNs and the criterion is the actual sale of ITNs.

Criterion validity is divided into predictive validity and concurrent validity (Bowling, 1991; Streiner and Norman, 1995 and Carmines and Zeller, 1979). In concurrent validity, " we correlate the scale with the criterion measure, both of which are given at the same time" (Streiner and Norman, 1995). Predictive validity "refers to the extent a measure forecasts future performance" (Fink, 1993).

Establishing predictive and concurrent validity.

If the outcome is dichotomous (e.g. either somebody buys a net or does not), the results could be analysed using a measure of correlation which can be derived from a 2X2 table, such as the phi coefficient (Φ) (Streiner and Norman, 1995). According to these authors, the coefficient is related to the chi-square by the formula:

$$\Phi = \sqrt{\chi^2/N} \quad \text{---------- (1)}$$

Criterion validity can also be calculated from the 2X2 table illustrated by Table 3.2.

Table 3.2 2X2 table to evaluate criterion validity

	Actually paid	Did not pay
WTP	A	B
Not WTP	C	D

Criterion validity is then calculated from the 2X2 table using the following equation:

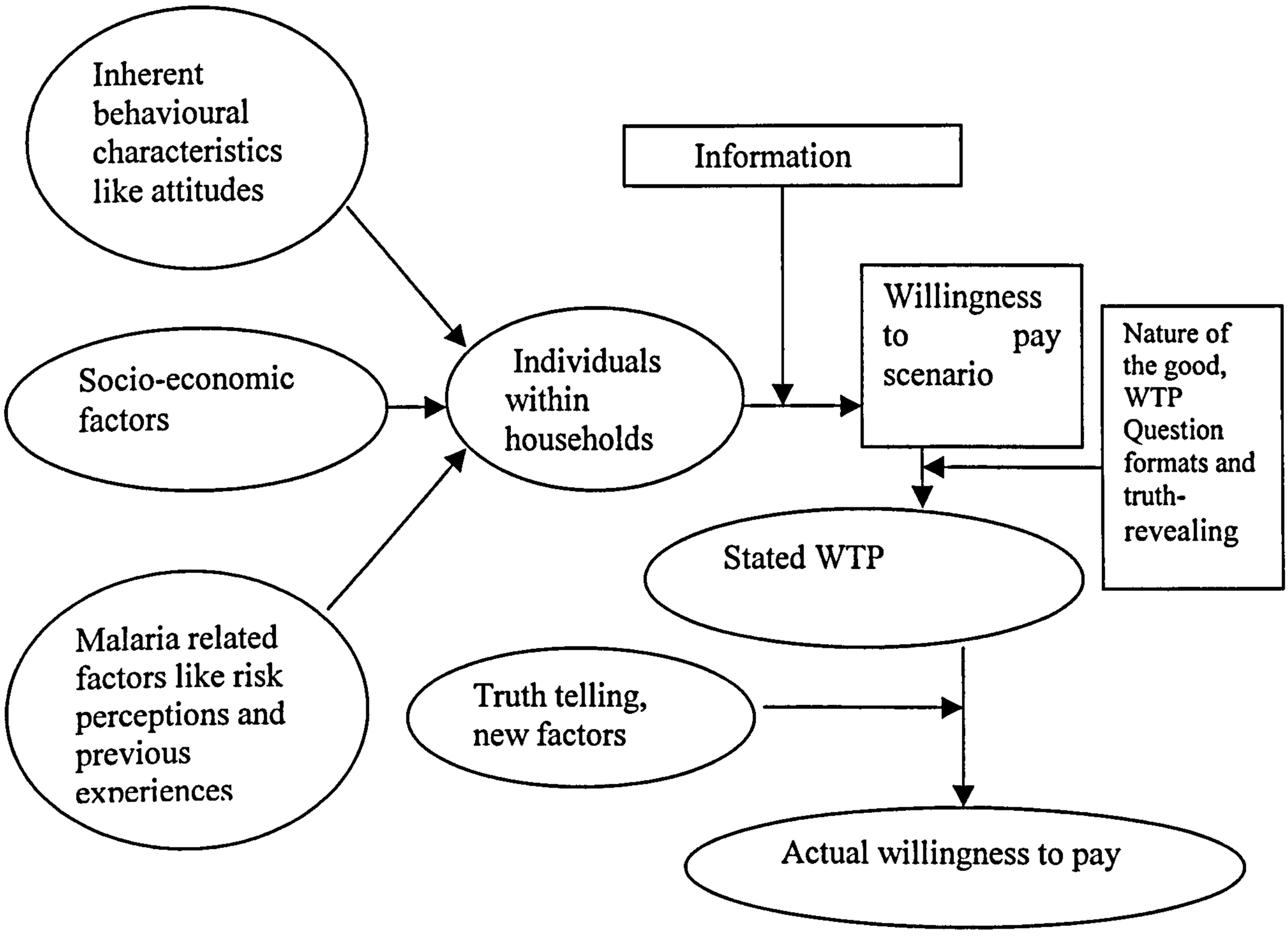
$$\Phi = |BC - AD| / \sqrt{[(A+B)\,(C+D)\,(A+C)\,(B+D)]}$$

However, if the measures were continuous, a Pearson correlation coefficient calculated through regression analysis will be used. In either case, one is looking for a strong association between the new measure and the already existing one. (However, Streiner and Norman (1995) did not state what the acceptable levels of strong association should be used to judge performance).

3.6: Study Framework of Reliability and Validity of CVM Responses

The essential elements that will guide further aspects of the study as gleaned from the theoretical and methodological review are presented in figure 3.1 and explained below.

Figure 3.1 The study framework of valid and reliable WTP elicitation



Explaining the framework

The Figure shows that individuals and households are constantly exposed to many dynamic and static factors that determine their demand for goods and services. However, in contingent valuation, the external information about the good and scenario presented by the interviewer could stimulate the thought processes in respondents. However, the nature of the good and the familiarity of the respondents with the question format will play a major role in deciding how genuine the stated WTP will be.

The translation of stated WTP into actual WTP brings a cascade of factors into play. A central one is how the consumers perceive the sales mechanisms and its accessibility. The consumers' perception is mainly with regards to trust, quality assurance and financial management (for instalment payments). Nevertheless, the potential ultimate factors involved at this stage include to what extent the respondents told the truth while stating their WTP, and the influence of factors that can cause a divergence between stated and actual WTP. All these will now lead to the actual WTP, which is the final outcome of the entire cascade of interactive factors.

Therefore, this study will attempt to measure determinants of divergences between stated and actual WTP and the link they will provide to truth-telling in CVM surveys. Neill et al. (1994), in determining why stated and actual WTP differed in their study, were unable to devise an incentive-compatible provision rule to elicit valuations that were demonstrably truthful. They also contended that research concerning the validity of CVM values requires a means by which values reflecting truthful real economic commitments can be obtained: the need is for valuation institutions with truth-telling as the dominant strategy. Fundamental to these is using appropriate question formats to ensure more reliable and valid estimates of WTP.

3.7 Conclusion

Future research calls for a deeper understanding of reasons for CVM biases and optimal ways of determining and improving the validity and reliability of WTP estimates. Only by doing this will these biases be either eliminated or minimised in such a way that the hypothetical WTP data will be a valid representation of consumer surplus. Further research may involve de-briefing the respondents through in-depth interviews, focus group discussions or questionnaire surveys to understand the socio-cultural, economic and other factors that govern their WTP decisions for particular goods and services and why stated and actual WTP might diverge. The knowledge to be gained from these exercises will aid the development of better CVM question formats that would elicit more reliable and valid consumers' valuation of their welfare.

An important key towards having reliable and valid WTP measures is to ensure greater realism in the scenario and the question formats so that the valuation will make sense to the respondents. In effect, this involves presenting the scenarios in ways that mimic the price-taking situations the people are familiar with. In the Nigerian context, this means that WTP elicitation should be modelled after the haggling technique that both buyers and sellers apply in the markets. The bidding game and the binary with-follow through techniques more closely resemble the haggling technique but there are still key differences.

In summary, in order to determine the WTP for health care goods and services whether in CBA or in demand modelling, the CVM is an acceptable method. It is grounded in welfare economics. However, the reliability and validity of the method has been questioned because of divergences between stated and actual WTP and inconsistencies with WTP value elicitation using different question formats. Thus, the design of appropriate context-specific formats could improve the reliability and validity of CVM. It would also be necessary to use appropriate survey and statistical analytic procedures for testing the reliability and validity of WTP estimates to ensure that conclusions are robust. Developing a better framework to tackle these tasks and move forward the CVM technique involves a review of the CVM empirical literature. This review would help provide the perspectives and prospects of applying the CVM technique for eliciting reliable and valid WTP for ITNs/health care in Nigeria.

CHAPTER 4: LITERATURE REVIEW OF EMPIRICAL WILLINGNESS TO PAY STUDIES

This chapter presents a critical review of the relevant empirical literature on willingness to pay that was determined through the contingent valuation method. The gaps in knowledge, weaknesses, and opportunities for improvement or additional work are also evaluated and discussed.

The specific objectives of the literature review were to: determine the question formats that have been used in past contingent valuation studies, and the rationale for the choice for the question formats chosen; determine which reliability and content validity tests have been performed, together with the methods used; explore the variables included in tests of construct validity, together with the model specification; determine how criterion validity was determined in the studies that compared stated and actual WTP; identify the recommendations on ways to move forward the WTP technique from the reviewed studies; and to extract parts of the literature vital for development of the thesis, and use them to provide a framework for designing the materials and methods for the study.

The WTP literature in English was searched and included that from both developing and developed countries since 1970. The key words for searching the databases were linked to the abstracts and only those articles whose abstracts fell into the inclusion criteria were reviewed in detail. The key words were: Willingness to pay, willing* to pay, willing* near pay, WTP, willingness to accept, ability to pay, contingent valuation, demand for health-care/Africa & Nigeria, affordability/Africa, revealed preference, expressed preference, Cost-benefit analysis/health care, consumer choice and valuation of health states.

The data sources comprised both electronic and non-electronic databases. The electronic databases were Popline, Medline, Health Star, HHIC, BIDS and Econlit. The non-electronic sources included books, conference proceedings, grey literature, working documents, published articles and other sources. An iterative literature search was also done and contact established with some active researchers.

The criteria for selecting the empirical papers for detailed review were: all literature on WTP in health care; WTP articles focusing on developing countries; WTP articles that investigated criterion validity in all sectors and WTP studies that determined reliability in all sectors. The criterion for selecting literature for the detailed review was that they must have addressed WTP fully in the paper and not just merely mentioned it. The literature were grouped into three categories; those that explored construct and content validity, those that compared stated and actual WTP, and those that determined the reliability of WTP estimates.

The review questions for face/content validity were whether the studies assessed content validity of the question formats and the scenario used, the methodology they used (especially whether face/content validity was explicitly determined through field-work), together with their findings.

The review questions for the articles that investigated construct validity were to; determine the type of good valued; the survey techniques (mail, telephone, in-person etc); the type of question formats used by each study and the rationale for using them; whether WTP determined was for CBA or used for market research; the type of validity tests conducted; the models used and the variables in econometric models; the variables that explained WTP and the type of diagnostic tests performed on the econometric models.

The review questions for articles that explored criterion validity include; the first three review questions applied to literature on construct validity; the study design used to assess criterion validity; and the data analytic methods. The questions for the literature review on reliability of CVM were the same as used for literature on criterion validity.

A total of 113 articles were reviewed from approximately 410 articles screened. Most (80) of the articles were from developed countries while 33 were from developing countries. Papers on health care dominated with 83 articles while there were 30 articles from other sectors. The majority (85) of the articles focused on assessing construct validity in health care, while just 17 and 12 articles investigated criterion validity and reliability respectively. All the 113 articles were reviewed for content validity as no article was focused specifically on this type of validity.

The literature review is presented in three major sections. Sections 4.1 to 4.4 present and critically discuss the findings of the issues explored in the literature focusing on the above questions, with reference to content validity, construct validity, criterion validity, explanations of divergences in WTP and reliability. Section 4.5 provides the conclusion of the review by highlighting the gaps in knowledge, ways of improving the WTP technique together with the implications of the literature review for the development of the structured haggling technique.

4.1 Content Validity

None of the studies reviewed explicitly determined the content validity of the question formats they used from the general population before and after their study. The ten studies that mentioned content validity of the question formats relied on personal views and face validity to state that the question formats used were content valid. The failure of the studies to explicitly gain insights from the general population that form the potential respondents and usual practitioners of price-taking in the market has resulted in many divergent views of what question format is content valid, and in what context.

Donaldson et al. (1997b) and Ryan et al. (1999) posited that the open-ended question format had no face validity, while dichotomous choice and payment card methods did. Donaldson et al and Ryan et al's arguments for preferring these two formats were that in dichotomous choice, people are usually presented with a price in shops to which they can only respond yes or no, and with the payment card because people shop around looking for the best bargain.

However, Donaldson and Shackley (1997) argued that the payment scale method makes it difficult to elicit a maximum WTP. According to Donaldson et al. (1997b), without any reference point, as in open-ended questions, many respondents estimate costs and state them as their WTP. The problem is to decide which argument to accept, since they also represent subjective conclusions of the authors, with no empirical backing. The conflicting arguments were also buttressed by Donaldson et al. (1997c) who felt that there was no strong reason for the strong case against the use of open-ended questions.

Sorum (1999) preferred the open-ended question format over close-ended, because "close-ended methods risk setting up artificial bounds on subject's answers". Sorum also felt that the dichotomous choice format was impractical since it required a large sample size, while the bidding game was too time-consuming for the multiple valuations done in that study. However, Brookshire et al. (1980) stated that the iterative bidding method has been found to be an effective data collection mechanism. Onwujekwe (2001) also argued that the bidding game and the dichotomous choice with follow-up format resembled price-taking in that study area.

While no study had empirically determined the content validity of any question format, 12 of them used such inputs to improve the questioning and WTP scenario. This is because "unrealistic questions are more susceptible to bias[14]" (Johannesson, 1992). In one study the investigators noted that individuals had difficulty valuing uncertainty and complicated life threatening situations in a short interview (Berger et al., 1987).

Whittington et al. (1992) selected the two starting-points for their bidding iteration based on initial focus groups and pre-testing they undertook. Also, Gonzalez-Caban and Loomis (1997) and Onwujekwe et al. (1998, 2001) obtained information together with reviewing evidence on the nature of the good in question, before developing the final CV scenario. This included information on the preferred payment systems, sale of the goods, common names for the health conditions and the nature of the goods.

Weaver et al. (1996) claimed content validity of the scenario by exploring a variety of ways of presenting the service to be valued, such as showing respondents drug packages, before determining whether the verbal descriptions were comprehensive and understandable for the respondents to identify them. Bala et al. (1998) used focus groups and pre-testing to determine the appropriate starting bids. Also, Berwick and Weinstein (1985) used initial focus groups to generate hypotheses for more systematic study, and then used multiple scenarios to elicit WTP[15].

---

[14] Unrealistic questions have no bearing on reality, e.g. total cure for rheumatoid arthritis or cancer.
[15] Multiple scenarios are where the attributes of the market is presented piece meal to the respondents before determining their WTP.

The findings thus show that the content validity of all existing question formats has not been proved in health care and in other sectors. Mimicking price-taking in the market as the basis for choosing an appropriate question format was posited by Hanemann (1991) and adopted by the NOAA panel in recommending the dichotomous choice question format (NOAA, 1993). The use of the dichotomous choice question format gained increased use after the recommendation by NOAA (Smith, 2000). However, there has been no comparative empirical study of content validity of one question format over others and further research is needed to prove which one is the best.

It is necessary to determine the market behaviours in different contexts so that using mimicking price-taking as the basis of choosing appropriate question format will become operationalised properly. This is because price-taking behaviour will differ, depending on the nature of the good, the socio-cultural, religious belief and location of the respondents. For instance, the differences in usual price-taking between Nigerians and the western world has been presented (Basden, 1938; Balogun, 1991).

An illustration of the need to determine market behaviours to inform the type of WTP question format to be used was the study by Clarke (2000) who sought to know why a respondent's WTP declined when s/he was asked a follow-up question. One of the explanations Clarke proffered was that the follow-up question comes as something of a surprise in Australia. The surprise at being asked a different amount in the follow-up question might induce a negative reaction. However, a different scenario happens in Nigeria where respondents would expect a follow-up question in preparation for bargaining. This example further illustrates the cultural contexts that need to be considered in using different WTP question formats in different areas and contexts.

Therefore, the prescription of universal question formats should be discouraged, and it is better to use context-specific question formats for the elicitation of more reliable and valid estimates of WTP. However, the disadvantage is that question formats may become too specific and studies may have to devote a lot of funds in developing the context-specific methods. Nevertheless, what may be required in many cases may be a modification of any of the existing methods to suit the context of the study.

The development or refining of question formats to be context-specific may be possible through empirical studies with the potential respondents and people like traders who usually use such price-taking mechanisms to sell goods and services. These sets of people are well positioned to advise the researchers on the formats that would better elicit WTP.

Thus, urgent work is needed to determine the content validity of different question formats in different settings and for different goods. In particular, such studies are needed in sub-Saharan Africa where the CVM is still in its infancy and where market behaviours are different from the Western world, which is now mostly using the dichotomous choice system to elicit WTP.

The content validity of the scenario elements and the questioning also needs to be improved for improved WTP validity and reliability. Brown and Green (1981) recommended that question formats should be designed so that the respondents are first enabled to discover their preferences before stating them, as this will aid their understanding of the good and why they would want to pay. Less cognitively challenging scenarios should be used to make answering WTP questions easier.

Qualitative research methods like focus group discussions with the populace, observations of market behaviour and interviews with the actual practitioners of market WTP elicitation (traders) could be used to determine and also improve the content validity of different question formats and the scenario. Then, with tests of construct and criterion validity together with reliability, it would then be possible to know whether better content validity of an question format translates to lower occurrence of hypothetical bias and/or strategic bias.

4.2 Construct Validity

A total of 85 studies that had determined stated WTP were reviewed (see Appendix 1). A minority (36%) of these studies was conducted in developing countries, while 64% were conducted in developed countries. Also, 90% were in health care and 10% in other sectors. These studies determined the willingness to pay for a variety of goods and services and the detailed outcomes of the already set criteria for reviewing these studies are presented in Appendix 1. The individual aspects for determining construct validity are described in brief in the following sub-sections.

4.2.1 Contexts of the WTP studies examined

More than 90% of the studies in health-care from developed countries estimated perceived benefit of the good and services while under 4% were concerned with market research, and the rest examined WTP from the perspective of both market research and valuation of benefit. Unlike the WTP literature from developed countries, more than 90% of research in developing countries was devoted to market research. Only two studies (Binka and Adongo, 1997; Whittington et al., 1996) focused purely on valuing benefits.

The WTP literature reviewed in other sectors showed that the technique has been used for only market research in sub-Saharan Africa (n=5) and in one study in Haiti, while it has been used to determine benefit elsewhere. The five studies in sub-Sahara Africa were in the water and sanitation areas and were intended to shed more light on the demand for such services and factors associated with such demand, and possible extension to price setting. Three studies from outside sub-Saharan Africa (Asgary and Willis, 1997; Gonzalez-Caban and Loomis, 1997; Alberini and Krupnick, 1998) were concerned with the benefit of different types of environmental improvements.

The use of the CVM is more established in other sectors in developing countries than in the health care sector. Also, better methods and reporting occur in these sectors. This could be due to the fact that most of the reviewed studies in other sectors were commissioned by The World Bank which had relevant trained people in the technique to conduct the studies, whereas in health care it was an all-comers affair!

4.2.2 Question formats used

The open-ended format was the most commonly used, followed by the dichotomous choice (binary) format (see Table 4.1). However, 14% of the studies used more than one question format, mostly to compare their performance in eliciting WTP and as a test of convergent validity between the two methods.

Table 4.1 Types of question format used

Type of Question format	n (%)
Open-ended	25 (29%)
Binary	20 (24%)
Bidding game	16 (19%)
More than 1	12 (14%)
Binary with follow-up	5 (6%)
Payment card	4 (5%)
Not stated	3 (3%)
Total	85 (100%)

4.2.3 Econometric and other models used to assess construct validity

Only 78% (66) of studies that elicited stated WTP determined the construct validity of their estimates and most of those that did (80%) used econometric modelling (Table 4.2). The detail of the literature on construct validity is presented in Appendix 1.

The most common methods for determining construct validity of the WTP estimates was through theoretical validity testing via probit/logistic econometric modelling, as 27% of the studies used either one of them. They were followed closely by ordinary-least squares multiple regression. Only three studies had strictly used models for limited dependent variables. A small percentage (23%) of the studies made no attempt to establish the construct validity of the WTP estimates. Though econometric models were the dominant construct validity tests, thirteen of the studies used non-econometric methods such as analysis of covariance and Kruskal-Wallis tests to determine construct validity, and six of this later group used tests of convergent validity to determine construct validity.

Table 4.2 Types of models used for construct validity

Model	n (%)
Probit/Logistic	23 (27%)
OLS	19 (23%)
No explicit test of construct validity	19 (23%)
Non-econometric tests of theoretical validity	13 (15%)
More than 1 econometric method	8 (9%)
Tobit	3 (3%)
Total	85 (100%)

The number of studies that assessed the construct validity of their estimates of WTP in health care represented an increase over the findings of an earlier review. Diener et al. (1998) found that only 50% of previous CV studies in health care had determined the construct validity of their findings. OLS models were mostly used before 1995, while the probit/logistic models became more fashionable after that period. This reflects the increased used of binary question format, which yield binary dependent variables, usually modelled with logistic or probit estimators. The earlier uses of OLS models reflect the predominance of the open-ended question format, which yield continuous variables. The explanatory powers of most of the models were low, though Johannesson et al. (1993a) argued that this is common in cross section data.

The use of different models to analyse the same data set would determine the robustness of the parameters and point to differences and similarities between estimators. Hence, Whittington et al. (1992) compared four econometric models using different versions of the dependent variable. They found that the magnitude of the impact of the key explanatory variables (the coefficients) on WTP were larger when estimated with either the OLS or the Stewart Maximum likelihood estimator, than with the probit/logistic estimator. Dalmau-Matarrodona (2001) found that comparative parameter estimates from both the Tobit and Double-Hurdle (DH) models revealed significant differences. Dalmau-Matarrodona performed a likelihood ratio test since the DH nests the Tobit model, and rejected the Tobit model at standard significance levels.

Nature of zero WTP responses and econometric modelling for construct validity

Elicitation of zero WTP amounts is common in CVM studies that used the open-ended or bidding game formats. These limited dependent variables resulting from the large number of zero WTP amounts need a different modelling approach from dependent variables that are not limited (Donaldson et al., 1998). However, it was interesting to note that Tobit and other models for limited dependent variables have been hardly used. Some of the studies explained the zeros away as protest votes or designed the studies in ways that zero and negative responses were excluded from the estimation. Donaldson et al. (1997c) asked monetary WTP questions to only those respondents who answered yes to the question of whether they were willing to pay. This meant that absolute WTP values were inflated by excluding those who would gave zero WTP valuation. Excluding the zero responses from the analysis is not optimal, as valuable information necessary for explaining WTP is lost.

Only Donaldson et al. (1998) and Gyldmark and Morrison (2001) had tried to unravel the nature of the zero WTP. The explanations that they gave of why some respondents may state a zero WTP include genuine zero in instances that the respondents felt that the good was not worth buying or were unable to pay. Other reasons were: the respondents could have had very small WTP amounts that they approximated to zero, they did not understand the question, they were unable to make a clear decision or they were protesting at being asked the WTP question. Gyldmark and Morrison (2001) explored the nature of their zero WTP responses by determining whether WTP systematically varied by five income classes, gender and age groups.

In some studies, the zero responses were not excluded, but the continuous WTP estimates were converted to a binary variable. Studies in this group include Whittington et al. (1990, 1992, 1993, 1996) who elicited continuous WTP through the bidding game, but preferred to use the binary responses to the first yes and no questions in the bidding game. In justifying their use of models for binary dependent variables, Whittington et al. (1990, 1992, 1993, 1996) argued that the first binary responses in the bidding game represented the most objective estimates of WTP. However, Whittington et al. did not have any empirical evidence to back-up their assertion. Also, if Whittington et al's assertion was true, then one could question their rationale for using the bidding game repeatedly in these studies.

Donaldson et al. (1998) compared the Tobit, Heckman and OLS models, while Gyldmark and Morrison (2001) compared Heckman and log OLS models. Donaldson et al. (1998) found that the Heckman selection model was the best for modelling limited dependent variables, but Gyldmark and Morrison (2001) found that log OLS was better than the Heckman selection model for the same type of dependent variable. However, Donaldson et al. (1998) suggested that either the Heckman selection model or the Tobit model should be used depending on whether the zero responses were true zeros, protest votes or a mixture of reasons. Also, Gyldmark and Morrison (2001) stated that investigation into why respondents state zero WTP would help to reveal whether they were the result of the respondents' confusion, a true zero valuation or simply a strategic or protest response bias.

Gyldmark and Morrison (2001) used the OLS for their modelling by arguing that both the Heckman selection model (from a limited empirical testing) and Tobit model were inappropriate for their data, because their elicited WTP amounts were not censored at zero. Gyldmark and Morrison's 'vague' argument was that "given that the payment card they used started at zero, respondents may have been discouraged to state a negative value. But since even if you own an insurance policy you are not obliged to use it, a negative WTP seems implausible. So the distribution is unlikely to be censored at zero".

Nevertheless, Gyldmark and Morrison (2001) justified their use of OLS instead of Heckman selection model based on a recommendation by Duan et al (1993). Duan et al argued that if zero WTP is determined by some exogenous random parameter, then the probability of paying zero is not-nested with the WTP function and it would then be more appropriate to consider a log-transformed OLS model instead. However, it is questionable whether this still leaves biased estimates as the log-transformed OLS will only estimate the WTP of only those who had stated a positive WTP.

4.2.4 Variables included in the models for assessing construct validity

The explanatory variables included in the base models for assessing theoretical validity could be classified into three broad groups. The first group of variables includes socio-economic and demographic factors, with the key variables being the age of the respondent, sex, household size, education of the respondent, household income and proxies for income like assets. The second group of variables is those directed at determining the perception of the respondents about the good or service being valued. The third group of variables includes those that represent the respondent's personal or household experiences with the good.

Because the hypothesised relationships of different independent variables with WTP were different in many of the studies, the results were interpreted differently in the studies. Nevertheless, income and education were consistently hypothesised to be positively correlated with WTP.

The correct signs on the significant variables reveal theoretical consistency in the CVM results (Asgary and Willis, 1997). However, in models where no significant variables were found, the interpretation was that the responses seemed to be more a reflection of guessing rather than rational considerations (Muller and Reutzel, 1984), as there was no variable that would explain the relationship of the elicited WTP to determinants of consumer choice. The WTP estimates would be interpreted to be theoretically not valid in such cases since there are no economic explanations of the elicited WTP.

The variables in some cases, though statistically significant, may not obey the set hypotheses. Whittington et al. (1991) in determining WTP for water in Nigeria found that WTP and income were negatively and significantly correlated, instead of their hypothesised positive relationship. Whittington et al. argued that this could be due to measurement problems with the income variable, or perhaps the higher income households were more likely to respond strategically, or many higher income households had already solved their water problem, and were thus less interested in paying. Whittington et al (1996) also found that an asset (private water connection) that was supposed to signal households who were better-off than others was negatively correlated with WTP. This they claimed was because those households may be less vulnerable to water-borne diseases than others that live in houses without private connections.

In some model estimations, variables that failed to obey prior hypotheses were dropped from the econometric models. For instance, Gonzalez-Caban and Loomis (1997) dropped income from their model because it was negatively correlated with WTP. However, one could argue that it is wrong to drop such variables, since they could be providing information that it is more in tune with reality than the researchers subjective views used in developing the hypothesis. The practical approach is to see whether there are reasons that could explain the unexpected behaviours.

Whittington et al. (1990) proffered three reasons why variation in WTP bids cannot be explained by variables suggested by economic theory after finding strategic bias in their study. First, economic theory may not be an appropriate conceptual framework for explaining the behaviour and preferences involved. Second, economic theory may be correct, but the contingent valuation method may not be a sound method for collecting information to estimate the water demand relationships suggested by such theory. Third, errors in execution of the research, such as poor questionnaire design, could lead to invalid inferences about the relationship between the WTP bids and the independent variables.

4.2.5 Diagnostic tests

A huge weakness found in the econometric modelling of WTP was the paucity of diagnostic testing. This is more pertinent with OLS, Tobit and other models that rely on strict adherence to the assumptions of the classical linear regression model to generate unbiased estimates.

Only eight of the sixty-six studies that determined construct validity performed any sort of regression diagnostics. However, two of these studies did not present the results of such tests and whilst the study by Donaldson and Shackley (1997) showed that the model was incorrectly specified (through the RESET test), the authors still went ahead to describe and make inferences from the parameter findings.

In one of the few studies that did diagnostic testing, a probability plot together with a scatter plot of the residuals versus the fitted WTP values were used (Olsen and Donaldson, 1998). These plots revealed the underlying distribution to be approximately normal with constant variance, and the authors concluded that the OLS model used was appropriate for the analysis. Ideally, it should be mandatory for any econometric model used to determine the theoretical validity of WTP to be accompanied by diagnostic tests and not just tests showing the fit of the model like chi-squares and adjusted R-squares. The appropriate diagnostic tests for different types of models should be included in future studies.

4.2.6 Convergent validity

As a test of convergent validity, Smith (2001) compared the relative sensitivity of time-trade-off (TTO) and WTP values obtained at different levels of change in health status. Smith found that WTP was a more sensitive measure of change in health status than TTO. This was because the WTP was more sensitive than TTO in distinguishing between different dimensions of health and was also more sensitive to differences in quality of life between different levels of health. O'Brien and Viramontes (1994) also compared WTP to TTO and Standard gamble (SG) and found some convergence, while Bala et al. (1998) found divergence between WTP and SG.

The basis for the theoretical expectations in some of the tests of convergent validity was not clear. Clarke (2001) found that the contingent valuation and travel cost method yielded different WTP estimates with the travel cost method being lower. Clarke's conclusions were that CVM estimates might also reflect altruism, or that potential biases known to affect both methods could have been the culprit. However, it could be argued that the divergence in WTP estimates found by Clarke (2001) accorded with theoretical expectations since CVM estimates the total consumer surplus and so should exceed the travel cost method. Thus, a divergence is expected if two estimations are not actually valuing the same thing. In fact, this is why cost of illness is seen as the lower bound of WTP.

All these results suggest that the field of convergent validity testing is confused. The rationale behind the expectations that estimates from theoretically different methods would converge were not explained in the studies. It is not clear to me why WTP estimates of Hicksian compensating variation should converge with TTO or SG. The different methods value different aspects of consumer choice and are also designed differently. A divergence in such cases may be a greater indicator of construct validity unless there is a strong argument that they measure something similarly.

Higher levels of convergence would be expected between WTP elicited by different question formats than comparing with estimates of TTO or SG, as like is compared with like. Such comparisons were to determine the comparative reliability and validity of WTP elicited by different formats. Thus Donaldson (1990), Donaldson et al., (1997b,c), Johannesson et al. (1991) and Neumann and Johannesson (1994) compared different question formats and came up with different findings. Clarke (2000) compared the single and double bounded dichotomous choice (DC) question formats and found that the likelihood ratio test for probit models of the two question formats was statistically significant, rejecting the null hypothesis that the mean of the underlying distribution was identical for both formats.

Other comparisons of different question formats as tests of convergent validity were geared at determining the performance of the NOAA (1993) recommended DC method with the open-ended format. Desvousges et al. (1993) found that the means from the DC data were much larger than the open-ended data, with the differences often statistically significant. They therefore called for a reassessment of the DC format. Loomis et al (1997) also found that DC data exceeded open-ended WTP estimates.

All in all, most of these studies that assessed both theoretical and convergent validity were experimental and were geared to establish the construct validity of the WTP method. However, a recurring theme in almost all of them was that in order to advance the use of the WTP technique, there is a need for further testing of construct validity, together with assessing its criterion validity and reliability.

4.3 Criterion Validity and divergences in WTP

The CVM is based on a hypothetical market in which respondents are not actually required to make the contributions they claim to be willing to pay (Foster et al., 1997). According to Foster et al, "this feature of CVM has provoked widespread criticism and led to experiments comparing real and hypothetical payment aimed at ascertaining whether there is any systematic divergence between the two".

Many social scientists are sceptical of the accuracy of responses to hypothetical questions, yet few studies assess the quality of this type of data (Dickie et al., 1987). Dickie et al. also argued that comparisons of hypothetical and actual WTP data would be of immediate value for assessing the relative usefulness of hypothetical data as well as for indicating ways to improve demand-revealing data collection methods. Test of criterion validity is the strongest validity criterion as it compares whether stated WTP will equal actual WTP. It is also devoid of the assumptions and analyst influences that are common in tests of construct validity using econometric and other approaches.

Although there have been over a thousand applications of the contingent valuation method (CVM) (Carson et al., 1994), few of these studies have tested the validity of the CVM responses, especially criterion validity (Loomis et al., 1996). However, the literature on criterion validity is gradually growing in all sectors, although the nature of the goods valued differ.

4.3.1 Contexts of the studies and question formats used

Whilst the contingent valuation method was originally designed to value the benefits of public goods, only Foster et al. (1997) have used a public good to determine criterion validity, while others have used private goods (see Appendix 2). This is because of the difficulty of getting the public goods to use, since actual payments are not usually made for most of them.

Of the 18 studies reviewed, most used the dichotomous choice (binary) question format (50%), followed closely by the open-ended system (33.3%). Four of the studies were in the health care sector (three examined WTP for ITNs in Africa and the other WTP for pharmacists' services in USA), while the rest were in other sectors. Most (77.8%) of the studies had relied on laboratory type simulated markets[16]. The three studies conducted in Africa used more realistic settings in determining criterion validity. For example, Dgedge et al. (1999) sold nets to respondents within their villages as part of a malaria control programme.

Where criterion validity was applied, three out of the four studies in health-care did so in the framework of Insecticide-treated nets' programmes and were conducted in sub-Saharan Africa, using the bidding game. The methodologies used in those studies were not presented in a way that one could make objective conclusions on the validity of the WTP technique. For example, Dgedge et al., (1999) determined the stated WTP in 1995 and the actual WTP in 1997. This was a long time frame to determining criterion validity as a person's WTP could have changed in this long period. Rowley et al. (1999) and Onwujekwe (2001) interviewed a subset of the sample to determine the divergence between actual and stated WTP for insecticides for re-treating mosquito nets. This method, though necessary if there are resource constraints to interview all the original respondents, is flawed by the fact that the sub-sample is non-representative.

4.3.2 Methods used to assess criterion validity

Different study designs were used to compare stated and actual validity by the studies. For instance, Johannesson et al. (1997,1998, 1999), Blumenschein et al. (1998), Foster et al. (1997), Cummings et al. (1995) and Neil et al. (1994) used both between (different respondents) and within-groups (the same respondents) to compare stated and actual WTP. However, Blumenschein et al. (2001) and Frykblom (1997) used entirely different respondents to compare stated and actual WTP, while Dgedge et al. (1999), Rowley et al. (1999) and Onwujekwe (2001) used the same respondents for comparison.

[16] Simulated markets are artificial markets created just for the purpose of the study.

Most of the studies in developed countries paid the people to take part in the studies (e.g. Blumenschein et al. (2001), Johannesson et al. (1998), Blumenschein et al. (1998), Cummings et al. (1995), Neil et al. (1994), Frykblom (1997), Loomis et al. (1996), Neil et al (1994) and Loomis et al. (1997)). A few of the studies actually gave the respondents loans with which to fulfil their actual WTP pledges (e.g. Johannesson et al. (1998), Cummings et al. (1995), Frykblom (1997) and Loomis et al. (1996)). The majority of studies with the exception of Foster et al. (1997), Dgedge et al. (1999), Rowley et al. (1999) and Onwujekwe (2001) explicitly informed the respondents that the tests were for experimental purposes.

However, many of the studies that reported comparisons of actual and stated WTP as tests of validity could be better classified as reliability tests. This is because actual WTP was not determined in many studies (72%) by directly offering the goods to the all the respondents (since they did not have enough of the good). In other words there was no test of criterion validity. Rather, the respondents were asked to state the price they would now pay, assuming the good was for sale. Comparisons were then made between the first and second stated WTP amounts, with the second stated WTP being regarded as actual WTP.

4.3.3 Comparative criterion validity of different question formats

The studies that compared the criterion validity of different question formats ended-up with mixed results and the findings did not favour any format. For instance, there was no statistically significant difference between the open-ended and binary formats (Frykblom, 1997; Loomis et al., 1997; Brown et al., 1996; Kealey et al., 1988). Another study found no difference in the criterion validity of the bidding game with the binary-with-follow up format (Onwujekwe, 2001). Loomis et al. (1997) thus argued that the studies suggest that it may be premature to abandon the use of open-ended questions.

The literature reviewed showed that the comparisons of the criterion validity of different question formats did not favour any format's elicited WTP, though the dichotomous choice method is now the most widely used CVM question format. Since the studies that compared the criterion validity of different question formats found no statistically significant differences in their comparisons, a hypothesis could be that there is no significant difference between different question formats for determining actual WTP.

However, the studies that used single question formats like Neil et al (1994) showed evidence that open-ended hypothetical surveys do not always accurately elicit the real economic commitments from individuals. In the light of findings of the latter studies that have shown no difference in criterion validity of different question formats, the use of single formats in such studies and making conclusions based on the findings will be misleading as it may lead to condemnation of a particular format.

Since criterion validity is the ultimate test of validity because of the objectivity involved, it should be the basis for informing on the best question format to be used in different contexts. The result of comparative criterion validity of different question formats cast doubt on the NOAA (1993) recommendation. Instead of prescribing a universal question format, the recommendation could be the use of at least two elicitation methods per valuation exercise, and using the WTP estimates from the method with better criterion validity as the best estimate.

4.3.4 Divergences between stated and actual WTP

The potential divergence between stated and actual payments has sometimes been described as "hypothetical bias" (Cummings et al., 1986). However an important question is, "if the stated and actual WTP differed, can we conclude that the differences are ambiguously attributable to differences between real and hypothetical payment?" (Neil et al., 1994). The three reasons Kealey et al. (1988) proffered as to why stated and actual WTP could diverge are; (1) errors in the contingent values, (2) errors in actual expenditures and (3) errors in both sets of values.

The majority (67%) of studies found divergences between stated and actual WTP, with hypothetical WTP overestimating actual WTP. However, Loomis et al. (1996) found that the correlation between stated and actual WTP improved when respondents were reminded to be honest in stating their WTP. However, only Kealey et al. (1988) empirically investigated the reasons for the divergences using a probit model and found that people who stated higher WTP and had higher behavioural intention with respect to buying the commodity were more likely to change their minds, while those with greater experience were less likely to change their answers.

The calibration[17] of stated WTP responses was used by many studies to limit the divergences between stated and actual WTP. In calibrating stated WTP, Johannesson et al. (1998, 1999) treated only the respondents who were very certain of paying their stated amounts as having positive WTP. In the same vein, Blumenchein et al. (1998, 2001) treated the definitely sure people as having positive WTP. However, Blumenchein et al stated that a major weakness with treating only definitely sure yes responses as yes responses, was that there is little theory to support this approach. However, it was posited that "the standard hypothetical and conservative (calibrated) hypothetical WTP provided a range for actual WTP" (Johannesson, 1998).

Calibration of responses makes hypothetical WTP equal to actual WTP, and it could be done using socio-economic characteristics of the respondents or the follow-up questions to identify real yes responses (Johannesson et al., 1999). Johannesson et al. (1998, 1999) found that "definitely sure" yes responses were found to significantly underestimate the real yes responses and thus provided a lower bound of real WTP. Blumenschein et al. (1998, 2001) found that definitely sure responses corresponded to real purchase decisions. Similarly, Champ et al. (1997) found that hypothetical donations significantly exceeded real donations but there were no significant differences if only subjects that were very certain of their yes responses were counted as real yes responses.

The artificiality of formal experiments to compare stated and real WTP may engender a bias in and of itself. This is so in the sense that subjects are placed in a purchase situation which they would not necessarily have chosen for themselves (Foster et al. 1997). In the case of divergences noticed for a public good (where stated WTP was at least four times the real payments), the primary effect of the hypothetical context may be to reduce the extent of extreme free-riding, as opposed to creating incentives for strategic over-bidding (Foster et al., 1997).

[17] For instance, follow-up questions are used after eliciting WTP to calibrate the responses into certainly yes, probably yes, probably no and certainly no. The method of calibration varied from study to study.

The market used to determine actual WTP could explain why most studies that used simulated markets experienced high divergences, while those that used actual markets had little or no divergence. In the latter case, the people would still have found themselves in that "market" since goods they required were put on sale, thus limiting their incentive to give spurious stated WTP values. However, in simulated markets, the respondents were put into another hypothetical situation, with many studies explaining to the people that the study was just experimental, and some studies even paid people to participate. These factors could influence the respondents to give spurious WTP estimates, which are interpreted as the actual WTP.

To understand why stated and actual WTP differ, it is important to analyse the motivation in individuals' answers to hypothetical questions and to develop a theory of how individuals respond to hypothetical questions (Blumenschein et al., 1998). Foster et al. (1997) found that the divergences between real and hypothetical behaviour is largely attributable to the higher positive response rate obtained in a hypothetical setting, and that the mean positive valuation does not vary greatly between the two.

Thus, "an important lesson to be learnt from validity experiments is to debrief and probe respondents about the differences in their decision processes in hypothetical versus actual cash decisions" (Loomis et al., 1996). This is especially as the pre-test conducted by Loomis et al. showed that respondents were in a different frame of mind when considering hypothetical and actual WTP scenarios. Using the knowledge gained from the de-briefing and probing, statements to combat the hypothetical nature and place them in a real market frame of mind can be developed to better improve the match between intended and actual behaviour (Loomis et al., 1996). Nonetheless, some approaches have been introduced to decrease the divergences between stated and actual WTP no matter which question format is used.

A major issue with criterion validity is that the divergence between actual and stated WTP should be better characterised. Divergences in WTP could be attributed to changes in demand and other taste factors between the periods when stated and actual WTP were determined. The variables that were mentioned as possible causes in those few studies were; income, age, gender, shop, race, market price and interaction variables (Dickie et al., 1987), price, sex, age, money (Blumenchein et al., 1998) and income, age, market and bid (Loomis et al., 1997). The divergences between stated and actual WTP could also be from the operational flaws of the research methods used. Hence, most of the study designs that have been used to determine criterion validity especially in developed countries, were flawed from the beginning and the actual WTP situations could be interpreted as hypothetical situations.

Genuine divergences could be better determined in real settings and not laboratory experiments, and using goods the people would really like to buy. In real settings, sufficient quantity of the good should be offered for sale to all the respondents, unlike in laboratory experiments that usually have one item of the good and people are not confronted by the kind of decision they make in real life situations. Then, factors that could cause divergences should be determined, as human beings have dynamic tastes. Also, there could be economic factors like change in income, changes in prices of substitutes that could lead to divergences between actual and stated WTP.

Therefore, divergences between stated and actual WTP caused by genuine reasons should not be classified as a sign of hypothetical bias, but rather behaviour in line with theoretical expectations. Thus, some divergences may actually be in support of the validity of the WTP technique and the question format used to elicit it.

All in all, a possible solution for improving criterion validity of WTP estimates is to use goods that the respondents would ordinarily like to buy, as the discrepancy between hypothetical and actual conditions is smaller when the commodity is more important to the respondent (Kealey et al., 1988).

Another possibility is to use question formats that will assist in making the respondents tell the truth in the surveys. It is also better to use payment and sales mechanisms that the people in the study context are either used to or which have been recommended by them, as better realistic settings for the project could help to make the WTP estimates more criterion valid. However, if it is not possible to mimic the payment and sales mechanisms used for the same or similar goods, then the potential respondents could be interviewed before the survey to determine their sales and payment preferences to be included in the scenario.

Thus, limiting the problems of divergences between actual and stated WTP potentially calls for more realism in setting up the studies. This is in terms of using appropriate content valid question formats, sales strategies and payment mechanisms for the exercise. In addition, explicitly determining the causes of divergences, possibly through a triangulation of methods, will lead to objective classifications of cases with hypothetical bias and those that are not.

4.4 Reliability

There was general paucity of information about the reliability of CVM and the different question formats as very few studies have investigated reliability in either the health-care or other sectors. Just 4 out of the 12 studies reviewed attempted to determine the reliability of the WTP technique in the health care sector (see Appendix 3).

4.4.1 Contexts of the studies and question formats used

Equal numbers of the studies used public and private goods to assess the reliability of generated WTP and all were conducted in developed countries. One of the studies in addition compared the reliability of the dichotomous choice and open-ended question formats. The question formats used varied: 5 used dichotomous choice; 4 used open-ended format; 2 used the bidding game; and one study used both the bidding game and open-ended questions.

4.4.2 Methods used to assess reliability

Test retest was the most common form of reliability determined and was used by 11 of the studies. Test-retest is the standard method used to investigate reliability of contingent valuation (Teisl et al., 1995). Desvouges et al. (1993) assessed intra-analyst reliability and Reiling et al. (1990) assessed the temporal reliability of the contingent valuation method. 11 of the studies administered questionnaires to elicit WTP for the tests of reliability, while Whitehead and Hoban (1999) used a telephone survey for the same purpose. The period between test and retest varied from 5 years (Whitehead and Hoban, 1999) to 2 weeks (Kealey et al., 1988); there was no most common time difference used in the studies.

Various analytical techniques were used by the studies to assess reliability. However, 6 studies used more than one analytical method and the rest used just one method. The analytical methods were correlation coefficients (n=6), tests of differences in means or medians (n=8) and regression analyses to assess whether the determinants of WTP changed between the two periods (n=7).

Compounding the picture of under-investigation of the reliability of WTP is the use of both field and analytic methods that do not measure reliability. The study by Golan and Shechter (1993) may not be considered to have assessed reliability since the general fit of a regression model was used as the test of reliability and as argued in Chapter 3, this is not a test of reliability. Desvousges et al. (1993) and Whitehead and Hoban (1999) could be rightly said to have only looked at construct validity. Carson et al. (1997) agreed that the regression analysis, which they conducted, was a conventional test of construct validity and was not a reliability test.

It was not easy to objectively critique the conclusions about whether CVM was reliable or not from some of the studies because of the varying analytical methods and lack of criteria set by the studies on the range of results that would be regarded as reliable. However, the reliability of WTP estimates was easier to comprehend in studies that used tests of differences in medians or medians, since insignificant correlation between WTP generated at different periods clearly showed that the two estimates were similar and thus, WTP was reliable.

However, WTP was unconditionally reliable in 8 of the studies. Sorum (1999) found that the WTP estimates became reliable when the extreme WTP amounts were deleted from the analysis. Whitehead and Hoban (1999) also found that WTP became reliable once the explanatory factors that caused changes in WTP after the five-year interval between test and retest were controlled for. The only study that compared the reliability of two question formats (Kealey et al., 1988) found that both methods were similarly reliable.

4.4.3 The exploration of divergences in test-retest and reliability of WTP estimates

Some of the studies determined reasons for lower than expected test-retest reliability coefficients, as implicitly a reliable measure is "where the 'true value' must be unchanged and sample characteristics, such as income, must be held constant during the test-retest period" (Reiling et al., 1990). This is not a realistic expectation as income and other determinants of WTP could change over time, and thus low test-retest reliability might not signal unreliable WTP estimates or vice versa. Thus, a reliable WTP estimate is one that reflects the constancy of values when preferences and choice sets do not change and reflects changes in values when preferences or choice sets have changed (Teisl et al., 1995).

Hoban and Whitehead (1999) found that WTP was more reliable when socio-economic and attitudinal variables were added to regression analysis. Reiling et al. (1990) posited that the period between test and retest would increase the probability that the "true value" might change, or make the respondents not take the valuation exercise seriously or remove themselves from the experiment. Hoban and Whitehead argued that low reliability WTP estimates might still mean that measurement was reliable, if there are logical reasons to show that the change in WTP was consistent with change in WTP determinants like income, perceptions and priorities.

Increased realism of the contingent valuation survey was recommended by Mitchell and Carson (1989), as a strategy for increasing the reliability of WTP estimates. To increase the realism of the hypothetical market in assessing test-retest reliability for preservation of Mono Lake, Loomis (1989) included coloured maps of the lake as part of the scenario and this enhanced the reliability of the WTP estimates.

The length of time between test and retest was a critical issue debated by some of the studies. If the time is too short, the respondents are likely to remember their previous WTP values, and if too long, determinants of demand and 'true' WTP could change thus making the WTP seemingly unreliable. Teisl et al. (1995) used independent pre-test and post-test control groups to allow testing of reliability when recall occurs or the true value changes and found that WTP measures were reliable regardless of those factors. However, there is no universal recommendation on acceptable time interval between test and retest nor the types of reliability tests that should be performed. Investigating reasons that reliability coefficients may be lower than expected is an important part of WTP reliability studies so that more informed interpretation of reliability coefficients would be done.

4.5 Conclusion

The evidence from the review of literature is that CVM is increasingly being used to better understand the method and to aid decision-making. In sub-Saharan Africa and other developing regions, the emphasis has been to use WTP for market research – to understand demand and set prices for goods and services (Weaver et al., 1996).

However, CVM, especially with regards to the question formats and survey methods may not be transferable between developed and developing countries without modifications because the context, people, market situation and major uses are quite different. For example, it would be inappropriate to use self-administered questionnaires, mail or telephone surveys in most developing countries, especially those in sub-Saharan Africa with low literacy rates where respondents will need assistance to complete the questionnaires.

There is a need for appropriate validity and reliability tests that could be used to assess the elicited WTP values, which could also help in the decision on which question formats could be used in different areas to better elicit WTP. However, no single test of validity is definite (Mitchell and Carson, 1989). Kealey et al. (1988) argued that on the basis of the comparisons between contingent values and actual expenditures, it is not possible to determine how closely either value approximates the theoretically correct Hicksian value. So, multiple tests may be needed in a study to better assure the validity of the findings. In addition, the results of studies point to the need for research to show how biases attributable to hypothetical payment might be identified and mitigated.

In determining the construct validity of WTP estimates, a big gap in the studies was the lack of diagnostic testing of models on which tests of construct validity are based. There was also very little comparison of alternative econometric models for the same data set to see which gives a better explanation of the estimated WTP. There is urgent need for literature to report on test results, otherwise spurious models give misleading interpretations of the effects of different variables on WTP. Thus, each construct validity model should use at least two appropriate models for modelling the estimated WTP, subject them to diagnostic tests and then choose the better one in order to assure that the reported construct validity is accurate.

In tests of criterion validity, a major problem was that most of these studies only had one item of the good and sold to only one person. Ideally, they should have had many units of the good and asked people to now pay what they stated they would pay as a real test of criterion validity. Many of the studies also used laboratory-type experiments that potentially led to biased estimates of criterion validity of the CVM. Furthermore, in many of the studies, stated and actual WTP were determined using different sets of respondents and the sample sizes used were inadequate for comparable criterion validity, accounting for their use of diverse methods for data analysis. The framing of the questions in some of the studies also informed the respondents that they were either in the group that would only respond to hypothetical WTP questions or in the group that would be used to determine actual WTP, thus predisposing to biased estimates.

These points mean it is unlikely that robust conclusions about criterion validity can be made, based on most previous studies. A better design would be to use the same people for determining both actual and stated WTP, within a relatively short-time period, to ensure that all respondents would be offered the good for sale and then directly determine the level of criterion validity using any of the tests. It would then be possible to also determine whether there were divergences or not between stated and actual WTP and the causes of those divergences.

Standard analytic methods should be used so that criterion validity across studies can be easily compared. Positive predictive validity and phi correlation coefficient tests from psychometrics could be used. The usual practice of comparing the stated WTP amounts elicited from respondents whose hypothetical WTP were elicited with those whose actual WTP were elicited could be regarded as a test of reliability. A few studies modelled stated and actual WTP using either logistic or probit models and then used likelihood ratio index test to see whether the two models were similar. Only one had used standard tests of criterion validity like positive predictive validity, and the phi correlation coefficient has not been used by any study.

There is generally very little work on the reliability of the CV method. Thus, simply determining the reliability of WTP estimates will be an important addition to knowledge. It will also be better if a wider variety of types of reliability were determined, since test-retest has been assessed in over 90% of studies. There should also be a standard definition of what reliability means and analytic methods in CVM by borrowing from psychometrics, since many studies have defined and measured reliability in ways that do not fit into any of the reliability categories.

An unresolved question is which question format to use in different settings, especially as none has been shown to be superior. One approach would be to present alternative question formats to potential respondents, then determine their relative content validity before deciding on the one to use for the survey. Improving content validity may minimise divergences between stated and actual WTP.

It is important to continue the work on testing the reliability and validity of the CV method, especially to systematically compare and evaluate different question formats (Johannesson et al., 1993a). Donaldson et al. (1997b) recommended that question formats are required which make it more obvious to the respondent what the questionnaire is trying to achieve. In addition Johannesson (1992) and Frykblom (1997) stated that future research should include systematic comparisons of reliability and validity of different question formats to find out which question mode best elicits actual WTP.

CHAPTER 5: EMPIRICAL RESEARCH METHODS

This chapter describes and justifies the empirical research methods used in this thesis. The aim of the empirical study was to determine the validity and reliability of the CVM and that of different question formats, by determining the stated and actual WTP for ITNs and the stated WTP for ITNs re-treatment. The specific objectives were:

1. To assess perceptions, attitudes and practice with respect to malaria and ITNs.
2. To develop and determine the reliability of and validity of an indigenous CVM question format (called the structured haggling technique) that mimics the price-taking process in Nigeria markets.
3. To compare the content, construct and criterion validity of the novel CVM question format with two related existing ones.
4. To determine the factors that cause variation between stated and actual WTP.
5. To investigate the factors that explain actual WTP.
6. To compare the inter-rater and test-retest reliability of the novel CVM question format to that of the bidding game and binary with follow-up format.

The study hypotheses were that: the CVM is both a reliable and a valid technique; there will be a convergence between stated and actual WTP for ITNs; and not all divergences between stated and actual WTP will be due to bias. Another hypothesis was that; the structured haggling technique will have the highest validity and reliability out of the three question formats in the study.

Section 5.1 focuses on the context of the study. Section 5.2 is devoted to qualitative aspects of the study. It presents the methods used for determining the study populations' general perceptions concerning malaria and their health seeking behaviour, as well as for establishing the content validity of the question formats. Sections 5.3 to 5.4 describe the survey methods. Section 5.3 presents the first survey on stated WTP, used for construct validity and inter-rater reliability tests, and describes means for establishing the content validity of the question formats, the scenario and questionnaire. Finally, section 5.4 deals with methods used in determining actual WTP for assessing test-retest reliability and criterion validity of the CVM and the three question formats. This section also presents methods used to assess divergences between stated and actual WTP, together with factors explaining actual WTP.

5.1 Context of the study

The general research procedures used will be presented in this section. Thus, the sub-sections will present descriptions of the study areas, sampling methods and sampling size. This section also outlines the structure of all the research methods, matching the specific empirical research objectives with research methods used, and present the schema for the fieldwork. The problem of malaria, use of ITNs and a description of the nature of the goods (ITNs and the re-treatment) to be valued are provided. Furthermore, the rationale for choosing the question formats and price-taking mechanism among the Igbos of Nigeria are presented here. The basic method for the development of the structured haggling technique and the nature of the goods to be used for the valuation are also presented. Finally, the general data management and criteria for comparing the three question formats are presented.

5.1.1 Study area

The study was conducted in three malaria holo-endemic villages namely; Amaetiti, Ahani and Enugu Akwu in Oji-river local government area of Enugu State, Southeast Nigeria. Enugu State is one of five states in the South Eastern part of Nigeria and shares borders with Abia State to the South, Ebonyi State to the East, Benue State to the Northeast, Kogi State to the Northwest and Anambra State to the West.

Though a land-locked State, Enugu (the capital city), is located approximately two and half driving hours away from Port Harcourt, Calabar and Warri, all of which are coastal cities with major shipping ports. Enugu is also located within an hour's drive from Onitsha and 2 hours' drive from Aba, both of which are trading centres in Nigeria (Enugu state government (ESG), 2000). The city is also located within 5 driving hours from Abuja and 7 driving hours from Lagos, the administrative and commercial headquarters of Nigeria respectively (ESG, 2000).

Lying partly within the semi-tropical rain forest belt of the south, the State spreads (with a land area of approximately 8727.1km^2) towards the north with its physical features changing gradually from tropical rain forest to open wood-land and then to the Savannah (ESG, 2000). The mean temperature in the state in the hottest month of February is about 36^{0}C (97^{0}F), while the minimum temperature of about 20^{0}C (68^{0}F) is recorded in November (ESG, 2000). The lowest rainfall of about 0.16cm^3 is normal in February, while the highest is about 35.7cm^3 in July (ESG, 2000).

The study villages form part of the twelve that comprise the Achi autonomous community, which is located 5 kilometres from the local government headquarters called Oji-River town and 45 kilometres from the state capital, Enugu. It has an estimated population of 45,000 people. The villages are divided into two broad groups called Achi-uno and Achi-agu, because of links to common ancestors. There are six villages in each sub-division, with the three villages used in the study located in Achi-agu.

Achi is linked to Oji-River town by a single lane road covered with asphalt. However, the road is filled with potholes and presents a formidable challenge to users during the rainy season. This road passes right through the town. There are no tarred roads within the villages, and dirt roads and bush paths provide means of access to the interiors. The Oji River runs very close to Achi, and actually forms one of the borders of one of the study villages, Enugu-akwu.

The people are of "Igbo" (pronounced eebow) ethnic group and Christianity is the major religion, with subsistence farming being their major source of livelihood. The Igbo people comprise the third largest ethnic group in Nigeria with a population of more than 20 million people (NPC, 1999). Most Igbo communities are densely populated and the people are renowned for their spirit of enterprise and business acumen (Basden, 1930). Achi has a traditional ruler called the "Igwe" who has a cabinet of traditional ministers called 'Ozo', 'Ichie' or 'Nze' as the case may be. In addition, these titled men are the leaders of their respective villages.

Ordinary mosquito nets are not sold in these villages, but they are sold in the urban towns like Enugu and Onitsha. It usually takes about one hour to reach Enugu and an extra 20 minutes to reach Onitsha by bus. This project was the first contact of the villagers in the study area with ITNs. The main local market (Ozu-Uda) is held once every 4 days (the Igbos have a traditional four-day week). However, there are very small daily markets that sell mostly vegetables, fish and food seasoning. Bargaining for goods is the norm in these markets.

Each of the study villages has a nursery and primary school. Achi has two post-primary schools; one private and one public. There is currently no electricity supply in the town and people use bush lanterns as source of light. However, a few rich households have electric power generating sets. There is no pipe-borne water and people get their water supplies from streams, the river, water collected from the rains and from water sellers.

There is a public general hospital in the town. There are also a number of health centres with one of them located in Ahani (one of the study villages). Three private hospitals/clinics and two maternity homes complement the public providers. There are a number of patent medicine stores in each of the study villages (Ahani has one, Amaetiti two, and Enugu-akwu four), however, it was not verified whether they were licensed or not. Itinerant drug providers also visit the community on the major market days. Numerous herbalists and other unorthodox health care providers (not using western medicine) abound.

The major malaria vector in Achi *is Anopheles Gambiense*, while Plasmodium falciparum causes more than 90% of all malaria cases (MVCU, 2000). These villages are also all hyperendemic for onchocerciasis (Okonkwo et al, 1991). There are two major seasons here; namely the Rainy season from May to October, which is the period of increased mosquito and malaria incidence; and the Dry season from November to April.

5.1.3 Nature of the goods to be valued: Insecticide-treated nets and re-treatment

Malaria is the number one cause of mortality and morbidity in Nigeria amongst all age groups (FMOH 1992). Malaria control has been integrated into primary health care (PHC) and service delivery comprises passive case detection and treatment, with no previous emphasis on prevention or the use of insecticide-treated nets (ITNs) prior to year 2001. However, since the third quarter of year 2001, the Federal Ministry of Health made the use of ITNs part of the national malaria control strategy.

The use of mosquito nets is not new in Nigeria, as they have been in use for many decades as door, window or bed nets. What is new is the use of ITNs. They have been proven to be effective tools for the control of malaria, especially with childhood malaria (Armstrong-Schellenberg et al., 2001; Abdulla et al., 2001; Curtis et al., 2000; Binka et al., 1996; Nevill et al., 1996; D'Alessandro et al., 1995a,b). Using World Bank guidelines, Goodman et al. (2001) found the ITNs to be cost-effective.

ITNs and their re-treatment may be assumed to be normal and private goods, but with potential positive and negative externalities depending on the extent of use and coverage in the communities (due to the potential mass killing effect of the nets on mosquitoes). However, the evidence is not clear on these factors and the nets are hence treated as private goods in this thesis. Nevertheless, the ITNs may qualify as merit goods because of their public health significance.

ITNs are different from ordinary nets in that they are treated with insecticides, and thus will be perceived to be better than untreated nets (Njunwa et al., 1993; Winch and Makemba, 1993; Premji et al., 1995; and Lines, 1996). Also, the facts that they kill mosquitoes and other insects like bedbugs, as well as reducing the mortality and morbidity due to malaria, are the qualities the consumers will be paying for. However, the individuals may not necessarily perceive these other benefits unless they are informed about them. Nevertheless, the need for regular re-treatment for the nets so that they do not become ordinary nets will be explained to the respondents in the scenario before WTP elicitation, as this factor may alter how they view the usefulness of the ITNs.

The Nigerian government in its recent quest to "Roll back malaria" has just launched a five-year strategic plan and wants to assure a high coverage of the population with ITNs on a user fee basis. UNICEF is currently pioneering the sales of ITNs in the country. The distribution is through the public-sector on a user fee basis, at the prevailing market price-rate at any point in time. However, preliminary reports are that there is very low numbers of nets sold and the reasons for these are not known.

5.1.4 Price-taking among the Igbos

The contingent valuation approach can be used to determine the willingness to pay for ITNs. A methodological difficulty however, is determining the question format to use in order to reliably and validly measure WTP for ITNs, as existing formats may not be suitable for a sub-Saharan African setting like Nigeria. None of the question formats closely mimic price-taking among the Igbos as described in the following paragraphs. The usual price-taking system among the Igbo people of Nigeria is based on a bargaining system, which currently applies to the purchase of untreated nets and other mosquito preventive tools in the market.

The mechanics of price-taking among the "Igbo" was described succinctly in a classic study by Basden (1938). The "Igbo" according to Basden is an inveterate bargain-hunter. "The vendor normally states the first price. The initial figure s/he quotes is always considerably in excess of the price s/he will eventually be content to accept and the first offer on the part of the prospective buyer is well below the sum he or she will ultimately pay. Occasionally, the two will come to mutually acceptable terms fairly quickly; usually, a liberal allowance of time and patience is absorbed. Depending on the tenacity of the seller and buyer, this bargaining process may go on for either a few or several times before a final bargain is struck after which money and goods normally exchange hands" (Basden, 1938).

Sometimes after the bargain is struck, the buyer may defer purchase until another day, or the seller may either not have the quantity the buyer wishes to purchase or the exact specifications of the buyer. On the new day, if things have changed, the seller can decide on a new price, which the consumer may or may not agree to. However, the consumer may decide on the new day to bargain further for a lower price than mutually agreed initially due to lack of money, more information about the product, knowledge of its price from other sources, reduced price of substitutes, reduced perceived need for

the product etc. If the consumer does not have enough money to completely purchase the good, s/he may make a small deposit to signal his/her commitment, which the seller may or may not accept.

The bargaining process described by Basden will surely apply to ITNs when they become available widely in the open markets and shops. However, fixed prices are operational in supermarkets and some shops mostly in the cities, mostly catering for a few select members of the population. Even then, all their customers still go to the open markets where bargaining reigns supreme. Hence, bargaining is the dominant price-taking mechanism, embraced in more than 90% of market transactions in Nigeria.

Haggling price-taking system has been described as repeated bargaining and also occurs in all parts of Nigeria (Balogun, 1991), in many sub-Saharan African countries and in some contexts in western countries. It has been found that females are better buyers at bargaining than men, because they ended-up getting lower prices for goods (Ojediran, 1988; Agbonifoh, 1985). However, Balogun (1991) found that males were better.

An alternative form to the usual process of bargaining occurs when the buyer initiates the bargaining as occurs in parts of northern Nigeria. In this case, the buyer initiates the haggling process by quoting a very low price for the product. The buyer initiative could either be from prompting by the seller or from the buyer himself. Thereafter, the seller will quote a high price and the whole process as described above will then be repeated.

5.1.5 Choice of the question formats for comparison

The major criterion for the question formats to be compared in this study was that they must somehow mimic normal market purchase decisions in Nigeria. Of existing techniques, the bidding game and binary with follow-through techniques were chosen. They had positive predictive validity of 85% and 75% in a pilot study conducted in Nigeria (Onwujekwe 2001) and the bidding method is actually the most frequently used question format in CVM studies in Africa (Onwujekwe and Nwagbo, 2001). The last question format was the novel structured haggling technique that was developed for this thesis.

The "structured haggling technique" derives its name from the usual haggling behaviour that is common between buyers and sellers of goods and services. However, it has been structured for use in eliciting maximum willingness to pay amounts. It resembles the bidding game and the binary with follow-through format. Differences are that it has more steps to mimic the haggling process, so that the respondents that are willing to pay will be coaxed to state the possible highest amounts they can pay. The detailed final haggling structure used in the study will be presented with the results of content validity in Chapter 6. This is because tests of content validity through market observations, interviews and focus group discussions (FGDs) were used to produce the final version of the structured haggling technique used in the study.

The development of the structured haggling technique involved a review of the existing literature on haggling and CVM question formats. Informal consultations were held with colleagues on price-taking in Nigerian markets, together with best approaches for designing a good CVM question format. These were combined with the investigator's personal experience with bargaining for goods and services to produce the basic structured haggling concept. The compensating variation perspective was taken in measuring willingness to pay and an ex-post perspective taken in framing the scenario, as this is the context of the study.

The open-ended technique was not chosen because it does not reflect the market situation and is applicable only in donations or voluntary contributions. Thus, it can be used to measure altruistic willingness to pay. Also, the payment card does not reflect how people buy goods as goods do not have multiple prices for people to choose the prices they fancy. Although there are alternative prices, there is more than one seller and a single seller will not present different alternative prices and ask the buyer to choose one price. This is because the buyer will most likely choose the lowest price when faced with making a real purchase decision.

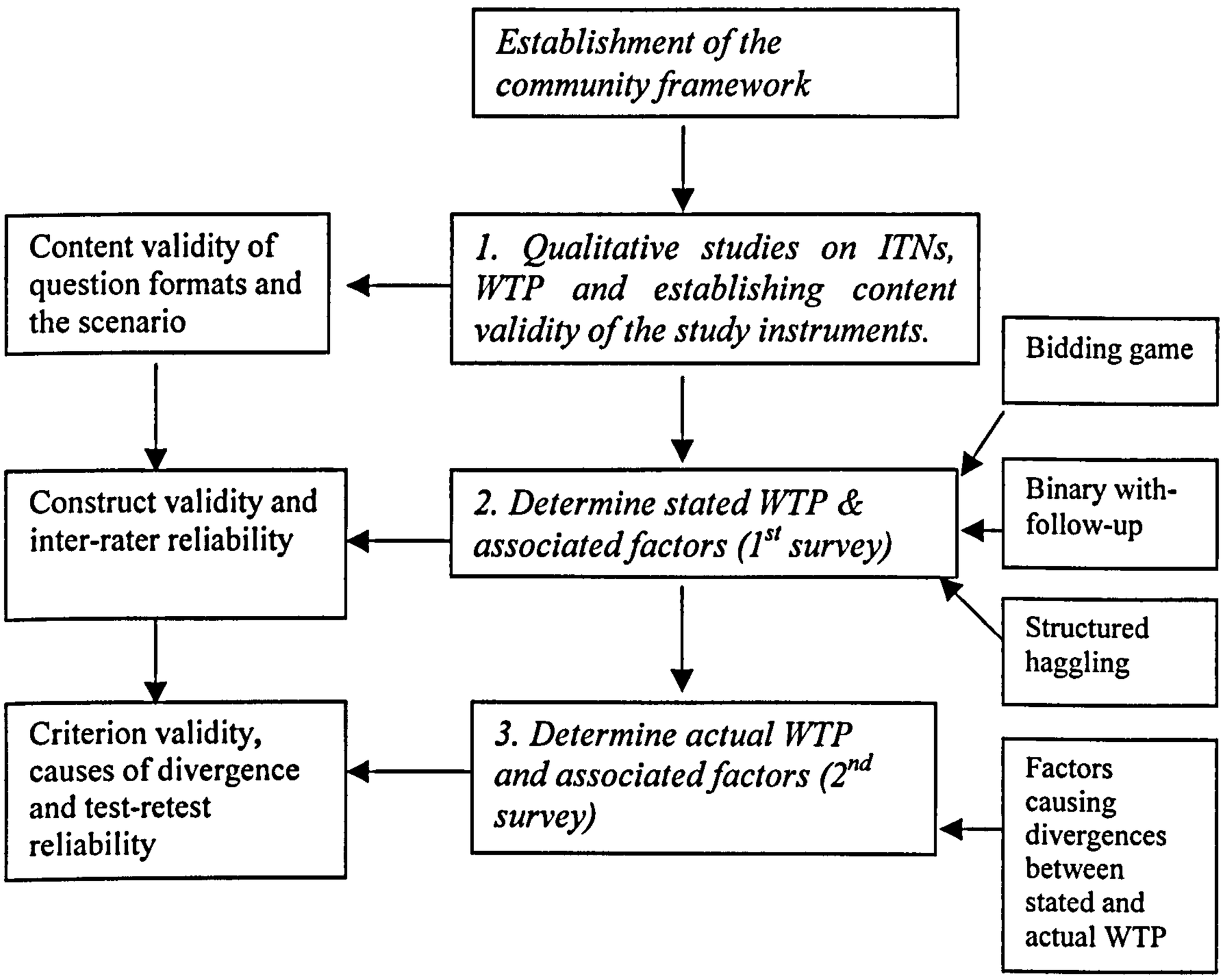
Finally, the binary system was not included as most Nigerians do not take or leave any price they see in the market, as price negotiations must take place. Balogun (1991) pointed out the cultural aspect of the bargaining behaviour of Nigerians "as being different from the European or American pricing system which usually has fixed prices in open markets".

5.1.6 The study outline

The empirical study was designed so that the specific objectives of designing a novel question format and of determining the reliability and validity of the WTP technique and the three question formats are met.

Figure 5.1 shows the flow of methods and data used to achieve the objectives, and the steps in its middle are described in detail in sub-sections 5.2 (Qualitative study), 5.3 (First survey) and 5.4 (Second survey).

Figure 5.1 Flow of methods used in the empirical study



Description of the various steps shown in figure 5.1

Establishment of community framework and mini-census

A community system to facilitate the study was established in each village. Formal meetings were held with the traditional ruler of Achi-agu, together with leaders in the three villages. They were briefed about the study and their consent obtained for the study to be conducted. They also nominated respondents for the baseline focus group discussions, together with candidates that could be the field workers. The venues for meeting and training the field workers were selected, and the schedule of community-based activities drawn up. A mini census to establish the household list used for the sampling was done, using educated community residents as the enumerators.

Qualitative studies

Baseline focus group discussions (FGDs) with different groups of men, women and youths were held in each village, after the mini-census, but before the surveys started. These FGDs were used to elucidate variables to be included in the questionnaires and to provide information to be used for developing the scenario for WTP elicitation, by eliciting information on preferred sales and financing mechanisms for the community-based sales of the nets. The means of encouraging truth-telling in the surveys were also determined. Furthermore, information on perceptions and expenditures on malaria and its prevention that are needed to provide the village perspective of the study and the WTP results to be generated were also considered.

Further qualitative studies involving market observations, interviews with traders and FGDs with villagers were used to establish and improve the content validity of the scenario, question formats to be used in the surveys and the questionnaires. The findings were used to produce the final design of the structured haggling technique, modify the binary with follow-up technique and finalise the design of the study instruments.

First survey

This involved interviewer-administered questionnaires to determine the stated WTP of randomly selected respondents for the ITNs and re-treatment services. The three different question formats were applied in sub-sections of each village. In addition, the supervisors also interviewed a sub-sample of the respondents for the test of inter-rater reliability.

Second survey

This involved the determination of actual WTP through sales of the ITNs to the respondents. However, a quantitative survey was also conducted using a modified version of the original questionnaire so that the factors explaining actual WTP were determined. In addition, the survey was used to discover factors that can cause a variation between stated and actual WTP. The WTP findings were then used to test for test-retest reliability, criterion validity, and the factors responsible for the change in WTP.

Survey data management

Data from the household interviews were entered into the computer on the same week the interviews were conducted and initially processed using EPI info software. There was manual checking of the data immediately in the field by the supervisors on the day of collection. I additionally re-checked all the questionnaires and questionable ones not detected by the supervisor were returned to the field and the interview re-conducted by the supervisor. Such questionnaires were excluded from the analysis. Summary statistics were continually compiled during the course of fieldwork, and discrepancies in the data and problems with the survey implementation detected and solved. STATA and SPSS software packages were used for the analysis.

Comparing the three question formats

The comparison of the appropriateness and performance of the question formats will use all the validity and reliability measurements. The criteria are: (1) Content validity: from the collective judgement of the traders and consumers. (2) Construct Validity: from numbers of hypothesis obeying statistically significant variables of estimates of WTP for ITNs and re-treatment respectively (3) Criterion validity: Relative scores on the phi coefficient. (4) Test-retest reliability: Pearson's correlation coefficient on WTP for personal nets, nets for others and net-re-treatment. (5) Inter-rater reliability: Pearson's correlation coefficient on WTP for personal nets and nets for other household members. The relative reliability and criterion validity of the WTP would be established using the criteria described in Chapter 3.

5.2 Qualitative Research

Qualitative research was comprised of two parts. The first part consisted of the baseline FGDs on the perceptions and practice of the villagers to malaria and its control. The second part was geared towards helping the development of the structured haggling technique and assessing the content validity of the different question formats together with the scenario.

5.2.1 Assessing perceptions of, attitude to and practice concerning control of malaria

An important objective of the focus group discussions (FGDs) was to elicit variables that would be included in the questionnaires in addition to informing community members about the aims, methodology and benefits of the study.

Three FGDs were held in each of the three villages, with separate groups of men, women and youths. The number of participants per FGD ranged from six to nine people and they were purposively selected so that all sections of the study areas were represented. The FGDs were held in central locations in each village. I was the moderator, while my research assistant recorded the discussions by taking notes. Each session lasted for about 2 hours. We started by introducing ourselves to the discussants. The participants then introduced themselves, by name, occupation and the section of the villages they were from.

The themes of the FGDs were introduced one after another by the moderator. These points focused on: determining the most common diseases of the people; local names of malaria; the different types of malaria, its symptoms and causes; perception of mosquito nuisance; peoples' health care seeking behaviour, mosquito control effort; and approaches to the control of malaria.

Discussions were also held on methods for eliciting truthful willingness to pay amounts from the people, the modality for payment and community-based sales of the ITNs. However, before discussing the modes of eliciting truthful WTP, the moderator explained to the participants the scientific cause of malaria, the symptoms, treatment and prevention. This was followed by a brief explanation of what ITNs are, before the people were asked how truthful WTP for ITNs could be determined. We concluded by thanking the participants who responded with a vote of thanks. The guide for the FGDs is attached as Appendix 4.

Data analysis and result

The records of the discussions were transcribed on the same day the FGDs were held. Content analysis was used to categorise responses into domains representing common themes. The themes were: common diseases in the villages; awareness about malaria; classification, symptoms, health seeking and prevention of malaria; knowledge and attitude to ITNs as well as whether people would be willing to pay for them; best ways for determining WTP in order to generate truthful responses; possible causes of divergence between stated and actual WTP; and peoples' preferred payment and distribution mechanisms during ITNs sales. The result is presented in Appendix 5.

5.2.2 Testing and improving the content validity of the question formats

The detailed steps taken in this study for determining content validity, which were market observations, interviews with traders and FGDs with consumers, will be described in the following sub-sections. The interviews and FGDs were also used to improve the scenario. Purposive sampling was used to select the respondents for the interviews and participants in the FGDs.

Market observation

Observation of how traders actually sold untreated mosquito nets (ITNs are not yet available in the markets) was conducted at the Mosquito-nets Trading Line; Main market, Onitsha[17] urban town of Nigeria. I sat by the shops and observed how buyers and sellers interacted and bargained for prices. This was to learn how customers were attracted to individual shops, the initiation of the haggling process and how the final prices were agreed upon. Other aims were to record the average number of haggling

[17] Onitsha is about 80 kilometers from the study villages and is located in a neighbouring state. However, Onitsha is also an Igbo town, and similar market behaviours obtain between buyers and traders in the town and in the study villages.

iterations it took before final prices were agreed upon, differences in bargaining between males and females and how the goods finally exchanged hands between buyers and sellers. The observation took place over 3 days.

Interview with traders of mosquito nets

Interviews were held with seven purposively selected sellers of untreated mosquito nets and focused on the question formats and the scenario (see Appendix 6 for the interview guide). The basic structured haggling technique together with bidding game and binary with follow-up method were presented to the respondents. These interviews determined what the traders thought about the scenario and the question formats, especially in terms of resembling actual processes in the market for arriving at final prices. They also determined the extent to which the question formats would capture the price consumers would actually pay, and how they could be improved to resemble the normal price-taking processes for mosquito nets and similar goods.

FGD with consumers

A focus group discussion was held with a group of purposively selected eight consumers (a mixture of four men and women). The aim and discussion points for the focus group, held in the study district were similar to the questions posed to mosquito nets sellers (Appendix 6).

5.3 The first survey: Determination of stated willingness to pay

The first survey was used to generate data for determining the construct validity and inter-rater reliability of the question formats. The data generated were used in conjunction with those from the second survey to determine criterion validity, divergences in WTP and test-retest reliability. The individual steps taken in this first survey are described below.

5.3.1 Development and structure of the questionnaire

Variables were identified through FGDs and the literature review, including their possible hypothesised relationship with WTP. The questionnaire was divided into four parts. Section A was devoted to personal data about the respondent and general information about the household. The second section (B) focused on the perceptions of, recent experiences with and expenditures on malaria. This section directly preceded that where WTP was determined. It served to get the respondents to start thinking seriously about malaria and mosquitoes, as according to Gonzalez-Caban and Loomis (1997), before asking how much respondents would pay for the good, it is important to allow them an opportunity to reflect on why it may be important to them.

In Section C, the WTP scenario and question formats were administered. The scenario included showing the respondents an ITN. The tests for content validity were used to decide on the starting-points for the three question formats. A follow-up check question was asked after the WTP questions in the case of ITNs and re-treatment respectively to those who were not willing to pay, in order to find-out whether they really were not willing to pay or whether they were protesting at being asked their willingness to pay. Mitchell and Carson (1989) posited that the protest responses are usually not considered valid representations of the individual's WTP or benefits, and are normally not included when WTP is computed. The protest responses only included those who stated that they were not willing to pay because they felt that it was the duty of the government to provide the ITNs.

Section D of the questionnaire was used to determine the households' experiences of and expenditures on other illnesses and proxies for income (food expenditures and the value of home produced and consumed food). The last part determined households' assets (Appendices 7 and 8 contain the Igbo and English versions of the questionnaire, illustrated using the bidding game method). The BWFU and Structured Haggling methods are presented in the result section under content validity in Chapter 6, section 6.1.3.

5.3.2 Variables

The variables included in the questionnaire, their measurements and hypothesised theoretical relationships with willingness to pay for ITNs are shown in Table 5.1. The variables were selected based on the literature review, the focus group discussions and considerations of demand theory.

The variables were the WTP estimates, socio-economic variables like status of the respondent in the household, the number of people living in a household and their age composition, the sex, age, occupation, level of education and marital status of the respondent. As most respondents are usually reluctant to provide information on household income (Whittington et al., 1992; Onwujekwe, 1999), the food cost of a household was used as the proxy for income since most households spend more than 50% of their income on food (Onwujekwe, 1999). This was calculated as the monetised value on home produced and consumed food, together with the expenditures on food consumed.

Malaria-specific variables were the households' perceived risk of contracting malaria, whether or not malaria was present in the household within the past month prior to the interview and the average monthly household expenditure to treat and prevent malaria respectively.

Data on ownership of movable assets were also collected to shed light on household wealth. In the same vein, Whittington et al. (1992) in their study among the Igbos of Nigeria relied on indirect measures of households' ability to pay such as asset variable, the housing quality dummy, and the measure of storage capacity, because they were unsuccessful in collecting income information.

Table 5.1 The independent variables for stated WTP

Variables	Measurement	Hypotheses (relationship with WTP)
Status in the household	1 = household head 0 = otherwise	Household heads being the ones that control the purse strings will be more willing to pay than others
Number of household residents	Continuos quantitative measure	The more the number, the more willing a household will be, because they will have more resources to pool
Sex	1 = male 0 = female	Males will be more willing to pay since they usually control purse strings and are the key decision makers
Age	Continuous measure in years	The active workforce will be more willing to pay since they have more money, thus the under 60s will have more willing to pay
Number of years of formal schooling	Ditto	Education increases level of awareness. Thus increased education will lead to increased WTP
Marital status	1 = married 0 = never married	Married people will be more willing to pay, because they'll like to protect their spouses and children
Perceived incidence of malaria	1=perceived occurrence 0 = perceives no occurrence	The more the perceived incidence, the more willing to pay
Actual incidence of malaria	Continuous data on occurrence of malaria in last month	The more the actual incidence, the more the WTP
Last month's expenditure to treat malaria	Continuous monetary measure	The more the expenditure, the more the WTP
Dummy of maltreat	1 = made expenditures 0 = made no expenditure	Ditto
Last month's expenditure to prevent malaria	Continuous monetary measure	Ditto
Dummy of exp. to prevent malaria.	1 = made expenditures 0 = made no expenditures	Ditto
Actual incidence of other illnesses	Continuous count o occurrence of other illnesses in last month	This may lead to depletion of household resources, thereby decreasing WTP
Last month's expenditure to treat other illnesses	Continuous monetary measure	Ditto
Dummy of exp. to treat other illnesses	1 = made expenditures 0 = made no expenditure	Ditto
Previous purchase of nets	1 = ever purchased any type of net 0 = otherwise	People who have ever purchased nets will be more willing to pay since they had earlier decided that they are important
Ownership of radio	1 = household owns a radio 0 = otherwise	It shows enhanced socio-economic status and will lead to increased WTP
Ownership of bicycle	1 = household owns a bicycle 0 = otherwise	It shows enhanced socio-economic status and will lead to increased WTP
Ownership of grinding machine	1 = owns a grinding machine 0 = 0therwise	It shows enhanced socio-economic status and will lead to increased WTP
Ownership of motorcycle	1 = owns a motorcycle 0 = otherwise	It shows enhanced socio-economic status and will lead to increased WTP
Ownership of car	1 = owns a car 0 = otherwise	It shows enhanced socio-economic status and will lead to increased WTP
Last week's Food cost (expenditure + home produced)	Continuous monetary measure	This is the proxy for income. Thus, the more the food cost, the more WTP
Occupational group1	1=unemployed/unskilled lab. 0 = otherwise	This group will have negative WTP since they don't have money
Occupational group2	1 = farmer 0 = otherwise	They will have negative WTP because of lack of disposable cash
Occupational group3	1 = skilled labourer/trading 0 = otherwise	They will have a positive WTP since they have disposable cash
Occupational group4	1 = Formally employed 0 = otherwise	They will have a positive WTP since they have disposable cash
Occupational group5	1 = professionals & mid/big time business 0 = otherwise	They will have a positive WTP since they have disposable cash

5.3.3 Recruitment and training of field workers

The field workers and their supervisors were selected from within the study villages. An inclusion criterion was that all interviewers must have at least post-primary education, and should be able to read and write the Igbo language very well. In the training, we started with many field workers, but some were eliminated because they could not grasp the subject matter or were not able to read the Igbo language well enough. The final questionnaires used for study were in the Igbo language.

It was necessary to use local field workers since they were familiar with the households in their villages and were more likely to elicit truthful answers than outsiders. It was felt that the respondents would be less suspicious of the motives behind the questions, in contrast with an external person, where they may have suspected that the exercise was for tax purposes. Furthermore, asking people sensitive questions about food expenditures and assets require some degree of openness and familiarity between interviewee and interviewer, as that will help assure the interviewee of confidentiality and the interviewers are better positioned to detect false responses.

Each of the field workers mastered only one question format. This served to create expertise and ensured that the interviews were well conducted. O'Brien et al (1998) used a similar system, where their interviewer-administered questionnaire interviews were undertaken by six interviewers, each working within a specific stratum of experimental design in order to minimise confusion. However, the supervisors mastered all question formats so that they could conduct the second interviews for determining inter-rater reliability.

The field workers were trained on the initial design of the questionnaire. However, they participated in the pre and pilot testing and were able to help improve and appreciate the changes made. This also ensured that they had a very good mastery of the questionnaires.

5.3.4 Pre-testing and piloting

Pre-testing of the question formats was conducted with consumers and traders. Pilot tests of the questionnaires were undertaken with 120 residents of a neighbouring village to the study area. Some of the wording, content and coding of responses were modified from results of the pilot testing.

5.3.5 Translation of the questionnaire to Igbo

The result of the pre-test was used to modify the questionnaires and then translate them into Igbo. Two people versed in the Igbo language, working independently, translated the questionnaires into Igbo. A meeting was held with both translators to clarify areas of disparity that were noticed. Then, the consensus Igbo version was read to some selected people from the villages to ensure that all words and expressions were in the local dialect. This feedback was then used to produce the penultimate version of the questionnaire used for field-work. The final version was produced after the pilot-testing.

5.3.6 Sampling for construct validity

The sample size was calculated based on the formula for a population survey that will be based on random sampling. The EPI info 6 software package was used for the calculation. The inputs to the sample size calculation were; the combined average of 2000 households for the three villages, 95% confidence level, power of 80% and 76% true positive (positive predictive validity) from results of a previous study by Onwujekwe et al, (2001) and the worst acceptable percentage of 71% of true positive.

Thus, the calculated sample size per community was 246 households per question format. However, since all question formats would be used in each village, 300 households were selected from each of them to account for refusals and non-usable questionnaires. This was informed by the previous study by Onwujekwe et al. (2001) where 7% of the questionnaires were unusable for analysis. A mini-census was conducted in the villages to produce the household list that served as the sampling frame. Systematic random sampling was then used to select the households, by including every 2nd household in the household list in the study. The household heads or their representatives (if the household head was not available) from the selected household were interviewed.

5.3.7 Sampling for inter-rater reliability

There is no existing consensus on what method to use in calculating sample size for reliability studies. Nunnally (1978) recommended that a study should involve at least 300 subjects, while Guilford (1950) and Kline (1986) recommended 200 subjects. However, Streiner and Norman, (1995) argued that a standard graph could be used accurately for most purposes to determine the sample size for testing reliability. Cohen (1987) recommended that for a two-tailed test, with an alpha of 0.05, an effect size of 0.30, one needs 84 subject for a power of 0.80.

However, the sample size in this study was determined based on the reliability coefficients of previous studies that have used the bidding game. Those studies found a test retest reliability of 0.66 (O'Brien and Viramontes, 1994) and 0.80 (Flowers et al., 1997). A similar system as was used to determine the sample size for construct validity using the EPI info 6 programme was also used here.

The inputs were: an average of 2000 households in the three villages, at a 95% confidence level and the power of 0.80, the minimum of 85 respondents were needed for the test of reliability. This was considering an expected reliability coefficients of 0.75 (5% below Flowers et al.) and the worst acceptable coefficient of 0.66 (as found by O'Brien and Viramontes, 1994). The determined sample size was similar to the number recommended by Cohen (1987).

In order to take care of refusals and non-usable questionnaires, 150 respondents were selected in each village. The respondents for the second survey in the test of inter-rater reliability were selected using systematic random sampling by including every 2nd household in the big sample of 300 respondents from each village.

5.3.8 Sub-division of the samples from each village and conduct of interviews

The sample from each village was divided into three equal parts and a question format applied to a third of each village. This was to control for differences in villages' characteristics affecting the comparative performance of the three question formats.

5.3.9 The interviews

Interviewer-administered questionnaires were applied to the selected household heads or their representatives (if the household head was not available). Ramsey et al (1997) noted that interviewer-administered surveys are time consuming, resource intensive, and subject to variability in the interviewer training and experience. Nevertheless, face-to-face interviews have been recommended for contingent valuation surveys (Mitchell and Carson, 1989; NOAA, 1993) to minimise problems with non-interviewer-administered techniques such as low response rate and inability of the respondents to fully understand the scenario and the question format. Our detailed and lengthy training ensured that the misgivings raised by Ramsey et al. (1997) regarding variable interviewer performance in the use of face-to-face interviews did not arise.

5.3.10 Special interviews for determining inter-rater reliability

The second interview by different interviewers using exactly the same questionnaire took place after a maximum of five days, instead of the originally planned maximum interval of three days. This was because some respondents were either not available or needed some time to recover from the first survey (it took an average of 30 minutes to complete) before submitting themselves to another.

The sampling for the sub-sample was done prior to commencement of the entire survey as already described. However, the first interviewers informed the second interviewers about the people they had interviewed the previous day and if the person's name was on the second interviewers' list, then they tracked the person down for the interview.

5.3.11 Encouraging truth-telling during the conduct of the stated WTP surveys

A number of measures were taken to encourage truth-telling. Payment cards (proxy for promissory notes) were issued to all respondents willing to pay for the nets immediately after the survey. Their names were also written down in registers in their presence. They were also informed that they have been registered to buy the ITNs and should bring the cards and first instalment payments when the nets arrive. They were allowed to change their maximum WTP bids at the next stage if they so wished. This was to encourage them to state truthful answers to the WTP questions.

5.5 The second survey and sales of ITNs

The two main objectives were to see whether people would do what they said earlier in the stated WTP survey and the second was to compare how the different question formats used predicted actual WTP. The survey together with the sales were used for test of criterion validity and test-retest reliability together with determining causes of divergence between actual and stated WTP. In addition, the factors that would explain actual WTP were determined.

Development of questionnaire

The survey instrument was expanded to include variables that could cause a divergence between stated and actual WTP. These variables were arrived at from the qualitative studies. The questionnaire was pre-tested during the content validation of the study tools. The Igbo and English versions of the questionnaire are included as Appendices 9 and 10.

Sampling for criterion validity and test retest reliability

No sampling was done, as all the original respondents seen were interviewed. The nets were offered for sale to all the respondents whether or not they were willing to pay positive WTP amounts during the first survey.

Sales of ITNs

The nets were offered for sale to the respondents within one month of the survey. The respondents that were available were first interviewed with the questionnaire before they purchased the nets. The sales were also open to all community members. However, the net sellers were also the interviewers during the initial survey and therefore they were able to take note of whether the person buying had been interviewed in the first survey.

The respondents' households were informed of the arrival of the ITNs, orally and in writing. Also, town criers made announcements in the various villages for anybody willing to buy the nets to do so. Sales were conducted in central locations in the villages and in the interviewers'/sellers' homes. However, the sellers were permitted to take the nets to the homes of people who asked. The supervisors confirmed that people that bought the ITNs were the real owners of the nets and were not surrogate buyers.

The price at which a net was offered for sale was 350 Naira ($3.5). (US$1 = 100Naira). A fortnightly instalment payment system was used but consumers had the option to pay all at once if so desired. Ideally, the price would have been set at the level of median WTP, so that there would be a larger sample of those whose stated WTP were up to the sale price. However, this was not possible since the government had fixed 450 Naira ($4.5) as the cost of a net, and the officials were not prepared to reduce the price for the study. This was because they felt that it would distort the market for the nets. Moreover, distributions of the stated WTP responses showed that less than 5% of the people stated WTP amounts up to 450 Naira. Thus, selling at the government rate would have meant the inability to determine criterion validity, especially the positive predictive validity aspect. So the decision was made to sell the nets at 350 Naira, with the project subsidising the cost of the nets. The 100 subsidy was the highest the project could offer because of the very limited funds available for the fieldwork.

The net sellers and the respondents were not aware that there was a subsidy (since there was no local market), thus no potential distortions in behaviours were experienced. Some other studies that have determined criterion validity offered the goods at prices that were lower than the market price. These include Blumenschein et al. (1998) who offered the sunglasses at $1.00 and $5.00 while the market price was $9.00 and Cummings et al. (1995) who offered two of the goods they used at lower prices than obtained in shops.

Questionnaire survey to determine levels of WTP

If a respondent wanted to buy an ITN, a retest interview was requested prior to a purchase. If a member of a respondent's household wanted to buy an ITN, an interview with the original respondent was requested. If the person was at home[18] an interview was conducted (but the net was sold in advance of this interview) but if the person was not at home, the interview was not done. At the end of the net sale period, interviewers sought out the respondents who had not made contact with the interviewers/ITN sellers, and interviewed those that agreed to be interviewed.

[18] Which would have been, on average, a 15 minute walk away

5.6 Data Analysis

All the methods used to analyse the qualitative work and the first and second surveys are presented in this section. The first three parts detail the data analytic tools used for investigating content, construct and criterion validity. The section on criterion validity is further used to analyse factors explaining actual WTP and comparing them across the three question formats. In the fourth section, the analyses for inter-rater and test-retest reliability are provided. The final section is devoted to the methods used to investigate the possible causes of divergence between stated and actual WTP, together with stated WTP in the first survey and stated WTP at the retest.

5.6.1 Data analysis for content validity

Frequency tabulations were used for the quantitative data from the market observations. Content analysis was used to analyse the other purely qualitative data. The responses during the interviews and FGDs were divided into domains on the content validity of each question format and the scenario.

5.6.2 Data analysis for construct validity of ITNs and re-treatment

Descriptive statistics

The data for each question format was pooled across the three villages. The demographic and other characteristics of the respondents were tabulated and compared across the three question formats' groups using Kruskal-Wallis statistics, to see how similar the groups were. The mean and median values of the WTP estimates were generated, together with the confidence intervals around them and compared across the three groups. In addition, the Box-plots of all the continuous variables were plotted to show their distributions, to help inform the choice of functional forms for testing construct validity. Data transformations into several functional forms for non-normally distributed WTP and other variables were then done so that the best linear relationships were selected for the tests of construct validity.

Construct validity of ITNs

The construct validity of ITNs (for own nets) was determined by econometric modelling of WTP and the independent variables. These variables, their measurements and hypothesised theoretical relationships with WTP were set out in Table 5.1. Specification searches employed included determining the best models for limited dependent variables that resulted from the WTP for ITNs in three question formats. However, there was also specification search for only the positive WTP for ITNs responses. Diagnostic tests for heteroscedasticity, multi-collinearity, functional mis-specification and normality were conducted.

In the following parts of this section, the methods for construct validity tests for WTP for ITNs will be presented. I first describe the rationale for the choice of the models. This is followed by description of the variables used for the econometric modelling together with their hypothesised relationships with WTP for ITNs. The model specification begins by the determination of the nature of the zero WTP for ITNs, as this shed light on whether they were true or false zeros. This is followed by a description of the modelling techniques for limited dependent variables and for the positive responses alone. The processes used to compare the different models and different question formats would be spelt-out.

Choice of model(s) to use for determining the construct validity of WTP

The Tobit, Heckman selection model estimated by the Maximum Likelihood estimator and log OLS were compared for the test of construct validity of the full sample WTP. The OLS and truncated regression models were compared for the positive WTP.

If the dependent variable is limited in some way, OLS estimates are biased even asymptotically (Kennedy, 1998). OLS in this case fails to account for the qualitative differences between the limit observations (those with zero WTP) and the non-limit observations (Donaldson et al., 1998). Omitting the limit observations creates bias. Ignoring them will be throwing away information, but including them as though they were ordinary observations also creates bias (Kennedy, 1998). However, Duan et al. (1993) argued that if there is no selectivity problem; in WTP terms; if the decision to state a positive amount WTP is not nested with the level of WTP, then log OLS is the appropriate model.

There are several alternative models for limited dependent variables and the easiest to estimate is the Tobit model (Kennedy, 1998). Donaldson et al. (1998) also recommended the Heckman selection model depending on the underlying assumptions behind the zero responses. They posited that if the zero observations reflect genuine WTP values of zero, the Tobit model is appropriate for the estimation. However, if the zero responses have other than a genuine zero WTP, ranging from reporting errors to protest responses, the Heckman model is better since a more flexible specification of the censoring mechanism is required in that case.

The choice of model will be influenced by the use to which the results are put. Thus, since the use of WTP data is two-pronged; benefit estimation and pricing, it will also be necessary to determine the appropriate model in each case. In this vein, benefit estimation requires societal valuation of the good in question and any appropriate model, has to include the zero values. However, in pricing, one may only need the information about positive WTP since the interest could be on only those who would demand the good and the econometric models then may not include the zero values. This makes the OLS appropriate in such cases.

In the case of modelling of positive WTP, a possible failing of the OLS estimation occurs with truncation of data. Such truncation at zero causes the lower tail of the distribution of the error term to also be truncated (Hoddinott, 1992). Thus, the mean of the true error term differs from the estimated error term (which by definition is equal to zero) (Hoddinott, 1992).

Whether truncated or OLS regression is more appropriate depends on the purpose of the estimation (STATA, 1985-1999). For instance, if one is interested in the mean of a woman's working hours conditional on the sub-sample of market labourers, OLS is appropriate. However, if it is the mean of the woman's working hours regardless of market or non-market labour status, OLS estimates could be seriously misleading (STATA, 1985-1999). Applied to pricing for ITNs, this means that OLS is appropriate if the interest is in the WTP of only those with positive WTP, while it may not be appropriate if one is interested in an unconditional sample of the respondents.

Brief description of the econometric models for limited dependent variables: These are the Tobit and Heckman models.

The Tobit model

Tobit estimation was originally developed by Tobin, (1958). Tobit models refer to regression models in which the range of dependent variable is constrained in some way (Amemiya, 1984). Tobit can also estimate models that are censored from above, together with models censored from both sides, the two-limit tobit (STATA manual ver 6; p197-8). The Tobit model is a special case of a more general model incorporating sample selection (Kennedy, 1998). A second equation called the selection equation determines whether an observation makes it into the sample. Selection causes the sample to be non-random, drawn from a sub-population of a wider population (Kennedy, 1998; p251). According to Kennedy (1998; p252), in Tobit model, the sample selection equation is the same as the equation being estimated, with a fixed, known limit determining what observations get into the sample.

The Heckman selection model

Heckman's model (Heckman, 1976) is a simultaneous equation model consisting of two equations (Amemiya, 1984). The first part of the Heckman procedure estimates the expected error and the second stage reruns the regression with the expected error as an extra explanatory variable (Kennedy, 1998). The expected value of the error term requires observations of the explanatory variables for the limit observations, so the Heckman procedure only works with censored data. According to Kennedy (1998; p 256), the Heckman procedure does not perform well relative to sub-sample OLS when the errors are not distributed normally and the sample size is small. Also when the amount of censoring is small, the correlation between the errors of the regression and selection equations is small and the degree of collinearity between the explanatory variables and selection equations is high. The variables in the selection equation should strongly affect the chances for observation, but not the outcome under study (STATA manual ver 6; p18). According to the manual, the command assumes that missing values of the dependent variable are unobserved (not selected).

Models' specification approach

Before the application of the specific models, the nature of the zero responses was first determined. Various data transformations into the natural logarithm and square root functions were done for the continuous variables before the final functional forms used were selected. Regression diagnostic tests were conducted and the best model used for the interpretation of the parameters and for comparing the different question formats.

Theoretical model

Based on consumer demand theory, it is hypothesised that a respondent's willingness to pay for ITNs (W_i) is a function of the personal and household socio-economic characteristics, malaria/mosquito specific variables and attributes of the nets.

$$W_i = a + F_iB + e_i \quad (1)$$

Where F_i is a vector of the respondent's and household's characteristics and the attributes of the good under valuation, a and B are parameters of the model, and e_i is a random term with a standard normal distribution. Equation 1 depends on the individuals' utility function

$$U_i = U(Y, P, N, Q) \quad (2)$$

With utility of the individual (U_i), monetary income (Y), the prices of other goods and services expressed as a vector (P), other demographic and economic factors that might influence ability to pay or constrain behaviour (N) and the character of the goods under valuation (Q).

Nature of the zero WTP responses

The nature of zero WTP responses was investigated to determine whether the zeros represented true zero responses and whether respondents that stated a positive WTP differed systematically with those that had zero WTP.

The first step taken was the creation of a dummy variable (WTPbin) with the value 0 if the respondent had a zero WTP and 1 if otherwise. Bivariate analysis between the dummy and a series of explanatory variables was undertaken using tests of differences in means and in proportions.

Logistic regression of WTPbin to shed more light on the groups' characteristics, and helped determine variables for inclusion in Heckman's selection equation. A general-to-specific modelling approach was used in order to arrive at the best model. The goodness of fit measures assessed were the log-likelihood, the chi-squared (which uses the log-likelihood to test the overall significance of the model) and the proportion of correct predictions made by the estimated model.

The independent variables with the smallest t-statistic, and whose removal did not adversely affect the other coefficients nor the prediction of the models were removed sequentially. The F-test for the hypothesis that the coefficient of that variable is zero was used to decide whether the variable would be finally dropped or re-entered into the regression. The variables were finally dropped if the probability associated with the F-test was more than 0.10.

Modelling of the estimators

Full and specific models were estimated using the approach in the logistic regression. However, since there is no existing guideline on the variables that should be in the Heckman's selection equation, the variables that were statistically significant in the logistic model, but not in OLS for positives were used as the differentiating variables. Thus, these variables are correlated with the decision to state a positive WTP, but not with the level of WTP. The other variables in the first part of the Heckman's model were those that were statistically significant in the OLS for positive model but not in the logistic model. Hence, these variables are not correlated with the decision to state a positive WTP, but with the level of WTP.

Regression diagnostics

Generalised residuals were used to check for heteroscedasticity and normality Donaldson et al., 1997a). The Ramsey RESET test was used to check for functional mis-specification. The RESET test is a t-statistic and the critical value is 1.96 at 5% level. Normality is measured as a chi-square and the critical value is 5.99 at the 5% level (Donaldson et al., 1998).

Scatter plots of the residuals versus the predicted values were used to check for heteroscedasticity. In addition, the Cook-Weisberg test for heteroscedasticity was also used for the OLS models. The numbers of statistically significant variables that obeyed hypotheses and the WTP values predicted by the models were used as additional criteria for comparing the econometric models and question formats. In the case of predicted WTP, models whose predictions were closer to the raw WTP were adjudged better on that criterion. All these tests were combined to determine the best econometric models and informed on the performance of the question formats.

Construct validity for ITNs re-treatment

Construct validity was determined by econometric modelling of WTP for ITNs re-treatment. The specification searches included determining the best OLS models since the data sets were neither censored nor truncated. The description of the variables used for the econometric modelling together with the hypothesised relationships with WTP for ITNs re-treatment were as for ITNs. The only addition was that the WTP for ITNs was added as an explanatory variable. The hypothesis was that the higher the WTP for ITNs, the higher the WTP for the re-treatment. The modelling approach and comparison of construct validity of the question formats mirrored that of WTP for ITNs.

Comparison of the econometric models and construct validity of the question formats

The regression diagnostics were used to compare the econometric models within each question format in order to determine how good they were for explaining WTP for ITNs using the same data set. Comparisons across question formats were used to determine the overall better performing models and question formats whose generated WTP, the models best fit. The better performing models for full sample and positive WTP were used for the final interpretation of the coefficients and to compare the relative construct validity of the different econometric models and question formats. The comparison of construct validity was achieved by determining the numbers of statistically significant variables that had correct hypothesised constructs from each question format.

Note that the modelling approach and comparison of construct validity of the question formats in determining the construct validity of WTP for re-treatment mirrored that of WTP, for ITNs. The only exception was that only the OLS model was used in this case since the data sets were neither censored nor truncated. Construct validity was also compared across the question formats to determine how they relatively performed. The number of statistically significant variables that obeyed the theoretical expectation was used as the basis for judging their relative construct validity.

5.6.3 Data analysis for criterion validity and determinants of actual WTP

Descriptive statistics

Tabulations and descriptive statistics of the respondents that actually bought the nets were computed. Statistically significant differences in the demographic characteristics of three question formats' groups were assessed using Kruskal-Wallis and Mann Whitney U tests.

Criterion validity

Criterion validity was determined using the phi correlation coefficient and positive predictive validity. The phi coefficient (Φ) was determined, together with the confidence intervals around it. In interpreting the phi coefficient, it was posited that the test will not be useful unless it correlates significantly with the criterion; and similarly, the higher the correlation, the more valid is this test for this particular criterion (Streiner and Norman, 1995). However, has been argued that even modest correlation (e.g. a correlation of 0.30) between test and criterion can prove quite useful for selection purposes (Nunally, 1978). The positive predictive validity of all the methods was also determined. Positive predictive validity means the number of respondents whose stated WTP were equal to, or more than the price of the nets that actually bought the nets.

Determinants of actual WTP

Analysis of determinants of actual WTP was used to compare the characteristics of those who bought the nets with those who did not, in order to know whether there were systematic factors guiding actual net purchases. These characteristics were also compared across the three WTP question formats' groups to see how similar they were.

A dummy variable called actual WTP was created with 1 representing a person who bought a net and 0 if otherwise. Tests for differences in means were used to see whether there were statistically significant differences in the socio-economic and stated WTP variables for the two groups. Then econometric modelling of the actual WTP with the independent variables was effected through logistic regression.

Most of the variables were as measured and hypothesised for stated WTP and are not repeated. Only the new variables that were additionally expected to explain actual WTP are included (see Table 5.2).

Table 5.2 Additional independent variables for actual WTP

Variables	Measurement	Hypotheses (relationship with WTP)
Stated WTP amount	Continuous monetary scale	The higher the stated WTP, the more the actual WTP
Distance to sales spot	Continuous scale measured in minutes	The more the distance, the less the actual WTP
Expenditures on unplanned activities	Continuous monetary scale	The more the expenditures on unplanned activities, the less the actual WTP

5.6.4 Data analysis to assess reliability

Inter-rater reliability

The data on inter-rater reliability was analysed using Pearson's correlation coefficient. The variables were; whether willing to pay, level of WTP for ITNs for personal use, level of WTP for other household members and level of WTP for re-treatment. The reliability coefficients of the demographic and asset variables were also determined and formed a check of the robustness of the reliability of the WTP measurements. This is because, if the two sets of interviewers' differed considerably in measuring the demographic characteristics and assets of the households, then the robustness of the WTP reliability coefficients may be in doubt.

Test-retest reliability

The data elicited during the first survey to determine hypothetical WTP and the second survey to also determine WTP at point of purchase was analysed using Pearson's correlation coefficient. The variables assessed were the level of WTP for personal nets, level of WTP for other household members and demographic factors. Only the willingness to pay for ITNs was assessed since there was not enough time to implement net re-treatment. The WTP amounts determined during the second survey were not more than 350 Naira (the sale price of the nets). Thus, for comparing the WTP from the first survey and the re-test, the WTP amounts in the first survey that were more than 350 Naira were uniformly reduced to 350 Naira.

5.6.5 Analysis of factors that can cause a divergence between levels of WTP in the first survey and the retest and between stated and actual WTP

It was recognised that there will be divergences between the elicited stated WTP in the first survey and stated WTP in the survey that was conducted with the sale of ITNs, together with divergences in stated WTP in the first survey and purchases of ITNs (actual WTP) in some instances. Understanding these factors is relevant for understanding the results of both criterion validity and test-retest reliability.

However, in criterion validity, divergent WTP referred to changes in WTP relative to actual purchase decisions, while in test-retest reliability, it refers to changes in levels of stated WTP between the two surveys. Thus, in criterion validity, the divergences are with respect to respondents that stated 350 Naira or more in the first survey, but that did not buy nets or respondents that stated less than 350 Naira but that bought nets. However, in test-retest reliability, divergences refer to cases where the levels of WTP stated in the first and retest survey are different.

The causes of the divergences were investigated using additional variables built in to the questionnaire in the second survey. However, the variables in the first survey were also examined for answers to causes of divergences. Furthermore, the ratios of divergence between levels of WTP in first survey and the retest in the three question formats were determined.

Variables

Most of the variables were as measured and hypothesised for stated WTP and are not repeated here. The variables presented in the Table 5.3 are special change variables that were hypothesised to explain the divergences between actual and stated WTP.

The variables were derived from the factors that may affect demand (from demand theory) and literature on variables affecting willingness to pay. The major one is change in income[18], especially cash availability. Others include; time-to-think, the procedures used to elicit WTP; the reliability of the WTP values reported by the same individual (Whittington et al, 1992), recent household expenditures influencing cash available for net purchase, mosquitoes' nuisance and incidence of malaria, the extent of information the households receive from external sources about nets (cost, characteristics, availability), attitude of community leaders and the urge to copy the neighbours.

There were no expected directional theoretical relationships of the variables with divergent WTP since the test in this study was a hypotheses-generating process. However, it was speculated that giving people time to think about the ITNs and their WTP might work in both ways. For some respondents, it might increase their WTP because they could have thought more about their household income and expenditures

[18] Income was the major cause of divergences between actual and stated WTP (Rowley et al, 1999, Onwujekwe et al,1999; and Bhatia, personal communication).

or could have learnt more about the ITNs, and thus better appreciate its benefits, while for others the opposite mighty happen.

The more the WTP question format stimulated the thought processes of the respondents by somehow resembling the way they bargain for similar goods in the market, the less chances were that there would be a variation between stated and actual WTP. An increase in household income was expected to lead to a positive variation between the two WTP measures.

Similarly, the fact that another person had interviewed the respondent is expected to convey to them the seriousness of the matter, and might lead to an increase in WTP. The more the positive external information a respondent received the more their WTP. Also factors that might increase the actual WTP include the positive attitude of community leaders to the project, positive information from external sources, and copying the neighbours to buy or to not buy the nets.

Table 5.3 Variables that can cause variations in WTP between the first and second survey

Variables	Description	Measurement
Time to think effect	It measures whether the time to think increased WTP	2 = no effect 1 = increased 0 = decreased
Change in household income	It measures whether household income increased or decreased within the two periods as a dummy variable	2 = no effect 1 = increased 0 = decreased
Effect of multiple questioning	Whether the fact that another person also asked for the WTP acted to increase or decrease the amount quoted	2 = no effect 1 = increase 0 = decrease
Information from external sources	It measures whether or not the respondent received further information about ITNs from external sources	2 = no effect 1 = increase 0 = decrease
Attitude of community leaders	Whether or not the positive attitude of community leaders influenced net purchase	2 = no effect 1 = increase 0 = decrease
Copying the neighbours to buy the nets	It measures whether or not the respondent was buying the nets because the neighbours or associates were doing the same.	2 = no effect 1 = increase 0 = decrease

Data analysis for divergences in stated and actual WTP

A variable called divergent WTP was created with 1 representing somebody whose stated WTP in the first survey was at least the price of the net and did not buy (a decrease); 2 was somebody whose stated WTP in the first survey was less than the price of the net and bought (an increase); and 3 was somebody whose stated WTP in the first survey was at least the price of a net and actually purchase a net (no change).

There was a testing of means of the three groups of respondent by socio-economic groups using the Kruskal-Wallis test, to see whether they were statistically significantly different. Chi-square analyses, through cross-tabulations of the binary variables and Friedman test for the continuos socio-economic variables with divergent WTP were used to assess bivariate relationships. Finally, chi-squares through cross-tabulations of the special change variables with divergent WTP evaluated variables that explained divergences in WTP.

Data analysis for divergences in stated WTP at first survey and at retest

A dummy variable called divergent WTP was created with 1 representing somebody whose stated WTP in the first and second surveys differed and 0 if otherwise. A second dummy variable called divergent WTP2 was created to investigate the direction of the divergences and to understand factors that would lead to either an increase or a decrease when there is a divergence. Thus, in divergent WTP2: 1 represented an increase in WTP and 0 a decrease in WTP when there is divergence. The statistical analysis mirrored that for investigating divergences between stated and actual WTP already presented above.

5.7 Summary of the Research Methods

In summary, the study first established the content validity of the question formats and the scenario by using qualitative research methods on sellers and consumers (see Table 5.4 and Table 5.5). This was followed by a survey to determine stated WTP from a sample of 300 respondents per question format. 150 of the same respondents were re-interviewed in a mini-survey within 1 to 5 days of the first survey to assess inter-rater reliability. Econometric modelling was used to determine the construct validity and Pearson's correlation coefficient was the tests of inter-rater reliability.

A second survey was conducted one month after the first survey to determine test-retest reliability and criterion validity, as the survey was accompanied with the actual sale of ITNs. The factors that could cause a divergence between WTP elicited in the first and second surveys were also investigated at this stage (see Table 5.4 and Table 5.5). The sample was all the respondents that were interviewed in the first survey and who were available one month later.

Table 5.4 Summary of the methods for the validity and reliability tests

Test	Modality
Content validity	Assessment of the scenario, the question formats and the general body of the questionnaire by net sellers, consumers and interpreters. The findings were used to modify the Structured Haggling and Binary with follow-up question formats.
Construct validity	Elicitation of stated WTP from 900 randomly selected households from three villages, with each question format used in a sub-sample in each of the villages. Econometric methods were used to test for theoretical constructs of the stated WTP.
Inter-rater reliability	150 randomly selected respondents from each village from the initial sample, were interviewed by another interviewer within 1-5 days of the first interview. Pearson's correlation coefficient was used to determine inter-rater reliability.
Criterion validity	All the respondents in the first survey were the sample. Criterion validity was determined by calculating the phi correlation coefficient and the positive predictive validity. Factors affecting actual WTP, together with those causing divergence between stated and actual WTP were modelled.
Test-retest reliability	The respondents seen during ITNs sales were re-interviewed prior to purchasing nets by their original interviewers. Pearson correlation coefficient was used to determine test retest reliability.

Table 5.5: A matrix of the reliability and validity methods and tests used

Test	Qualitative	Survey 1	Mini-survey	Sale of nets	Survey 2
Content validity	X				
Construct validity		X			
Inter-rater reliability		X	X		
Criterion validity		X		X	X
Test-retest reliability		X			X

CHAPTER 6: CONTENT AND CONSTRUCT VALIDITY

This chapter presents the result of the measures taken to establish and compare the content and construct validity of the various question formats and by inference those of the contingent valuation method in the study area. Thus, an objective of this chapter is to show the extent to which the question formats resembled the price-taking mechanisms of the consumers in the study area. Another objective is to present how the question formats, together with the scenario were improved in order to become more content valid in order to aid the elicitation of more valid WTP responses. The next major objective of the chapter is to see the extent to which the different question formats elicited construct valid WTP responses, by comparing how better content validity translate to better construct validity. The last objective is to determine the appropriateness of different econometric models in determining construct validity of WTP responses.

The first section (6.1) provides the evidence on whether the question formats and the scenario were content-valid. An objective of this section is to show the use of the input from the potential respondents (consumers) and the sellers of mosquito nets to improve the content validity of the structured haggling and binary with follow-up methods (BWFU), together with the scenario. The second objective was to determine the relative content validity of the three question formats in mimicking price-taking in the study area. The processes used in conducting the study that assured that valid and reliable WTP were generated are also presented as part of content validity.

In section 6.2, the results of the construct validity tests of the data generated by the three question formats are presented. The objectives of this section are to compare how different econometric models fared in fitting the data generated by the three question formats and to compare the construct validity of WTP elicited by the three formats. The demographic, malaria-related characteristics of the respondents, together with the descriptive results of stated WTP are first presented to give some perspective for the interpretation of the construct validity results. The construct validity for stated WTP for ITNs re-treatment is presented in section 6.3. The discussion on findings for both content and construct validity is presented in section 6.4.

6.1 Content Validity

The findings from the various steps undertaken to assess and improve the content validity of the three question formats are presented. Similarly, the steps taken to improve the content validity of the scenario are also provided. These included market observation of trading activities on untreated nets geared towards understanding the mechanisms for price-taking in the markets, interviews with sellers of mosquito nets and a focus group discussion (FGD) with consumers. The manner in which these processes led to the development of the version of the structured haggling method used in the study is also presented in the last part of the results of this section.

6.1.1 Market observations

A total of 18 mosquito-net buying transactions were observed in the market. It was found that 13 (72%) of the buyers were males while 5(28%) were females. They transacted for purchase of the nets from 4 different net sellers – all males.

The sellers initiated the haggling in all the cases, and the mode of payment at the conclusion of the bargaining was 100% cash and carry. However, a major bulk buyer that was acquainted with the seller was also given some new nets on credit, so that he could use them to assess demand and pay the seller on his next visit.

The difference between the starting haggle to the final price ranged from 0 to 120 Naira. This meant that while some prices did not change between the starting haggle and final prices, some prices were reduced down by up to 120 Naira, before arriving at the final price.

The average number of haggling iterations was 2.2. It was noticed that the bulk purchasers (information got from the traders) usually bargained not more than 2 times because they were aware of the prevailing prices, whereas the end users or single item purchasers bargained more than 2 times. Females had an average of 3 haggles per transaction, while men had an average of 2 haggles.

In one instance before initiating the haggling, the seller described the qualities of the different nets available in his shop. The description included their various advantages like relative airiness. The buyer then ensured that the net met his requirements and specifications vis a vis: length, height and strength. All these were done before the bargaining started.

In another case, there was a slight aberration to the haggling process. Here, the buyer did not haggle after the first price quoted by the seller. He instead asked the seller for the last price at which the net was sold. The seller then reduced the initial price slightly and then the buyer just bought at the reduced price without further demanding a price decrease.

It was noticed that buyers who were most knowledgeable about price (like regular buyers who were mostly small retailers) usually bought at the first price quoted by the seller. The sellers obviously know these people from constant interactions with them, especially as most of them were bulk purchasers. Thus, they did not waste time haggling but just gave the actual price.

However, in one instance, a price knowledgeable buyer still haggled but ended up buying at the initial price quoted. The sellers remarked that if some buyers are told a high starting haggle price for a net, they might either ask for cheaper versions or compare the prices of similar nets the seller has instead of jumping into haggling.

6.1.2 Interviews with sellers of mosquito nets

There were mixed reactions to the BWFU method as presented as a two-step procedure. One of the sellers remarked that it was used more in department stores and by wholesalers. He said that the buyer should be told that the first price is the actual price and then asked whether s/he was willing to pay that price. If not willing to pay, s/he should be asked how the amount s/he is willing to pay. If this amount is not as much as the intended sale price the seller would now tell the 'real' sale price to the buyer and ask whether willing to pay as a yes or no answer and then stop.

Other respondents, though stating that the BWFU method was not very common in everyday transactions, agreed that the system is used for goods with known prices. However, they felt that the buyers should not be told the exact sale amount at first, since buyers insist on haggling.

All but one of the traders said that the bidding game was not a good method for determining WTP. This stems from one part of the bidding iteration where the price was increased after the buyer had agreed to buy at the first quoted price. However, one respondent said that bidding technique is sometimes used in the market when the seller does not have the good in question but is imagining the prices. Here, if the buyer agrees to the first price, the seller may increase it to see whether the buyer will still be willing to pay, in case he gets the good from another seller. The seller would have explained to the buyer that the good was not available in his store and that he would have to look for it, and was not sure of the exact price.

The respondents all agreed that haggling was the system used at all times, by both retailers and wholesalers. They said that the starting haggle amount could be as high as the seller wanted. On the last stage of the structured haggling iteration where the buyer is asked maximum WTP in case of inflation or uncertainties, the sellers said that it was still applicable if the seller does not have the exact product but intends to get it from another seller. He can then ask the buyer his/her maximum WTP if the price is higher than what they had earlier agreed on. The step is also used if the buyer is not buying immediately but may come back on another day.

On the original step 2 of the structured haggling technique (as described in section 6.1.4), some of them said that the buyer should not be asked to pay-up immediately if his/her first offer is equal or more than the sale price. They stated that the seller should increase the price a little and ask whether they buyer was willing to pay. As they explained, selling at the buyer's first offer may signal to the buyer that either s/he is being cheated or that the net is not good enough.

One of the traders was of the view that the number of haggling iterations should not be limited. He also felt that women should have more haggles. Conversely, others said that there should not be more than 4 haggles, and that there should be equal number of haggles for both men and women.

Concerning the scenario, the traders said that the scenario was understandable and comprehensive enough. However, they said that it should also include the size of the nets, colours, material, durability and ability to retain the insecticide with repeated washing. They were sceptical about including that ITNs may cause burning and redness of eyes as part of cold-like symptoms in some people. They feared that its inclusion would misinform the people, thus resulting in their not buying the nets, because the respondents might misinterpret the symptoms to mean that they literally burn the eyes.

6.1.3 Focus group discussion with consumers

The bidding game was criticised because of the second stage of bidding, in which the respondent is bid up after having agreed to buy at the first price. They preferred the structured haggling technique as the best question format. However, they felt that the BWFU was also good if the nets would be sold in a community with a fixed price. The participants felt that the scenario was adequate, but that we should add how the side effects of ITNs could be treated.

They also suggested a modification in the second stage of the structured haggling technique. This relates to when the buyer's first bid is less than the intended sale price; instead of telling the buyer that the amount was too low, it was suggested that there should be another haggling step whereby the price is lowered. The added step as one of them explained, would enable the buyer to think more and also feel that the exercise is a serious one and that the good in question in beneficial.

6.1.4 Modifying the BWFU format and Structured haggling techniques to improve their content validity

Appendix 11 presents the modified version of the structured haggling technique, as was used for the study. The difference of the structured haggling technique from the original design was that a step was added for customers whose quoted willingness to pay amount in stage 2 is up to the sale price. Here, the seller extends the haggling by coaxing the buyer to pay more. This is now represented by step 3 where the seller (interviewer) reduces the original price offer to see whether the buyer will accept.

Another modification was inclusion of an added step for buyers whose quoted WTP was less than the intended sale price. Here, instead of telling them immediately that their WTP was too low, the seller (interviewer) will decrease the initial price offer and see whether the buyer accepts before going to the next stage of the questioning (new step 4). In addition, these latter group of buyers also have an additional step (step 5) coaxing them to state their true WTP. Appendix 12 presents the modified BWFU method. Here, the 3rd step of the BWFU was added, as a means of coaxing the respondents to state their maximum WTP, as a result of insight gained from interviewing the traders. The bidding game appears as was originally designed (Appendix 8).

6.2 Comparison of the Construct Validity of question formats for WTP for ITNs

This sub-section provides the findings of construct validity from the model specification exercise for WTP for ITNs. The purpose was to decide on the best econometric models for determining construct validity. It also shows how the different econometric models fit the WTP data from the different question formats. The criteria for determining the comparative construct validity of the question formats and the performance of the econometric models were described in Chapter 5.

The first part of this section is used to present the comparative demographic characteristics of the respondents in the three question formats' groups, together with non-parametric tests used to test how similar the groups were. The quantitative results of the perceptions, experiences and expenditures of the respondents and their households to malaria, mosquito nets (including ITNs) and other illnesses are presented. The results of the descriptive analysis of the elicited WTP are presented and compared for similarity among the three groups.

The second part presents the results of construct validity of WTP for ITNs and begins by characterising the respondents' decision on WTP (i.e. positive versus zero WTP values), in order to have an idea of the nature of the zero responses. Thus, the results of the testing of means of the two groups of respondents is presented, followed by econometric modelling using logistic regression of the decision to state either a positive or zero WTP.

Ordinary-least squares (OLS) regression and truncated regression for construct validity of positive WTP responses are then presented. The results of modelling using econometric models suitable for limited dependent variables are presented for the Tobit model and the Heckman selection model estimated using maximum likelihood.

The next section presents the summary of the findings of how all the econometric models performed in each question format group. The final section on construct validity of WTP for the ITNs compares the performance of the WTP data from the question formats, considering how different econometric models fit them.

6.2.1 General descriptive statistics

Demographic information of the respondents

Most of the sample, 268 (89.0%), 275 (91.7%) and 282 (94.0%) responded in the Bidding game (BG), Binary with follow-up (BWFU) and Structured haggling (SH) groups respectively. The rest refused to be interviewed due to lack of interest in the project. However, a few of the people that responded refused to answer some of the questions and their questionnaires were further dropped from the analysis. Thus, the usable numbers of questionnaires for the analysis were 261, 267 and 273 for the BG, BWFU and SH respectively (Table 6.1). The item non-response rates were not determined because of the small number of respondents that presented this problem.

The three groups were relatively similar across a range of characteristics as shown by the p values in Table 6.1. The respondents were mostly heads of households in the three groups; BG (65.5%), BWFU (74.96%) and SH (58.6%). Middle-aged people formed the majority of the respondents. The mean number of household residents was similar across the three groups; 4.1, 5.0 and 4.7 in the BG, BWFU and SH groups respectively. In all three groups, the majority of respondents were females; BG (55.6%), BWFU (51.3%) and SH (54.2%). Most of the respondents were also married. Farmers formed the majority of the respondents in the BG (69.7%) and the BWFU (61.4%) groups, while it was 46% in the SH group.

The three groups of respondents were similar for many of the demographic variables except for status of the respondent, the person's years of formal schooling, food cost and the number of household residents. The people in the SH group had slightly more years of formal education with 5.07 years, followed by the BWFU with 4.02 years and the last was the BG group with 3.52 years.

Also, the value of food consumed was higher in BWFU group than both BG and SH groups. The mean cost of food was lowest in the BG group (943.28 Naira), followed by the SH group (1216.00 Naira) and BWFU group had the highest (1276.98 Naira) Ownership of household assets was homogenous, except for radio.

Malaria Perceptions

More than 80% of the respondents across the three groups correctly identified mosquitoes as the cause of malaria (Table 6.2). However, 5.0%, 6.4% and 2.6% of the respondents in the BG, BWFU and SH groups respectively disagreed that mosquitoes cause malaria ($p>0.05$). The majority of the respondents concurred that nets could prevent mosquito bites (this was for all kinds of nets), but 22.5% of the respondents in the BWFU group did not believe that nets could prevent mosquito bites. These proportions were smaller in the BG (8.8%) and the SH (15%) groups ($p<0.05$).

It was found that 82.0% in the BG group, 58.1% in the BWFU group and 74.7% in the SH group believed that insecticides could prevent mosquito bites ($p<0.05$). The highest proportion was found in the BWFU group (34.5%), while the least was in the BG group (10.7%). About 80% of the respondents in all groups perceived that either they or other members of their household could contract malaria ($p>0.05$).

Table 6.1 Socio-economic and demographic characteristics of the respondents

Variables	Measurement	BG	BWFU	SH	χ^2 (p value)
Status	1 = head 0 = representative Total	171 (65.5) 90 (34.5) 261 (100)	200 (74.9) 67 (25.1) 267 (100)	160 (58.6) 113 (41.4) 273 (100)	16.49 (p=0.000)
No of household residents	Mean (S.D) Median Mode	4.12 (2.14) 4.00 3.00	5.03 (3.10) 5.00 2.0	4.73 (2.66) 4.00 3.00	15.15 (p=.001)
Age (years)	Mean (S.D.) Median Mode	49.52 (17.38) 49.00 45.00	48.74 (14.47) 50.00 50.00	48.04 (14.72) 46.0 40.00	1.98 (p=0.37)
Sex	1 = male 0 = female Total	116 (44.4) 144 (55.6) 261 (100)	130 (48.7) 137 (51.3) 267 (100)	125 (45.8) 148 (54.2) 273 (100)	0.835 (p=0.66)
Years of formal education	Mean (S.D.) Median Mode	3.52 (5.12) 2.00 0	4.02 (4.28) 3.00 0	5.07 (4.78) 5.00 0	19.23 (0.000)
Marital status	1 = married 0 = never married Total	233 (89.3) 28 (10.7) 261 (100)	234 (87.6) 33 (12.4) 267 (100)	238 (87.2) 35 (12.8) 273 (100)	0.53 (p=0.77)
Households previous month's food cost	Mean (S.D.) Median Mode	943.28 (786.21) 730.00 510.00	1276.98 (1008.43) 975.00 600.00	1216.00 (1102.34) 920.00 500	18.78 (p=0.000)
Household with a radio	1 = yes 0 = no Total	186 (71.3) 75 (28.7) 261 (100)	212 (79.4) 55 (20.6) 267 (100)	192 (70.3) 81 (29.7) 273 (100)	7.72 (p=0.021)
Household with a bicycle	1 = yes 0 = no Total	176 (67.4) 85 (32.6) 261 (100)	179 (67.0) 88 (33.0) 267 (100)	167 (61.2) 106 (38.8) 273 (100)	3.85 (p=0.15)
Household with a grinding machine	1 = yes 0 = no Total	34 (13.0) 227 (87.0) 261 (100)	34 (12.7) 233 (87.3) 267 (100)	43 (15.8) 230 (84.2) 273 (100)	1.51 (p=0.47)
Household with a motor cycle	1 = yes 0 = no Total	47 (18.0) 214 (82.0) 261 (100)	63 (23.6) 204 (76.4) 267 (100)	56 (20.5) 217 (79.5) 273 (100)	2.50 (p=0.29)
Household with a motor car	1 = yes 0 = no Total	21 (8.0) 240 (91.6) 261 (100)	23 (8.6) 244 (91.4) 267 (100)	20 (7.3) 253 (92.7) 273 (100)	0.32 (p=0.85)
Occupational group 1: Unemployed/unskilled lab/house wife	1 = yes 0 = no Total	2 (0.8) 259 (99.2) 261 (100)	3 (1.1) 264 (98.9) 267 (100)	3 (1.1) 270 (98.9) 273 (100)	242.46 (p=.000)
Occupational group 2: Farmer	1 = yes 0 = no Total	182 (69.7) 79 (30.3) 261 (100)	164 (61.4) 103 (38.6) 267 (100)	125 (46.0) 148 (54.0) 273 (100)	123.17 (p=.000)
Occupational group 3: Skilled labourer/trading/pensioner	1 = yes 0 = no Total	58 (22.2) 203 (77.8) 261 (100)	68 (25.5) 199 (74.5) 267 (100)	86 (31.6) 187 (68.4) 273 (100)	32.19 (p=0.000)
Occupational group 4: Formally employed	1 = yes 0 = no Total	14 (5.4) 247 (94.6) 261 (100)	20 (7.5) 247 (92.5) 267 (100)	23 (8.5) 250 (91.5) 273 (100)	0.45 (p=0.501)
Occupational group 5: Professionals/mid & big time business	1 = yes 0 = no Total	9 (3.4) 252 (96.6) 261 (100)	11 (4.1) 256 (95.9) 267 (100)	13 (4.8) 260 (95.2) 273 (100)	0.20 (p=0.665)

Table 6.2 The general perceptions of malaria amongst respondents

Variable	BG n (%)	BWFU n (%)	SH n (%)	χ^2 (p value)
Do mosquitoes cause malaria?				
No	13 (5.0)	17 (6.4)	7 (2.6)	2.315
Yes	225 (86.2)	230 (86.1)	243 (89.0)	(0.314)
Don't know	23 (8.8)	20 (7.5)	23 (8.4)	
Total	261 (100)	267 (100)	273 (100)	
Can nets prevent mosquito bites?				
No	23 (8.8)	60 (22.5)	41 (15.0)	11.041
Yes	217 (83.1)	185 (69.3)	209 (76.6)	(0.004)
Don't know	21 (8.0)	22 (8.2)	23 (8.4)	
Total	261 (100)	267 (100)	267 (100)	
Can insecticides prevent mosquito bites?				
No	28 (10.7)	92 (34.5)	47 (17.2)	32.704
Yes	214 (82.0)	155 (58.1)	204 (74.7)	(0.000)
Don't know	19 (7.3)	20 (7.5)	22 (8.1)	
Total	261 (100)	267 (100)	273 (100)	
Incidence perception				
No	41 (15.7)	39 (14.6)	56 (20.5)	3.67
Yes	220 (84.3)	228 (85.4)	217 (79.5)	(0.160)
Total	261 (100)	267 (100)	273 (100)	

Experiences with mosquito nets

Most of the respondents had never purchased any form of untreated mosquito net (Table 6.3). The proportions were uniform in the three groups (BG, 84.3%; BWFU, 84.3%; and SH, 85.7%). Also, very few of the respondents knew the price of untreated nets; BG (5.4%), BWFU (9.4%) and SH (2.6%). However, higher proportions knew where they could buy untreated nets. There was very little prior knowledge about ITNs, as only 1.5%, 2.2% and 1.8% of the respondents from the BG, BWFU and SH groups respectively had ever heard of ITNs before the study. Nobody had ever purchased or re/treated an ITN in the three groups.

Table 6.3 The experiences of the respondents regarding mosquito nets

Variable	BG n (%)	BWFU n (%)	SH n (%)	χ^2 (p value)
Ever bought untreated nets?				
No	220 (84.3)	225 (84.3)	234 (85.7)	0.26
Yes	41 (15.7)	42 (15.7)	39 (14.3)	(p=0.88)
Total	261 (100)	267 (100)	267 (100)	
Know the present cost of nets?				
No	247 (94.6)	242 (90.6)	266 (97.4)	11.58
Yes	14 (5.4)	25 (9.4)	7 (2.6)	(p=0.003)
Total	261 (100)	267 (100)	273 (100)	
Know where to buy an untreated net?				
No	201 (77.0)	143 (53.6)	197 (72.2)	35.626
Yes	60 (23.0)	124 (46.4)	76 (27.8)	(p=0.00)
Total	261 (100)	267 (100)	273 (100)	
Ever heard of ITNs?				
No	257 (98.5)	261 (97.8)	268 (98.2)	0.154
Yes	4 (1.5)	6 (2.2)	5 (1.8)	(p=0.93)
Total	261 (100)	267 (100)	273 (100)	
Ever bought ITNs?				
No	261 (100)	267 (100)	273 (100)	p<0.05
Yes	0 (0)	0 (0)	0 (0)	
Total	261 (100)	267 (100)	273 (100)	
Ever treated a net?				
No	261 (100)	267 (100)	273 (100)	p<0.05
Yes	0 (0)	0 (0)	0 (0)	
Total	261 (100)	267 (100)	273 (100)	

Actual incidences and expenditures on malaria and other illnesses

47.5%, 50.9% and 46.7% of all the respondents' households in the BG, BWFU and SH group respectively had case(s) of malaria attack in the month preceding the interview, with an average of approximately one person per household falling prey to malaria within that period (Table 6.4). This number of malaria attacks was similar across the three groups (chi-square = 2.43; p=0.30).

However, not all the malaria attacks resulted in expenditures in the households. The average expenditures on malaria treatment per household were 327.38 Naira in the BG group, 431.28 Naira in the BWFU group and 337.21 Naira in the SH group. This level of expenditures was also similar across the three groups (chi-square = 1.56; p=0.46).

Small numbers of the households spent money to prevent malaria in the three groups with the highest proportion in the BWFU group (33.0%) and the least in the SH group (18.7%). The average expenditures on malaria prevention per month was also highest in the BWFU group at 89.80 Naira followed by the SH group 41.01 Naira, while the BG group was the lowest with 36.17 Naira. This variable was heterogeneous across the three groups (chi-square = 14.73; p=0.001).

The proportion of households that fell prey to other illnesses were almost equal to those that had malaria in the BG group (51.0%), but were less in the BWFU (46.3%) and the SH group (33.5%). This variable was significantly different across the three groups (chi-square = 10.69; p=0.005). Post-Hoc analysis with Mann-Whitney U test showed that all the pairs of question formats' groups were statistically different from each other ($p<0.05$). The number of ill people per household was less for other illnesses than in malaria.

As in case of malaria, not all the household incurred expenses when illness arose. Nevertheless, the expenditure on other illnesses was more in the bidding game group at 577.64 Naira, almost twice as high as the malaria expenditure at 806.20 Naira in the BWFU group, but slightly less in the structured haggling group at 309.88 Naira when compared to malaria expenditure. The differences in the expenditures on other illnesses across the three groups was statistically significant (chi-square = 24.03; p=.000). Post-hoc analysis with the Mann-Whitney U test showed that while the bidding and BWFU

groups were similar (p>0.05), they were both different from the haggling group (p<0.05).

Table 6.4 Actual experiences and expenditures on malaria and other illnesses

Variables	Definition and measurement	BG n (%)	BWFU n (%)	SH n (%)	χ^2 (p value)
Households with malaria	Yes	124 (47.5)	136 (50.9)	128 (46.7)	
	No	137 (52.5)	131 (49.1)	145 (53.3)	0.93
	Total	261 (100)	267 (100)	272 (100)	(p=0.63)
Malaria incidence	Mean	0.67	0.76	0.68	2.43
	(S.D.)	(0.78)	(0.88)	(0.92)	(p=0.30)
	Median	0	1.0	0	
Households with expenditure on malaria treatment	Yes	118 (45.2)	133 (49.8)	120 (44.0)	
	No	143 (54.8)	134 (50.2)	153 (56.0)	1.75
	Total	261 (100)	267 (100)	273 (100)	(p=0.42)
Expenditures on malaria treatment	Mean	327.38	431.28	337.21	
	(S.D.)	(782.77)	(1011.85)	(727.68)	1.56
	Median	0	0	0	(p=0.46)
Households with expenditure on malaria prevention	Yes	63 (24.1)	88 (33.0)	51 (18.7)	
	No	198 (75.9)	179 (67.0)	222 (81.3)	13.95
	Total	261 (100)	267 (100)	273 (100)	(p=0.001)
Expenditures on malaria prevention	Mean	36.17	89.79	41.01	
	(S.D.)	(143.58)	(230.85)	(130.42)	14.73
	Median	0	0	0	(p=0.001)
Households with other illnesses	Yes	133 (51.0)	124 (46.3)	92 (33.5)	17.99
	No	128 (49.0)	143 (53.6)	181 (66.5)	(p=0.00)
	Total	261 (100)	267 (100)	273 (100)	
Actual incidence of other illnesses	Mean	0.63	0.63	0.50	
	(S.D.)	(0.71)	(.88)	(0.81)	10.69
	Median	1.00	0.00	0.00	(p=.005)
Households that spent money to treat other illnesses	Yes	129 (49.4)	118 (44.2)	79 (29.0)	24.97
	No	132 (50.6)	149 (55.8)	194 (71.0)	(p=0.00)
	Total	261 (100)	267 (100)	273 (100)	
Expenditures on treating other illnesses	Mean	577.64	806.20	309.88	24.030
	(S.D)	(2994.56)	(2720.17)	(1058.42)	(p=0.000)
	Median	0.00	0.00	0.00	

Description of stated willingness to pay

More than 75% of the respondents in all three groups stated positive willingness for the ITNs, with the highest proportion of 80.5% coming from the BWFU group (Table 6.5). The decision on WTP was not statistically different across the question formats' groups (chi-square = 2.93; p=0.231). The major reason that some people gave for not being willing to pay anything for the nets was lack of money (Table 6.6).

The unconditional (all respondents) mean WTP was highest in the BWFU (195.07 Naira; 95%CI = 178.7 - 211.5), followed by SH group (190.22 Naira; 95%CI = 170.87 - 209.9), while the least was the BG group (162.04 Naira; 95%CI = 145.2 - 178.9). Kruskal-Wallis test showed that the distributions of level of WTP across the three groups were statistically different (chi-square = 10.04; p=0.007).

The median unconditional WTP was the same for the BWFU and SH groups at 200 Naira, while the BG group had a lower median of 150 Naira. Post-hoc tests with Mann-Whitney U test showed that there was statistically significant difference only between the medians from the BG and BWFU groups ($p<0.05$).

The SH group had the highest conditional mean (all positive WTP amounts) WTP at 248.99 Naira, followed by BWFU group (242.26 Naira), and with the BG group bringing up the rear once more (212.78 Naira). However, the conditional median was the same for all three groups at 200 Naira respectively.

The mean WTP for re-treatment was higher in the SH (49.23 Naira; 95%CI = 45.30 - 53.15) and BWFU (43.43 Naira; 95% CI = 40.31 - 46.54), than in the BG group (37.06 Naira; 95% CI = 37.50 - 40.63). Kruskal-Wallis test showed that the distributions of level of WTP across the three groups were statistically different (chi-square = 10.04; p=0.007).

The median WTP for re-treatment followed the same pattern and was uniform in the BWFU and SH groups at 50 Naira, but was 30 Naira in the bidding group. Post-hoc tests with Mann-Whitney U test showed that there was statistically significant difference in only the medians from the BG and BWFU groups ($p<0.05$) and the BG and SH group ($p<0.05$).

The three measures of central tendency for WTP for ITNs, were all about 200 Naira for the BWFU group suggesting that the data was normally distributed. Conversely, the BG and SH data were not normally distributed. However, comparing the three measures of central tendency for WTP for net re-treatment, which were all about 50 Naira for the BWFU and SH groups showing that they were normally distributed. Conversely, the BG data was not normally distributed.

Table 6.5 Stated WTP for nets and re-treatment

Variable	Measurement	BG	BWFU	SH	χ^2 (p value)
WTPYN: Whether the respondent was willing to pay for ITNs	1 = Yes 0 = No Total	196 (75.1) 65 (24.9) 261 (100)	215 (80.5) 52 (19.5) 267 (100)	206 (75.5) 67 (24.5) 273 (100)	2.93 (p=0.23)
WTP for ITNs: Continuous monetary measure of respondents WTP for their own ITNs	Mean (S.D.) [95% CI] Median [25th - 75th percentiles] Mode	162.0 (137.8) [145.2-178.9) 150.0 [50.0-250.0] 0	195.1 (136.2) [178.7-211.5] 200.0 [100.0-300.0] 200.0	190.4 (162.7) [170.9-209.9] 200.0 [40.0-290.0] 0	10.04 (p=0.01)
WTP for ITNs re-treatment: Continuous monetary measure of respondents WTP for net re-treatment	Mean (S.D.) [95% CI] Median [25th - 75th percentiles] Mode	37.06 (25.64) [37.50-40.63] 30.00 [20.00-50.00] 20.00	43.43 (23.06) [40.31-46.54] 50.00 [20.0 - 50.0] 50.00	49.20 (29.0) [45.30-53.15] 50.0 [25.0 - 50.0] 50.0	15.46 (p=0.00)
Natural log transformed WTP for nets	Mean (S.D.) Median Mode	5.21(0.59) 5.30 5.30	5.38(0.51) 5.30 5.30	5.36(0.60) 5.30 5.30	
Natural log transformed WTP for ITNs re-treatment	Mean (S.D.) Median	3.45(0.56) 3.40	3.63(0.55) 3.91	3.75(0.63) 3.91	

Post-hoc tests with Mann-Whitney U (MWU)

WTP for nets:

BG vs BWFU [MWU = 29242.5 (p=0.001)]

BG vs SH [MWU = 31087.5 (p=0.081)]

SH vs BWFU [MWU = *32703.0 (p=0.217)]*

WTP for re-treatment:

BG vs BWFU [MWU = 29436.5 (p=0.002)]

BG vs SH [MWU = 28112.5 (p=0.000)]

SH vs BWFU [MWU = 33264.0 (p=0.354)]

Table 6.6 The reasons respondents were not willing to pay

Variable	BG n (%)	BWFU n (%)	SH n (%)
No money			
Yes	59 (90.8)	44 (88.0)	58 (86.6)
No	6 (9.2)	8 (12.0)	9 (13.4)
Total	65 (100)	52 (100)	67 (100)
No interest			
Yes	7 (10.8)	4 (7.7)	6 (9.0)
No	58 (89.2)	48 (92.3)	61 (91.0)
Total	65 (100)	52 (100)	67 (100)
No bed			
Yes	1 (1.5)	0 (0)	1 (1.5)
No	64 (98.5)	52 (100)	66 (98.5)
Total	65 (100)	52 (100)	67 (100)
Others			
Yes	1 (1.5)	2 (3.9)	3 (4.5)
No	64 (98.5)	50 (96.1)	64 (95.5)
Total	65 (100)	52 (100)	67 (100)

Further analysis of the positive WTP responses showed that very small proportions were willing to pay as much as the price of the nets (450 Naira per net) that had been fixed by the National Malaria Control Programme. A slightly higher proportion was willing to pay 350 Naira that the nets would be offered for sale in the project (Table 6.7). The lowest proportion was in the BG group where only 19(7.28%) of the respondents stated amounts that were enough to cover the price of the net. The proportion was higher with the SH and BWFU groups with 36(13.24%) and 36 (13.48%) respectively.

Table 6.7 Number willing to pay the market price of ITNs

Variable	Definition and measurement	BG(%)	BWFU(%)	SH(%)
WTP≥350	Those willing to pay equal to, or more than the price of the net (350Naira)	19 (7.28%)	36 (13.48%)	36 (13.24%)
WTP<350	Those not willing to pay up to the price of the net	242(92.72%)	231(86.52%)	237(86.76%)
Total		261 (100)	267 (100)	273 (100%)

6.2.2. Nature of the zero WTP responses

In the first part of this section, the results of the bivariate analysis of the relationship between positive WTP and the independent variables are presented. In the second part, the results of the logistic regression of WTP on independent variables are presented. The marginal effects of the variables are used to interpret their effects on WTP, since direct interpretations of the logistic coefficients, which are log odds, are less informative.

Bivariate analysis

In general, the respondents who were unwilling to pay were more likely to belong to the lower socio-economic groups since they spent less money on food and health care (Table 6.8). The people that were willing to pay had higher incidence of malaria and expenditures to both prevent and treat it and had previously purchased more untreated nets. This latter group of respondents also had more household assets. Food cost differed significantly between the respondents that were willing to pay and others in the three elicitation groups ($p<0.01$), showing that income was the major determinant on the decision to buy a net. However chi-square tests showed that not all the differences between the groups of respondents with positive WTP and those with zero WTP were statistically significant.

A key question is whether those that stated zero WTP would have stopped at zero, or whether some would have had negative WTP or whether they choose to state an unwillingness to pay, because they felt the amount they could afford to pay may have been ridiculed as too low. However, most of these people, in addition to spending less on malaria treatment and prevention, were of lower socio-economic group, it could be safe to assume that they could not afford to pay for the ITNs.

As shown in Table 6.8, food cost was consistently different between the WTPbin=0 (not willing to pay) and WTPbin=1 (willing to pay a positive WTP) respondents in the three question formats' groups ($p<0.01$), indicating that income was the major determinant of the decision to buy a net. The other variables that were significantly different between the WTPbin=0 and WTPbin=1 respondents in all three question formats were the number of household residents and expenditure on malaria treatment.

The chi-square tests in the BWFU and haggling group confirmed that household heads, people with more numbers of years of formal schooling and males were more likely to state a positive WTP. The other statistically significant variables that explain the differences in the decision to state a positive WTP are presented in Table 6.8 and confirm the assertions already given above that those not willing to pay positive amounts could not afford to pay.

Food cost and asset variables show that money really matters, together with previous malaria and mosquito experiences, the ability of the respondent to make decisions on behalf of the household together with ownership of assets for people to state a positive WTP.

Table 6.8 Bivariate analysis of the groups willing (WTPbin=1) and unwilling (WTPbin=0) to pay positive WTP amounts

	Bidding		BWFU		Haggling	
Variables	WTPbin=0 N=63 Mean (SD)	WTPbin=1 N= 198 Mean (SD)	WTPbin=0 N=63 Mean (SD)	WTPbin=1 N= 198 Mean (SD)	WTPbin=0 N=65 Mean (SD)	WTPbin=1 N= 207 Mean (SD)
Status in the household	.62 (.49)	.67 (.47)	.63 (.49)**	.78 (.42)	.47 (.50)**	.62 (.49)
Number of household residents	3.68 (2.00)*	4.26 (2.17)	3.88 (1.98)***	5.17 (2.54)	3.56 (1.80)***	5.00 (2.60)
Sex	.44 (.53)	.45(.50)	.29 (.46)***	.53 (.50)	.28 (.45)***	.51(.50)
Age	49.54 (18.72)	49.52 (16.98)	49.06 (14.31)	48.97 (14.25)	52.00 (14.05)**	46.80(14.73)
Years of schooling	3.06 (4.66)	3.66 (5.27)	2.31 (3.50)***	4.32 (4.36)	2.57 (3.71)***	5.86 (4.81)
Marital status	.89 (.32)	.89(.31)	.88 (.32)	.87 (.33)	.92 (.27)	.86(.35)
Occupational group 1: Unemployed/unskilled lab/house wife	.00 (.00)	.001 (.10)	.04 (.19)**	.005 (.07)	.00 (.00)	.002 (14)
Occupational group 2: Farmer	.68 (.47)	.70 (.46)	.75 (.44)**	.58 (.49)	.70 (.46)**	.51 (.50)
Occupational group 3: Skilled labourer/trading/pensioner	.19 (.40)	.23 (.42)	.10 (.30)***	.29 (.46)	.25 (.44)	.35 (.48)
Occupational group 4: Formally employed	.11 (.32)**	.004 (.19)	.06 (.24)	.08 (.27)	.005 (.21)	.11 (.31)
Occupational group 5: Professionals/mid & big time business	.002 (.13)	.004 (.20)	.08 (.27)	.03 (.18)	.003 (.18)	.005 (.22)
Perceived incidence of malaria	.17 (.63)	.15 (.36)	.13 (.40)	.16 (.39)	.14 (.35)	.23 (.42)
Actual incidence of malaria	.52 (.72)	.71 (.80)	.35 (.52)***	.87 (.91)	.42 (.59)**	.77 (.99)
Exp. To treat malaria (maltreat)	201.43 (489.93)	367.43 (852.33)	220.58 (699.86)***	482.24 (1068.88)	120.39 (304.00)***	405.87 (805.26)
Dummy of maltreat	.37 (.49)*	.48 (.50)	.33 (.47)***	.54 (.50)	.33 (.47)**	.48 (.50)
Exp. to prevent malaria (malprev)	19.05 (54.85)**	41.62 (161.67)	41.92 (104.23)	101.37 (250.92)	13.90 (47.93)*	49.54 (146.09)
Dummy of malprev	.14(.35)**	.27(.45)	.25 (.44)	.35 (.48)	.11(.31)*	.21(.41)
Actual incidence of other illnesses	.59 (.64)	.64(.73)	.35 (.52)***	.69 (.94)	.30 (.63)	.56(.85)
Exp. To treat other illnesses (illtreat)	350.48 (908.20)	584.02 (3400.34)	450.58 (1343.77)**	892.62 (2955.51)	184.53 (766.76)**	348.44 (1131.95)
Dummy of illtreat	.46 (.50)	.51 (.50)	.29 (.46)**	.48 (.50)	.22 (.42)	.31 (.46)
Food cost	680.90 (539.39)***	1026.76 (833.79)	809.77 (722.08)***	1389.98 (1036.3)	629.77 (733.67)***	1395.98 (1137.60)
Previous purchase of net	.01 (.30)	.18 (.38)	.12 (.32)	.17 (.37)	.003 (.17)***	.18 (.38)
Ownership of radio	.62 (.49)*	.74 (.44)	.63 (.49)***	.83 (.37)	.46 (.50)***	.78 (.41)
Ownership of bicycle	.65 (.48)	.68(.47)	.62 (.49)	.68 (.47)	.45 (.50)***	.66(.47)
Ownership of grinding machine	.008 (.27)	.15(.35)	.08 (.27)	.14 (.35)	.005 (.21)***	.19(.40)
Ownership of motorcycle	.14 (.35)	.19 (.39)	.15 (.36)	.26 (.44)	.19 (.39)	.21 (.41)
Ownership of car	.13 (.34)	.006(.25)	.04 (.19)	.10 (.30)	.002 (.13)**	.009(.29)

Significance of differences in variables * <.10, **<.05, ***<.01 (inserted beside the WTPbin=0 groups).

Logistic regression

Table 6.9 shows that the log of food cost was the only variable that was uniformly significant across all question formats. The results showed that as food cost increased, the probability of respondents stating positive WTP amounts increased and the marginal effects of food cost were 0.18, 0.06 and 0.10 in the bidding game, BWFU and structured haggling groups respectively. Few variables were, however, statistically significant and not all the statistically significant variables had the hypothesised signs.

Two of the variables in the reduced bidding game model (dummy of expenditure to prevent malaria with a marginal effect of 0.14 and the log of food cost) had the expected signs out of the statistically significant variables. In the BWFU group, three variables namely, number of years of schooling, actual incidence of malaria (with marginal effects of 0.01 and 0.20 respectively) and log of food cost had the hypothesised signs. In the haggling group, variables like number of years of schooling, ownership of grinding machine (with marginal effects of 0.10 each) and log of food cost also had the hypothesised signs out of the statistically significant variables.

The variables that disobeyed theoretical expectations in the bidding group were ownership of car ($p<0.01$) and occupational class 4 ($p<0.05$). Ownership of a car and belonging to occupational class 4 will decrease the probability of stating a positive WTP for ITNs by 0.42 and 0.47 (marginal effects) respectively. The dummy of the expenditure to treat malaria had a negative sign instead of the hypothesised positive sign in the BWFU group, while actual incidence of other illnesses and ownership of a motorcar had the wrong signs in the haggling group. In this case, using the marginal effects for interpretation, a unit increase in actual incidence of other illnesses would lead to an increased probability of having a positive WTP by 0.11. Conversely, a unit increase in ownership of a motorcar would decrease the probability of having a positive WTP by 0.09. The logistic regressions nevertheless confirmed the findings of the bivariate analysis that higher socio-economic groups and those with more exposure to malaria were more willing to pay for the ITNs. The reduced models for the three question formats were all statistically significant ($p<0.01$) and were able to correctly predict more than 80% of the observations. The likelihood ratio test showed that the full and reduced models were similar in the three question formats.

Table 6.9 Logistic regression to characterise the nature of zero WTP values

Dependent variable: WTPbin; 1 if WTP >0 and 0 if WTP = 0

	Bidding		BWFU		Haggling	
Variables	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff. (SE)	Reduced Coeff. (SE)
Status in the household	.10 (.43)	.13 (.42)	.38 (.51)	.51 (.50)	.20(.50)	.29(.46)
No of household residents	.07 (.10)	.07 (.10)	.12 (.11)	.10 (.10)	.003 (.08)	-.001(.08)
Sex	-.40 (.45)	-.43 (.43)	.42 (.49)	.35 (.46)	.54 (.54)	.43(.50)
Age	.01 (.02)	.01 (.01)	.02 (.02)	.016 (.02)	-.01 (.02)	
School years	-.01 (.05)		.17 (.07)**	.16 (.07)**	.14(.06)**	.16 (.06)***
Marital status	-.32 (.66)	-.29 (.58)	-.32 (.58)		-.51 (.66)	-.53 (.65)
Perceived incidence of malaria	-.54 (.48)	-.56 (.46)	-.73 (.59)	-.75 (.56)	.14 (.53)	
Actual incidence of malaria	.02 (.43)		3.09 (1.13)***	3.10 (1.12)***	.09 (.43)	.20 (.28)
Dummy of exp. treat malaria	.35 (.66)	.42 (.36)	-2.73 (1.30)**	-2.76 (1.29)**	-.51 (.66)	-.55(.61)
Dummy of exp. to prevent malaria	1.12 (.48)**	1.15 (.47)**	.26 (.46)		.05 (.67)	
Actual incidence of other illnesses	-.13 (.49)		.58 (.65)	.88 (.35)**	1.22 (.61)**	1.16(.60)*
Dummy of exp. to treat other illnesses	.24 (.69)		.42 (.80)		-.93 (.85)	-.92 (.81)
Previous purchase of nets	.73 (.55)	.74 (.55)	-.26 (.66)		1.39(92)	1.27 (.89)
Ownership of radio	.18(.43)		.41 (.47)	.37 (.44)	.49(.44)	.49(.43)
Ownership of bicycle	-.09 (.42)		-.14 (.45)		.24 (.42)	.30(.41)
Ownership of grinding machine	.71 (.64)	.67 (.61)	-.22 (.70)		1.60(.90)*	1.56(.89)*
Ownership of motorcycle	.63 (.56)	.67(.55)	-.63 (.57)	-.62 (.54)	-1.65 (.61)***	-1.62 (.59)***
Ownership of car	-1.98 (.69)***	-1.98 (.004)***	1.22 (.91)	1.30 (.90)	1.60(1.18)	1.53(1.16)
Log of Food cost	1.20 (.31)***	1.20 (.30)***	.96 (.30)***	.96 (.30)***	.99 (.27)***	.99 (.26)***
Occupational group 2: Farmer	.94 (1.15)	.98 (1.11)	.07 (1.07)	.003 (1.05)	-.18 (.94)	
Occupational group 3: Skilled labourer/trading/pensioner	.79(1.10)	.83 (1.08)	.47 (1.07)	.36 (1.06)	.091 (.939)	
Occupational group 4: Formally employed	-2.18 (1.12)*	-2.17 (1.10)**	-.74 (1.17)	-.83 (1.15)	.83 (1.21)	
Occupational group 5: Professionals/mid & big time business	.69(1.68)	.76(1.65)	-1.66 (1.30)	-1.76 (1.27)	.32 (1.35)	
Constant	-8.06 (2.39)***	-.8.11 (2.33)***	-8.11 (2.40)***	-8.07 (2.30)***	-5.96 (2.29)***	-6.38 (1.73)***
No of observations	261	261	267	267	273	273
Chi2	53.29***	52.98***	79.03***	78.04***	94.27***	91.87***
Pseudo R2	0.1884	0.1874	0.3007	0.2964	0.3258	0.3175
Log likelihood	-114.736	-114.891	-91.911	-92.621	-97.559	-98.760
Correct predictions	80.69%	80.31%	84.59%	84.64%	85.77%	86.14%
LR test chi2 (p value)	0.31 (0.99)		1.07 (0.96)		2.54 (0.86)	

Significance of parameters * <.10, **<.05, ***<.01

Note: The LR test reported is for the full versus reduced models and Probit models were also used to estimate the regressions and they gave the same results as the logistic models.

6.2.3 Comparing the estimators within the different question formats

The results of modelling the whole continuous distribution of WTP are now presented. The models were compared within each question format in order to determine how well they explained WTP. Tables 6.10 to 6.12 show the regression diagnostics that formed the basis for selecting particular models as best performing within each question format for the assessment of construct validity. The summaries of the comparisons are:

Tobit and Heckman's models

In the bidding game group, the diagnostics showed that the residuals from Heckman model were normally distributed. The RESET test failed to reject the null hypothesis of correct specification, while the Tobit was the wrong functional form with a t-statistic of 3.28 for the RESET test ($p<0.05$) (Table 6.10). A plot of the residuals versus the fitted values of WTP showed no evidence of heteroscedascticity in either model.

Table 6.10 Comparison of bidding game models

	Predicted values	Normality (χ^2 critical value at 0.95 = 5.991)	RESET ($t_{0.05} \cong 2$)
OLS LnWTP	5.18	5.70	0.325
Truncated regression	5.18	11.25*	0.312
Tobit	3.65	2.44	3.278*
Heckman	5.38	0.58	0.19

*$p<0.05$; Mean of raw Log WTP = 5.21 and Log +1 WTP = 3.97

In the BWFU group's data, the residuals from Heckman's model were normally distributed and the RESET showed that it was correctly specified (Table 6.11). Plots of the residuals versus the fitted values of WTP showed no evidence of heteroscedascticity in both models. The Heckman also predicted raw WTP closely. The residuals from the Tobit model were normally distributed and had the right functional form. The plot of the residuals versus the fitted values revealed no evidence of heteroscedascticity in the Tobit model. However, the Heckman's model had better Ramsey RESET values and predicted raw WTP better than the Tobit model. Therefore, for the BWFU method, the Heckman's model is the preferred model for construct validity tests.

Table 6.11 Comparison of the BWFU models

	Predicted values	Normality (χ^2 critical value at 0.95 = 5.991)	RESET ($t_{0.05} \cong 2$)
OLS	195.07	0.92	0.26
Truncated regression	232.14	6.51*	0.86
Tobit	178.77	4.85	1.89
Heckman	201.89	0.42	0.42

*$p<0.05$; Mean of raw WTP = 195.07

The regression diagnostics of the models from the data from the structured haggling group showed that the residuals were not normally distributed (borderline significance) and the RESET test was also slightly more than the critical value except in the Tobit model (Table 6.12). Plots of the residuals versus the fitted values of WTP showed no evidence of heteroscedascticity in either model. However, the Heckman predicted raw WTP very well. The Tobit was the wrong functional form with a t-statistic of 4.84 ($p<0.05$). Thus, as with the previous question formats, the Heckman model better than the Tobit for the structured haggling data.

Table 6.12 Comparison of the structured haggling models

	Predicted values	Normality (χ^2 critical value at 0.95 = 5.991)	RESET ($t_{0.05} \cong 2$)
OLS LnWTP	5.29	6.12*	1.33
Truncated regression	5.29	6.12*	1.31
Tobit	3.79	1.11	4.84*
Heckman	5.34	7.09*	2.16*

*$p<0.05$; Mean of raw log WTP = 5.36 and log + 1 WTP = 4.10

OLS and Truncated regression for all question formats

The OLS had better outcomes than the truncated regression in both the bidding group and BWFU data (Tables 6.10 to 6.12). However, they were indistinguishable in the structured haggling group. Thus, the OLS was selected as the better estimator of positive WTP.

Diagnostics for multicollinearity

There was no general evidence of multicollinearity in the data generated by any of the question format. The major correlation was with between the second and third occupational groups in the Bidding game (0.89), BWFU (0.88) and structured haggling (0.88). The correlation between other pairs of variables was less than 0.70.

6.2.4 Comparing the OLS estimator across the different question formats

The number of statistically significant variables that obeyed set hypotheses, the adjusted R-squared, and the regression diagnostics were used to compare the models for the different question formats. The OLS regression best fits the data generated using the BWFU followed closely by the structured haggling.

Most of the statistically significant variables had the hypothesised signs (Table 6.13). The log of food cost was statistically significant across the three question formats and also had the expected positive sign. The incidence of malaria was uniform in the BWFU and structured haggling groups and had the hypothesised positive sign. Also, there were six statistically significant variables in the bidding game that had the hypothesised signs, four in the BWFU group and three in the haggling group.

The coefficients showed that as food cost increased, the respondents were more likely to state a larger positive WTP. The economic effects of number of years of schooling, incidence of malaria and ownership of radio were considerable in the BWFU. The result showed that a unit increase, in the value of any of these variables would lead to an increase in the level of WTP from about 9 Naira (years of schooling) to 43 Naira (ownership of radio).

A variable that contradicted theoretical expectations in the bidding group was ownership of a car ($p<0.01$). Instead of being positively related to WTP, it had a negative sign. A unit increase in ownership of a car will decrease the level of log WTP for an ITN by 0.29 (all other variables being held constant in each case). In the BWFU, belonging to occupations 4 (formally employed people) and 5 (professionals and mid/big time business people) had massive effects as they would each decrease the level of WTP by 103.4 Naira and 126.6 Naira per any unit increase respectively.

The OLS explained about 17% of the variation in WTP the Bidding group, 18% in the Haggling groups and 20% in the BWFU group. The reduced models were all statistically significant ($p<0.01$). The models in the bidding and BWFU groups were both homoscedastic and normally distributed, while the haggling was heteroscedastic and not normally distributed and thus, robust standard errors were estimated for the interpretation of the haggling model. However, the Ramsey RESET test failed to find evidence of misspecification in any of the models.

Table 6.13 OLS regression of positive WTP

	BG		BWFU		SH (with robust SE)	
Dependent variable	Log WTP		WTP		Log WTP	
Variables	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff. (SE)	Reduced Coeff. (SE)	Full Coeff.(SE)	Reduced Coeff. (SE)
Status in the household	-.09 (.10)		9.66 (23.19)		.22 (.12)*	.28 (.09)***
Number of household residents	-.02 (.02)	-.01 (.02)	3.77 (3.67)	3.51 (3.36)	.02 (.02)	.02 (.02)
Sex	.18 (.11)*	.17 (.08)**	15.60 (20.32)	19.28 (16.25)	.06 (.12)	
Age	-.0002 (.003)		-.03 (.66)		.002 (.003)	
School years	-.01 (.01)		9.17 (2.40)***	9.09 (2.14)***	.012 (.01)	.01 (.01)
Marital status	.25 (.14)*	.31 (.13)**	-21.12 (25.14)		-.22 (.13)*	-.23(.08)***
Perceived incidence of malaria	-.03 (.12)		-10.37 (21.45)		.03 (.10)	
Actual incidence of malaria	.08 (.11)		39.64 (18.98)**	17.43 (9.15)*	.14 (.07)**	.14 (.06)**
Dummy of exp. treat malaria	-.09 (.17)		-40.61 (32.42)		-.18 (.14)	-.14 (.13)
Dummy of exp. to prevent malaria	.30 (.10)***	.32 (.09)***	-14.70 (17.46)		.07(.12)	
Actual incidence of other illnesses	.002 (.115)		-13.60 (14.23)		.01 (.08)	
Dummy of exp. to treat other illnesses	-.23 (.16)	-.24 (.08)***	28.85 (4.63)		.02 (.14)	
Previous purchase of nets	.03 (.11)		19.04 (22.74)		-.04 (.11)	
Ownership of radio	.08 (.10)		37.34 (21.19)*	42.72 (20.01)**	.15 (.11)	.17 (.11)
Ownership of bicycle	.02 (.10)		9.88 (18.28)		-.08(.09)	-.09 (.09)
Ownership of grinding machine	.27 (.13)*	.29 (.11)**	-44.28 (25.62)*	-41.09 (23.69)*	.10 (.12)	
Ownership of motorcycle	.14 (.11)	.16 (.10)	-30.16 (19.97)	-27.04 (18.97)	-.04(.11)	
Ownership of car	-.32 (.18)*	-.28(.16)*	31.83 (29.99)		.09(.15)	.10 (.12)
Log of Food cost	.08(.08)	.13(.07)*	29.33 (11.65)**	29.77 (11.20)***	.10 (.07)	.13 (.05)**
Occupational group 2: Farmer	.08 (.24)		-34.37 (43.52)	-31.22 (41.86)	-.16 (.17)	
Occupational group 3: Skilled labourer/trading/pensioner	.20 (.23)		-11.22 (42.84)	-10.18 (41.59)	-.13 (.18)	
Occupational group 4: Formally employed	.28 (.26)		-103.09 (49.72)**	-98.35 (47.56)**	-.13 (.20)	
Occupational group 5: Professionals/mid & big time business	.29 (.33)		-126.61 (57.22)**	-135.90 (55.67)**	.18 (.25)	
Constant	4.35 (.57)***	4.04 (.45)***	-55.79 (88.42)	-72.81 (80.69)	4.41 (.49)***	4.28 (.39)***
Number of Obs.	198	198	266	267	204	204
F stat	2.31***	5.586***	3.53***	6.41***	2.57***	5.51***
Adj R2	0.13	0.17	0.18	0.20	0.15	0.18
Standard error of estimate	.5445	.5315	123.58	122.12	.5542	.5551
Cook-Weisberg		1.50		1.30		13.39***
Normality		5.70		0.92		6.12*
Predicted values		5.18:ref 5.21		195.07:ref 195.07		5.29:ref 5.36
RAMSEY RESET		0.36		0.26		1.33

Significance of parameters * <.10, **<.05, ***<.01

Note: Robust standard errors were estimated for the interpretation of the haggling model.

6.2.5 Comparing the Heckman's estimator across the different question formats

In this section, the result of modelling the full WTP distribution is presented. Apart from the number of statistically significant variables that obeyed set hypothesis, other criteria comparing the performance of the three question formats include the goodness of fit measure which is the magnitude of the chi-square test and the magnitude of the chi-square of independent equations. The Heckman selection regression best fits the data generated using the Bidding game followed closely by the BWFU. The model did not fit the data generated by the structured haggling technique because sample selection was not a problem there.

None of the variables was statistically significant across all question formats (Table 6.14). However, food cost was statistically significant in the BWFU and structured haggling groups and with the same sign. Other variables that were statistically significant in pairs of question formats had different signs. For instance, marital status had the hypothesised positive sign in the bidding group but the opposite sign in the structured haggling group. Also, ownership of grinding machine and motorcycle had the hypothesised positive sign in the bidding group but the opposite signs in the BWFU group.

All six statistically significant variables had the hypothesised signs in the Bidding group (Table 6.14). These variables were sex ($p<0.05$), marital status ($p<0.05$), dummy of the expenditure to prevent malaria ($p<0.05$), dummy of the expenditure to treat other illnesses ($p<0.01$), ownership of grinding machine ($p<0.05$) and ownership of a motorcycle ($p<0.10$).

In the BWFU group, just four out of the eight statistically significant variables had the expected signs. These were the numbers of years of schooling ($p<0.01$), previous purchase of nets ($p<0.10$), ownership of radio ($p<0.05$) and food cost ($p<0.05$). Conversely, ownership of grinding machine ($p<0.05$), ownership of motorcycle ($p<0.10$), and occupations 4 ($p<0.01$) and 5 ($p<0.05$) had unexpected signs.

Four of the five statistically significant variables in the structured haggling group had the hypothesised signs. These were the status of the respondent in the household ($p<0.01$), the number of household residents ($p<0.10$), the food cost ($p<0.01$) and

occupational group 2 ($p<0.10$). The marital status of the respondent had an unexpected negative sign.

In the selection equation, almost all the variables that were statistically significant in the logistic model were also significant there. Furthermore, three of the four identifying variables that were hypothesised to explain the decision to state a positive WTP were significant in the bidding game data, except occupational group 4. In the BWFU, the incidence of other illnesses was the identifying variable that statistically significant in the selection equation, while the dummy of the expenditure to treat malaria was not. Finally, three of the four differentiating variables were statistically significant in the structured haggling model and the only statistically insignificant variable was the actual incidence of other illnesses.

The likelihood-ratio test showed that there were no statistical differences between the full and reduced models and thus the reduced models were valid. All the models from the three question formats' groups were statistically significant ($p<0.01$). The log likelihood was highest in the BWFU group, while the bidding and structured haggling groups were quite similar.

The chi-square tests for independent equations showed that while the Bidding and BWFU groups' two equations were independent ($p<0.05$), that of the structured haggling group was not ($p>0.05$). This is confirmed by the coefficient of correlation (rho), where there was very little correlation between the select and main Heckman's equations in the structured haggling group (0.27), while it was 0.90 and 0.95 in the bidding and BWFU groups respectively. Lambda was also insignificant in the structured haggling group's model.

The RESET test failed to find evidence of misspecification in the bidding game and BWFU models, but found that the model from the structured haggling was borderline mis-specified ($chi2 = 2.16$). The models from the three groups predicted raw WTP almost perfectly. Plots of the residuals versus fitted values showed the absence of heteroscedasticity in all models. While the residuals in the bidding and BWFU models were normally distributed, those in the haggling group deviated from normal.

Table 6.14 Heckman selection model of WTP

	BG		BWFU		SH	
Dependent variable	Log WTP		WTP		Log WTP	
Variables	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff.(SE)	Reduced Coeff. (SE)
Status in the household	-.13 (.09)	-.12 (.09)	-11.51 (21.11)		.21 (.11)*	.28 (.08)***
Number of household residents	-.03 (.02)*	-.03 (.02)	1.30 (3.19)	2.73 (2.97)	.02 (.02)	.03 (.02)*
Sex	.25 (.20)**	.24 (.20)**	17.06 (17.55)		.07 (.11)	
Age	-.001 (.002)		-.25 (.55)		.002 (.003)	
School years	-.01 (.01)		8.65 (2.26)***	9.24 (2.18)***		
Marital status	.29 (.15)**	.31 (.18)**	-13.74 (20.03)		-.22 (.17)*	-.20 (.11)*
Perceived incidence of malaria	-.07 (.10)		6.67(18.78)		.04 (.09)	
Actual incidence of malaria (hadmal)	.15 (.09)*	.05 (.05)	11.21(8.93)	11.30 (8.76)	.12 (.07)*	
Dummy of exp. to treat malaria (treatmal)	-.18 (.14)				-.16 (.13)	
Dummy of exp. to prevent malaria (malprev)	.19 (.10)	.19 (.10)**	-12.39 (14.70)		.07 (.11)	
Actual incidence of other illnesses (hadill)	.02 (.10)					
Dummy of exp. to treat other illnesses (treatill)	-.28 (.15)*	-.26 (.08)***	-4.05 (14.68)		.02 (.09)	
Previous purchase of nets	.06 (.10)		41.22 (10.03)**	37.60 (19.28)*	-.03 (.10)	
Ownership of radio	.10 (.09)		38.41 (20.88)*	44.80 (20.49)**	.14 (.10)	.14 (.10)
Ownership of bicycle	-.04 (.09)		13.80 (15.30)		-.11 (.09)	-.10 (.08)
Ownership of grinding machine	.23 (.13)*	.21 (.13)*	-53.20 (24.42)**	-58.99 (24.61)**		
Ownership of motorcycle	.16 (.10)	.18 (.10)*	-32.31 (24.42)*	-30.02 (17.15)*		
Ownership of car	-.16 (.18)	-.17(.17)	32.84 (27.63)	30.91 (27.75)	.12 (.13)	
Log of Food cost	-.03 (.08)		27.39 (12.41)**	28.15 (12.44)**	.11 (.07)	.12(.07)*
Occupational group 2: Farmer			-52.19 (42.52)	-60.54 (42.80)	-.41 (.27)	-.44 (.25)*
Occupational group 3: Skilled labourer/trading/pensioner			-39.28 (42.83)	-44.64 (43.32)	-.37 (.27)	-.40 (.25)
Occupational group 4: Formally employed			-131.56 (49.32)***	-139.28 (49.90)***	-.32 (.29)	-.38 (.27)
Occupational group 5: Professionals/mid & big time business			-176.05 (58.52)***	-188.50 (58.97)**	-.04 (.31)	-.06 (.30)
Constant	5.46 (.54)***	5.19 (.16)***	21.45 (98.76)	-12.78 (94.01)	4.76 (.58)***	4.75 (.56)***
Select						
Occupational group2	**.76 (.36)****	.78 (.37)**	.06 (.52)	.08 (.51)		
Occupational group 3	**.98 (.36)*****	.96 (.37)***	.44 (.53)	.50 (.53)		
Occupational group 4	**-.25 (.37)**	-.28 (.37)	-.42 (.60)	-.34 (.60)		
Occupational group 5	**1.14 (.59)***	1.11(.62)*	-.71 (.67)	-.64 (.65)		
Sex	-.17 (.18)	-.18 (.18)				
Married	.14 (.27)	.10 (.28)			-.27 (.35)	-.28 (.35)
Malaria prevention	.40 (.23)*	.43 (.23)*				
Had other illnesses	.04 (.17)	.04 (.18)	**.27 (.18)**	**.31 (.16)***	**.21 (.15)**	**.22 (.15)**
Grinding machine	-.01 (.29)	.01 (.28)-	.07 (.32)	.05 (.30)	**.86 (.44)***	**.87 (.44)****
Ownership of car	-.91(.33)***	.91(.33)***				
Log of food cost	.64 (.15)***	.68 (.14)***	.34 (.17)**	.29 (.16)*	.66 (.13)***	.66 (.13)***
Malaria treatment			**-.63 (.58)**	**-.44 (.53)**		
Years of schooling			.07 (.03)**	.06 (.03)**	**.07 (.03)*****	**.07 (.03)*****
Had malaria			.94 (.48)**	.82 (.44)*	.13 (.14)	.12 (.14)
Radio			.35 (.23)	.34 (.22)*		
Motorcycle					**-.61 (.30)****	**-.60 (.29)****
Constant	-4.32 1.01)***	-4.56 (1.0)***	-2.35 (1.12)**	-2.08 (1.10)*	-3.76 (.89)***	-3.74 (.89)***

/arthro	-1.69 (.31)***	-1.48 (.49)***	1.55 (.49)***	1.81 (.54)***	-.30 (.31)	-.28 (.30)
/lnsigma	-.47 (.06)***	-.49 (.06)***	4.71 (.07)***	4.74 (.06)***	-.64 (.06)***	-.63 (.06)***
Rho	-.93 (.04)	-.90 (.05)	.91 (.08)	.95 (.06)	-.30 (.29)	-.27 (.28)
Sigma	.62 (.04)	.61 (.04)	111.33 (7.32)	114.53 (7.19)	.53	.53 (.03)
			101.76	108.52	(.03)	
Lambda	-.58 (.05)	-.55 (.06)	(14.22)	(11.85)	-.16 (.16)	-.15 (.15)
No of observations	261	261	267	267	273	273
Censored	63	63	52	52	69	69
Uncensored	198	198	215	215	204	204
Log likelihood	-262.323	-264.023	-1384.304	-1385.927	-266.938	-265.473
Chi2 of model	39.15***	35.76***	62.46***	57.58***	47.24***	43.03***
Chi2 for test of indep. Eqns	11.00***	12.23***	5.43**	8.51***	0.12	0.95
Predicted values		5.38		201.89		5.34
Chi2 for Normality		0.58		0.42		7.09**
Ramsey RESET test		0.19		0.81		2.16**
Likelihood ratio test	Chi2 = 3.41		Chi2 = 3.25		Chi2 = 3.67	

p * <0.10, **<0.05, ***<0.01

Note: The values in bold are those of the identifying variables for each question format.

6.2.6 The Tobit and Truncated regression estimators

The results of these two models that did not perform as well as the Heckman and OLS models are provided in appendix 13 (Truncated regression) and appendix 14 (Tobit model).

6.3 Construct Validity of WTP for ITNs' re-treatment

The result of the modelling of WTP for ITNs re-treatment is presented here, together with the comparison of the performance of the WTP data from the three question formats. The full and reduced models are presented in Tables 6.15 and 6.16, while the robust estimations for the BWFU and structured haggling technique is presented in table 6.17.

Using the reduced models (Tables 6.16 and 6.17) for the interpretations, respondents' WTP for their own nets was statistically significant across the three question formats, and it also had the hypothesised positive sign. The other statistically significant variables were not uniform across the question formats, and occurred in either a pair or just singly in one of them.

The status of the respondent in the household was statistically significant in only the BWFU and structured haggling groups and it also had the hypothesised positive sign. Thus, household heads were more willing to pay for net re-treatment than their representatives. Also, while perceived incidence of malaria had the hypothesised positive sign in the bidding group, it had a negative sign in the BWFU group.

The marital status of the respondents and the actual incidence of malaria in the households were statistically significant only in the BWFU group, but had negative signs instead of the hypothesised positive sign. Conversely, number of years of formal schooling, dummy of the expenditure to treat other illnesses and ownership of a radio were all statistically significant, and with the expected signs.

Ownership of grinding machine had the expected positive sign in the bidding. As WTP for nets increased, the respondents were more likely to state higher WTP amounts for net re-treatment, though the economic effects in the three question formats were low. The economic effects of the rest of the statistically significant variables were not too considerable. The reduced models for the three elicitation methods were all statistically significant ($p<0.01$).

The regression statistics showed that the reduced OLS models explained more than 20% of the variation in WTP in all three groups (Table 6.16). However, plots of the residuals versus the fitted values, combined with estimation of the Cook-Weisberg statistics showed that the OLS models used in the BWFU and the structured haggling groups were not homoscedastic, although the Ramsey RESET test showed that all the models were correctly specified. Based on the regression diagnostics, the OLS regression best fits the data generated using the bidding game technique for WTP for net re-treatment, in comparison to the other question formats.

Table 6.15 Full OLS models for WTP re-treatment

Dependent variable = log WTP

Variables	BG Coeff. (SE)	BWFU Coeff.(SE)	SH Coeff.(SE)
Status in the household	-.14 (.09)	.22 (.11)*	.23 (.13)*
Number of household residents	-.02 (.02)	-.01 (.02)	.03 (.02)
Sex	.12 (.09)	-.13 (.10)	-.18 (.12)
Age	.0004 (.002)	.001 (.003)	-.001 (.003)
School years	-.007 (.009)	-.01 (.01)	.014 (.01)
Marital status	-.15 (.13)	-.21 (.12)*	-.15 (.12)
Perceived incidence of malaria	.19 (.11)*	-.19 (.10)*	.08 (.10)
Actual incidence of malaria	.13 (.10)	-.17 (.08)**	-.06 (.07)
Dummy of expenditure to treat malaria	-.22 (.15)	.23 (.15)	.21 (.14)
Dummy of expenditure to prevent malaria	.07 (.09)	.01 (.08)	.08 (.12)
WTP for personal net	.002 (.0003)***	.002 (.0003)***	.002 (.0003)***
Actual incidence of other illnesses	.04 (.09)	.08 (.06)	.09 (.08)
Dummy of exp. to treat other illnesses	-.13 (.13)	-.09 (.11)	-.25 (.14)*
Previous purchase of nets	-.03 (.09)	.11 (.10)	.02 (.12)
Ownership of radio	-.01 (.09)	.10 (.10)	.18 (.11)
Ownership of bicycle	-.02 (.09)	-.06 (.08)	.07 (.09)
Ownership of grinding machine	.20 (.11)*	.12 (.11)	.12 (.11)
Ownership of motorcycle	-.08 (.10)	-.08 (.09)	.01 (.11)
Ownership of car	-.09 (.15)	.07 (.13)	.002 (.15)
Log of Food cost	.01 (.06)	.02 (.06)	-.04 (.06)
Occupational group 2: Farmer	.17 (.21)	-.36 (.21)*	.29 (.31)
Occupational group 3: Skilled labourer/trading/pensioner	.35 (.21)*	-.30 (.21)	.28 (.31)
Occupational group 4: Formally employed	.12 (.23)	-.17 (.25)	.40 (.33)
Occupational group 5: Professionals/mid & big time business	.39 (.28)	-.36 (.30)	.23 (.35)
Constant	3.15 (.49)***	3.42 (.45)***	2.89 (.51)***
Number of Obs.	201	212	204
F statistic	3.32	3.04	3.62
P value	0.000	0.000	0.000
Adjusted R2	0.218	0.189	0.237
Standard Error	0.478	0.500	0.552
Cook-Weisberg	1.14 (p=0.2865)	13.06 (p=0.0003)	5.85 (p=0.0156)
Ramsey RESET	2.12 (p=0.0995)	0.76 (p=0.5151)	1.98 (p=0.1192)

Significance of parameters * <.10, **<.05, ***<.01

Table 6.16 Reduced OLS models for WTP retreat

Variables	BG Coeff. (SE)	BWFU Coeff.(SE)	SH Coeff.(SE)
Status in the household	-.11 (.08)	.21 (.10)**	.21 (.12)*
Number of household residents			.02 (.02)
Sex	.13 (.08)	-.12 (.09)	-.16 (.11)
School years			.01 (.008)*
Marital status	-.17 (.11)	-.19 (.11)*	-.15 (.12)
Perceived incidence of malaria	.20 (.10)**	-.20 (.10)**	
Actual incidence of malaria	.10 (.09)	-.14 (.07)*	
Dummy of expenditure to treat malaria	-.17 (.13)	.19 (.13)	.16 (.08)**
WTP for personal net	.002 (.0003)***	.002 (.0003)***	.002 (.0003)***
Actual incidence of other illnesses			.10 (.07)
Dummy of exp. to treat other illnesses			-24 (.13)*
Previous purchase of nets		.12 (.10)	
Ownership of radio			.20 (.10)**
Ownership of grinding machine	.16 (.10)*	.12 (.10)	.12 (.10)
Ownership of motorcycle	-.11 (.10)		
Constant	3.24 (.13)***	3.22 (.15)***	2.97 (.15)***
Number of Obs.	201	213	204
F statistic	8.05	7.43	7.88
P value	0.000	0.000	0.000
Adjusted R2	0.241	0.214	0.272
Standard Error	0.472	0.491	0.539
Cook-Weisberg	0.71 (p=0.4006)	8.24 (p=0.0041)	6.78 (p=0.0092)
Ramsey RESET	1.21 (p=0.3062)	0.36 (p=0.7784)	1.26 (p=0.2888)
Normality	0.14	0.05	0.17
Predicted values	3.396 (ref=3.482) diff=0.086	3.578 (ref=3.628) diff1=0.05	3.601 (ref=3.745) diff=0.144

The log transformation of the dependent variable and food cost did not correct the heteroscedasticity in the BWFU and structured haggling OLS models. Thus, robust estimates of the BWFU and Haggling models were done to produce homoscedastic standard errors (Table 6.18). The only difference noted in the robust models with reference to the behaviour of the parameters, was the actual incidence of malaria increased in significance from the 90% to the 95% level. There were also slight increases in the standard errors of the coefficients.

Table 6.17 OLS models of WTP for re-treatment with robust standard errors for BWFU and Structured haggling data

Variables	BWFU Coeff.(SE)	SH Coeff.(SE)
Status in the household	.21 (.09)**	.21 (.11)*
Number of household residents		.02 (.02)
Sex	-.12 (.08)	-.16 (.12)
School years		.01 (.007)*
Marital status	-.19 (.09)**	-.15 (.11)
Perceived incidence of malaria	-.20 (.10)*	
Actual incidence of malaria	-.14 (.08)**	
Dummy of expenditure to treat malaria	.19 (.12)	.16 (.09)*
WTP for personal net	.002 (.0003)***	.002 (.0003)***
Actual incidence of other illnesses		.10 (.07)
Dummy of exp. to treat other illnesses		-24 (.15)
Previous purchase of nets	.12 (.09)	
Ownership of radio	213	204.20 (.10)*
Ownership of grinding machine	.12 (.08)	.12 (.10)
Constant	3.22 (.14)***	2.97 (.16)***
Number of observations	213	204
F statistic	7.81	7.85
P value	0.000	0.0000
R2	0.248	0.311
Standard Error	0.491	0.539

6.4 Discussion

The discussion explores the interpretations of the qualitative research methods for asserting the extent of content validity of the question formats, while the implications for using different econometric models and the relative construct validity of data generated by different question formats are also explored. The discussion of the result of assessment of content validity is presented in the first section, while that for results of construct validity is presented in the second section.

6.4.1 Content validity

The assessment for content validity showed that although haggling was the most common price-taking behaviour in the market, the other two question formats were still used in market transactions in different situations. Nevertheless, the traders felt that the structured haggling technique would be the best method for determining willingness to pay.

However, if the buyer was conversant with the good and price, binary with follow-up was preferred, but if it is a new good, or the price is not known, then structured haggling was preferred. Thus, though haggling was mostly preferred, there was scope for the use of other types of question formats in order to have truthful responses, as found from the interviews with the traders and observing the market transactions.

An overarching factor that came out of the interviews with the traders and discussion with the consumers was that respondents must be adequately informed about the nets if reasonable and honest WTP would be elicited. This of course means that the scenario should be both comprehensive and understandable. This study addressed these concerns by using both pre and pilot testing in the latter stages of content validation.

The findings showed that the people believed that a survey methodology could be used to elicit the amount people would be willing to pay for goods and services. This signalled that CVM is potentially content-valid in south-east Nigeria, and possibly in other parts of the country. However, the techniques for eliciting consumers' choice varied in their relative content validity.

Question formats that were more realistic in mimicking the price-taking behaviour in the market were felt by both traders and consumers to be the most appropriate for eliciting valid WTP responses. Thus, the structured haggling technique was unanimously held to have the highest content validity, as it was the most realistic. However, the other two question formats could still be used, depending on the nature of market transactions, and thus they had some measure of content validity.

In order to increase the content validity of the question formats, both the traders and consumers suggested modifications to the structured haggling and BWFU techniques that would aid in eliciting the maximum amount the consumers are willing to pay for goods and services. The suggestion by the traders for increasing the realism of the structured haggling technique could be interpreted as an aid to help sellers extract the maximum WTP from the buyers. Conversely, the consumers' suggestion for the same purpose was to enable buyers that may actually wish to buy the good but stated low amounts to think more about what they really would pay. Thus, these suggestions balanced the supplier and consumer views in ensuring that the structured haggling would be content valid to both sets of people.

All the findings from assessing content validity of the question formats and the scenario were integrated into the final design of the study protocol - the scenario and structured haggling. The scenario was improved in that using burning of eyes to explain a side-effect of the ITNs was removed, since it was the wrong word as the nets do not burn the eyes, but give flu like symptoms in first week of use. Thus, the scenario was re-written to show that the nets could cause the flu-like symptoms. The treatment for the side effects was included in the scenario.

6.4.2 Construct validity of WTP for ITNs

Response rates

The response rates in the three groups were quite high at almost 90%. Most other WTP surveys have reported lower response rates. For instance, Johannesson et al. (1993) reported a response rate of 64%. Some people that refused said they had no time for the questionnaire. Some wanted the questionnaires left with them to complete, but this was impossible because the WTP questions had to be administered by an interviewer. In one instance, a man chased away one of the field-workers with a machete. However, it could be said that most of the refusals were because the respondents did not have the time to sit down for at least 30 minutes for the interview.

The questionnaires of respondents that provided incomplete information on the variables for construct validity were excluded from data analysis. The most common occurrence was that after the WTP elicitation, some respondents refused to continue with the interview. Thus meant that no data was collected on food expenditure and household assets among others. There were also a few respondents that omitted answering some demographic and other questions, and they were also eliminated from the analysis.

Similarity of the groups

In comparing three different question formats, "it is necessary to establish that differences in WTP elicitation using different methods are due to the different stimuli, and not due to external factors" (Loomis et al., 1996). The socio-economic and disease-specific characteristics of the three groups used for the questionnaire survey showed greater measures of similarity than heterogeneity.

The groups were similar in respect of the sex of the respondents. This was particularly important since females usually state lower prices for health goods (Brieger et al., 1989-90). However, groups differed according to food cost that was the proxy for income, with the BWFU being different from the two other groups ($p<0.05$), while the bidding game and haggling groups were homogenous.

The greater measure of similarity was not surprising since the respondents came from the same villages and were randomly assigned to the three different groups so that any variation seen in WTP could be attributed to the effects of the question formats. Nevertheless, dissimilarity may not be a problem if multiple regression-based testing shows that the differing variables are not significant determinants of WTP (Loomis et al., 1996). This was assessed by modelling the construct validity of the estimates of WTP.

The distributions of WTP estimates and independent variables for tests of construct validity

As expected, the decision of whether to pay was similar across the three groups. This was because the same starting-points were used and the question was exactly the same for all three groups. Thus, the effects of the question formats were not expected to come into play at this time.

The statistically significant differences in the levels of WTP generated by question formats showed that they possibly acted in different ways on the respondents' thought processes. The greatest level of variability, as indicated by the confidence intervals around mean WTP, was from the data from the structured haggling question format. This could be a reflection of the many steps involved thus giving rise to highly varied responses.

Modelling the construct validity of the stated WTP estimates requires the use of some of the socio-economic, malaria-specific and other illnesses data to explain the WTP. Understanding the resulting relationships and the choice of model specification requires that both the dependent variables and independent variables are described. This helps the selection of the correct functional forms for the analysis.

The fact that many respondents stated zero WTP means that the WTP estimates are limited dependent variables. This therefore means that the Tobit, Heckmans selection and log OLS models could be the appropriate estimators. The modelling requires linearity between the WTP estimates and the independent variables. This is aided if the WTP estimates are normally distributed. Thus, the WTP for ITNs were transformed to their logarithms in the bidding and haggling groups, while the BWFU was used in the original form since that was normally distributed. WTP for re-treatment was also transformed into the natural logarithms in all groups.

The natural logarithm of food cost, together with the levels of WTP in the bidding and structured haggling groups, became normally distributed after the transformation. The distributions of the continuous independent variables were such that they could be used directly for econometric modelling of WTP, with the exception of food cost that had to be log transformed to ensure that it is normally distributed. The expenditures on malaria treatment and prevention together with treatment of other illnesses had to be replaced by dummy variables because of the high numbers of respondents with zero expenditures on these items, thus potentially leading to majority of the cases been dropped from the econometric analyses.

Decision about the best econometric models for determining construct validity

The decision about the appropriate economic models to use for estimating models with limited dependent variables started with an understanding of the nature of the decision of respondents to either buy or not buy the nets. Then, the next steps to selecting and estimating the appropriate models involved specification searches and performing diagnostic tests on alternative models that could be used to estimate the full sample and only the positives.

The first steps were determining the natures of the zero WTP responses in the three groups. This was followed by the testing of the first assumption of the classical linear regression of linearity between the dependent and independent variables. This helped in the selection of the best functional forms of the variables for the modelling.

The regression diagnostic tests applied were graphic tests for heteroskedasticity for all models. In addition the Cook-Weisberg tests for OLS models was done. The RESET test was used to check for model mis-specification for all models. The normality of the generalised residuals of all the models was determined, together with a comparison of their predicted values with raw WTP values. All these diagnostic tests combined with the behaviours of the parameters in the models were used to determine the best models for interpreting the construct validity of the elicited WTP.

Nature of zero values

The investigation of zero WTP values suggested that the decision to pay for an ITN was propelled by need and enhanced by better income status. The question becomes whether those that stated zero WTP amounts also represented negative WTP or low positives. The information is important because it determines the choice of appropriate model for the analysis. Johannesson and Johansson (1997) have argued that it is impossible to have negative WTP for private medical commodities, since the consumer is free to accept or reject to 'buy'. Thus, the zero WTP values could be interpreted as true zeros, but the possibility of small positives that some of the respondents decided to state as zeros could not be ruled out.

That the decision to state a positive amount was related to ability to pay and need was confirmed by the findings of the logistic regression, as food cost was statistically significant in the three question formats' groups. It also had the hypothesised positive sign showing that as food cost rises, the likelihood of being willing to pay for ITNs also increases.

Parameters of asset holdings like ownership of a motorcar (Bidding group) or a motorcycle (Haggling group) though statistically significant had negative signs, instead of the hypothesised positive signs. An explanation for the divergence from the expected could be that because these people are mobile, they felt that they could easily drive to the major cities and purchase the nets, maybe, at presumed lower prices. However, in other settings, it could be argued that these people with a car could afford alternative means of protection against mosquitoes such as window screens and fans, and so might not be interested in ITNs.

Heckman selection model

The Heckman selection model best fit the data generated by the bidding game format. The results showed that there was no sample selection problem in the structured haggling data and so the correlation between the main and the selection equations was very small (0.27) in the structured haggling group compared with the very high correlation (>0.90) in the other two groups.

The test of independent equations was statistically insignificant in the structured haggling group. Furthermore, Ramsey RESET test showed that the Heckman model was the wrong model for the data from the structured haggling group. Therefore, while the Heckman's model were used to interpret the construct validity of the bidding and BWFU data, the log OLS was used for the data from the haggling group, as argued in Chapter 5.

All the statistically significant variables in the bidding group had the hypothesised signs, while four of the eight variables in the BWFU group had the hypothesised signs and four out of the five statistically significant variables in the structured haggling group had the hypothesised signs. Some explanations for these divergences have been already discussed in the models previously presented. The implication of these for construct validity is discussed later in this section.

Tobit model

The Tobit model did not fit the data generated by the haggling and bidding game methods. Nevertheless, most of the statistically significant parameters in the three question formats had the correct hypothesised signs. Food cost was the only variable that was statistically significant across the three elicitation groups, and with the right positive sign, thus amplifying the positive effect of income on WTP.

OLS and Truncated regression models

Using only the positive WTP data, OLS provided the best fit of the data generated from BWFU, while the truncated regression model was the best fit for the bidding game and structured haggling technique. In the truncated regression, the bidding game had the highest numbers of hypothesis-obeying statistically significant variables, the best predictions of WTP and the most correctly specified functional form.

The majority of the statistically significant parameters in both the OLS and Truncated regression models across all question formats had the hypothesised signs. Notable exceptions were car ownership in both the OLS and truncated regression models in the bidding game, ownership of grinding machine in truncated regression of the BWFU method and marital status in both the OLS and truncated regression models in the structured haggling technique. Car ownership, being indicative of a higher socio-economic group, had the unexpected negative sign (see discussion above).

The fact that food cost was statistically significant across all the models and question formats other than the truncated regression in the BWFU group showed that income is a major determinant of the amount that people would pay for nets. Actual malaria incidence, being statistically significant in 3 out of the six models also confirmed the characterisation of the zero WTP that higher malaria incidences would lead to higher stated willingness to buy nets for prevention of malaria.

Though Hoddinott (1994) argued in favour of the truncated regression models, the OLS models in this study proved to be a better specification for positive WTP and were more efficient than the truncated regression models based on the diagnostic tests. Compared with the adjusted R^2 obtained in previous studies, the OLS models in the question formats performed well in explaining variation in WTP.

Comparison of the models

The Heckman models showed some measures of similarity with the Tobit models in the three question formats. The Tobit models generally had fewer statistically significant variables, especially the hypotheses-obeying statistically significant ones, than the Heckman models. Nevertheless, 2 out of 5 statistically significant variables in the bidding game, 3 out of 4 statistically significant variables in the BWFU, and 1 out of 2 hypotheses-obeying statistically significant in the structured haggling technique were similar in the Tobit and Heckman models.

The Heckman models were generally better than the Tobit model, especially viewed from the fact that RESET test showed that the Tobit was the wrong model in the bidding and haggling groups. Donaldson et al (1998) also found that Heckman's model was better than the standard Tobit model. The consistent variables between the Heckman's and Tobit models were the expenditure to prevent malaria and ownership of grinding machine in the bidding game.

The problem with the Heckman model was the lack of guidance on the variables to be included in the first stage of the model. In this study, the a priori logistic regression result of the decision to state a positive WTP and OLS for positives, that were used to investigate variables that explain the levels of WTP (as explained in chapter 5), were used to select these variables. Almost all the variables that were statistically significant in the logistic model were also statistically significant in the first stage of the Heckman estimation.

The truncated and OLS models were very similar to each other within question formats, but exhibited greater measures of heterogeneity between different question formats. The OLS models mirrored more closely the models used to estimate limited dependent variables from the comparisons of the OLS model and Heckman MLE. However, as was also found by Donaldson et al. (1998), the Heckman model was better than the OLS in the bidding game and BWFU data. However, as Gyldmark and Morrison (2001) found, the log OLS was better than the Heckman model in modelling the structured haggling data.

The fact that different econometric models fit the data from the structured haggling group on one hand and the two other groups on the other hand, points to potential difference in the nature of the question formats. The lack of sample selection problem in the haggling group could be because the many steps involved helped stimulate the respondents to reveal their true choices. However, further studies would help unravel the differences or similarities in the hypothetical WTP data generated by the different question formats and it is premature to reach a conclusion with this preliminary study.

The ultimate criterion for selecting the best models rests on regression diagnostics. However, a structured approach is needed for determining the performance of the question formats for easy comparisons and generalising of findings. Previous attempts to classify the performance of models have not used a structured regimen. Johanesson et al. (1991) used indicators of goodness of fit and Donaldson et al. (1998) used regression diagnostic tests determine the performance of different models.

Modelling estimated WTP using different regression estimators is an excellent search process for determining the best models to use in estimating the parameters. The robustness of the different parameters across different models strengthens the interpretations and conclusions on the effect of the variables on WTP. Qualitative studies could also be used to understand why some variables deviate from the logical set hypotheses.

Although the Heckman selection model was the best performing model for limited dependent variables, no single model is perfect for all data sets as the log OLS was best for the haggling data. Potentially, different question formats generate different distributions of WTP, thus accounting for differences in the econometric models used to model them.

The mechanisms leading to the generation of different distributions of WTP by different question formats is not clear and requires further investigation. Those studies would provide information on why different question formats generate different distributions of WTP. One approach might be to de-brief people after the questionnaire to ask them what they thought about the question.

Evidence of construct validity of WTP for ITNs

The Heckman model provides the evidence of the construct validity of elicited WTP from the bidding game and BWFU data, while the log OLS is used to interpret the construct validity of the data generated by the structured haggling technique. Therefore, the three groups are not directly comparable for construct validity because of the different models used. The explanations for all these theoretical expectations were provided in Chapter 5.

In the bidding game, all the six statistically significant variables had the expected signs, showing a high level of construct validity of the elicited WTP. Thus, having a male respondent would lead to higher level of WTP than if the respondent was female and this would increase log WTP by 0.20. Also, married people were more likely to state higher levels of WTP than single people. The other variables that had the expected positive effect on WTP were occurrence of previous expenditures to prevent malaria, ownership of grinding machine and motorcycle. The occurrence expenditures to treat other illnesses decreased WTP for ITNs because household income that could have been used to purchase the nets had been depleted.

The data from the BWFU group was not as construct valid as that from the bidding game as only four out of the eight statistically significant variables had the expected signs. As expected, as the years of formal education increased, the level of WTP also increased. Also, expectedly, people who had ever bought untreated nets were more willing to pay for ITNs than others, and a unit increase in the variables would increase the level of WTP by 37.90 Naira. The cost of food and ownership of a radio were also positively related to WTP, showing that people belonging to higher socio-economic groups were expectedly more willing to pay for the nets.

The findings of stronger relationship of WTP with food cost and ownership of radio in the BWFU being indicative that WTP was positively associated with socio-economic class, was countered by the finding that respondents' whose households owned grinding machines and motorcycles were less willing to pay for the nets in contrary to expectations and the findings from the bidding group. Also, people belonging to the higher occupational groups (groups 4 and 5) that earn more income were also less willing to pay for the nets.

These findings viewed alone could lead one to conclude that the ITNs were perceived as inferior goods. However, another explanation is that some asset variables may not indicate a household's current socio-economic status since such assets could have been acquired when the households had money. However, the supportive evidence for the positive effect of higher socio-economic levels on WTP was provided by the data from the bidding group, where the assets were positively associated with WTP. These results could be argued as showing that there is no reason to expect groups to differ systematically, with respect to asset variables.

Of the four variables that were statistically significant in the structured haggling log OLS model, three had the expected signs. Of these, being a household head was associated with higher levels of WTP than if the valuation was not from a household head. Also, respondents whose households had experienced malaria a month prior to the survey were more likely to state higher levels of WTP. Similarly, households spending more on food were more willing to pay again reinforcing the income effect found from the data from other two question formats. However, unlike in the bidding game, single people were more willing to pay for the nets. A possible explanation for this divergence is that it is easier for single people to raise money for their net, since they have limited expenditures.

All in all, the findings from the three groups show that the CVM was able to elicit construct valid WTP, as most of the variables accorded with theoretical expectations. Some of the variables that diverged from expectations in the data from the one or two of the question formats had the expected signs in others, with the bidding game having the most construct valid variables. Larger sample sizes could potentially help to improve the power for statistical tests of construct validity of each question format. However, deeper understanding of the causes of the divergences of some variables from the hypothesised requires further study, so that better conclusions about construct validity could be made.

OLS for ITNs re-treatment

The OLS regression best fit the data generated from the bidding game group since that was the only instance where the model gave constant variance of the errors on the parameter estimates. The models used in both the BWFU and haggling data were heteroscedastic as shown by the plots of the residuals versus the fitted values and confirmed by the Cook-Weisberg tests.

The Ramsey RESET showed that the models were correctly specified for the three question formats. Robust estimation of the standard errors of the OLS models in the BWFU and haggling group was used to correct for the heteroscedasticity since the log transformation of dependent and some continuous independent variables did not solve the problem. Another method that could have been used was to estimate the model using weighted least squares, but it was felt that robust estimation would adequately demonstrate relationships of the independent variables with WTP for net re-treatment.

A major factor that validates both the WTP for ITNs re-treatment and the WTP for ITNs, was the highly statistically significant relationship between these two variables. Thus, as the WTP for nets rises, the level of WTP for its re-treatment also rises. This shows that the respondents were not just picking up values from the air, but were actually thinking of the goods in question and giving valid responses.

The other statistically significant variables were only so in pairs of question formats and not across the three groups. The highest number of hypotheses-obeying statistically significant variables was in the structured haggling group, while the BWFU group had the highest number of statistically significant variables whose signs ran counter to the set hypotheses. There, both perceived and actual incidences of malaria were negatively related to WTP for re-treatment. It was more surprising because actual incidence of malaria was consistently positively related to WTP for ITNs in some of the econometric models used to model the WTP for ITNs data across all question formats.

One reason why some of the explanatory variables behaved differently in construct validity of ITNs re-treatment when compared with that of ITNs, maybe that some respondents perceived the net re-treatment to have different characteristics. Alternatively, the time factor in the scenario about re-treating the nets after at least 6 months made them to think more of immediate needs, thus making them less willing to pay for net re-treatment. Further studies will be needed to get to the crux of this matter.

Conclusion

Determining the performance of the question formats and the econometric models may need to be more structured, for easy comparisons and generalisability of findings from different settings. Johanesson et al. (1991) chose one logistic regression model over another using indicators such as a log likelihood value closer to zero, higher chi-square, higher pseudo R^2 and a higher rate of individual prediction. Donaldson et al. (1998) used tests for normality, heteroscedasticity, and RESET to determine the performance of different econometric models for limited dependent variables.

In modelling WTP for ITNs, the Heckman model best fit the data generated by the Bidding game, the Tobit model and OLS best fitted data generated by the BWFU for WTP for ITNs and the Truncated regression best fitted data generated by the structured haggling technique. Similarly, the OLS regression best fitted the data generated from the bidding game for WTP for ITNs re-treatment. Hence, the within question formats and between question formats comparison of all the econometric models showed that the models fit all the data generated in different ways.

Thus, depending on the econometric model, different question formats will perform differently. Some perform better with some methods, and badly in others. Improving tests of construct validity will include better ways of measuring the variables, administering the questionnaires, as well as training and supervising field workers. Asserting the construct validity of WTP requires the diagnostic testing of different econometric models.

The inconsistencies noticed across models within the same question formats may signal that the variables were not measured optimally. To address this, various types of data transformations were done for the continuous variables before the final functional forms used in the regression were selected. It is hoped that this work will be viewed as

modelling for WTP data that emphasises comparison of different models and diagnostic testing before a final decision is made on the specific model to present to use in demonstrating the construct validity of stated WTP.

Future studies should be targeted at comparing different models for construct validity of stated WTP. These modelling should incorporate basic diagnostic tests. The stability of the parameters over the different models should also be tested. In cases where there are limit observations, it will be important to determine what the zeros actually mean. That is, whether they stand for true zeros, false zeros, protest votes or small positives, which the respondents decided to express as zeros, because the interpretation would decide on the models to use in determining construct validity of elicited WTP.

In cases where some variables deviate from the pre-set hypotheses and there is no way of explaining the deviations without using too many assumptions and speculations, it would be better to conduct qualitative interviews with people with the characteristics in order to understand these deviations. Also, studies are needed to determine the construct validity of different question formats by examining how the variables explain WTP, but these must be accompanied by diagnostic testing so that only parsimonious models are used for the interpretations.

Since data quality ultimately determines the performance of the models and the variables, future studies must ensure that they have enough trained and dedicated interviewers together with strict monitoring strategies during fieldwork. Allied to this, studies should ensure that the question formats and scenario they would choose to use, are content valid.

Further assessments of content and construct validity, together with criterion validity are needed before conclusive evidence on the performance of the question formats in rural Nigeria can be made. Like all CVM studies, a limitation is whether the scenario presented enough information to the respondents. The assessment of content validity and pre-testing done before the survey suggested that it did. However it is not easy to convey enough information in a short survey (Phillips et al., 1997).

CHAPTER 7: CRITERION VALIDITY

One of the objectives of this chapter is to provide an understanding about the extent to which stated and actual WTP converge, a major information gap in understanding whether the contingent valuation method (CVM) has criterion validity. This is achieved through the comparison of the criterion validity of three question formats. These comparisons are also used to understand the extent to which resembling the price-taking mechanism of the respondents affects criterion validity. Another major objective is to explain divergences between stated and actual WTP and to determine whether these imply lower criterion validity for CVM and for the different question formats. The last objective is to examine the determinants of actual purchases for ITNs and the policy implications of the findings.

The results presented in this chapter were derived from two different surveys and activities. The first set of data was from a survey, eliciting stated WTP and the second set of data were collected during the sale of ITNs for determining actual WTP. The sale of ITNs was accompanied by another survey that collected information on variables that could explain the divergences between stated WTP and actual WTP (purchase of the nets).

The results compare the stated WTP with the purchase of ITNs by the same respondents that were interviewed in the first survey. This will allow the criterion validity of both the CVM and the question formats to be assessed. The factors that could cause a divergence between stated and actual WTP are also presented and discussed, because they can provide information on whether such divergences always mean the presence of bias in the CVM and/or the different question formats. This has implications for whether or not WTP estimates can be said to be criterion valid even in the face of lower than expected criterion validity estimates. Within this framework, the factors that explained actual WTP are also presented and discussed.

Demographic and other characteristics of the respondents that were re-interviewed during the sale of the nets, together with a description of the special change factors that could cause a divergence between actual and stated WTP are first provided. Then, the descriptive statistics and results of econometric modelling of actual WTP are presented in the second section.

The third section links the result of stated WTP elicited in the first survey (presented in Chapter 6) with that of actual purchase of ITNs to provide the criterion validity results. The fourth section presents factors causing a divergence between actual and stated WTP, while the final section discusses the findings from the four sections.

7.1 General Descriptive Statistics

Contact was made with 205 (79%) of the 261 original respondents (those interviewed in the first survey) in the bidding game (BG), 215 (81%) of the 267 people in the binary-with follow-up (BWFU) group, and 195 (71%) of the 273 people in the structured haggling (SH) group during the sale of nets. It was possible to confirm that 3, 8 and 7 people in the BG, BWFU and SH groups respectively had either died or changed abodes before the sale of the nets. The rest of the original sample (53 in BG, 34 in BWFU and 71 in SH) were either not found or did not make contact with the field workers. One of the reasons was that the second survey was conducted during the harvesting season when many of the respondents had left their villages for the farm settlements. These settlements were an average of 5 miles from the villages, but because of lack of road access to those places and logistic support, it was not possible for the field workers to go and interview them there.

Of the people with whom contact was made, 196 (96%), 202 (94%) and 171 (88%) of the respondents in the BG, BWFU and SH groups respectively agreed to be re-interviewed. However, a few of the people that responded refused to answer many of the questions and their questionnaires were dropped from those analysed. Thus, the usable numbers of questionnaires for the analysis were 158, 166 and 144 in the BG, BWFU and the SH respectively.

A slightly higher proportion of the respondents in each of the three groups was female (Table 7.1). The three groups had similar characteristics with regards to the other demographic and socio-economic variables. However, the status of the respondent in the household, the average weekly cost of food to the household, the average monthly expenditure to treat malaria and the household ownership of radio differed across the question format groups. Also, farmers (occupational group 2) and skilled labourers/small time trading/pensioners (occupational group 3) were statistically significantly different across the three groups.

Table 7.1 Demographic and socio-economic factors of the respondents that were re-interviewed during the sale of ITNs

Variables	BG n (%)	BWFU n (%)	SH n (%)	X^2 (p value)
Respondents' status in the household 1 = head 0 = representative Total	 105 (66.5) 53 (33.5) 158 (100)	 143 (86.1) 23 (13.9) 166 (100)	 106 (73.6) 38 (26.4) 144 (100)	 17.57 (0..000)
# of household residents: Mean (S.D.)	3.96 (2.42)	4.31 (2.65)	3.85 (1.96)	2.19 (0.334)
Respondents' age (years): Mean (S.D.)	49.57 (14.26)	50.58 (15.04)	50.47 (15.53)	0.51 (0.774)
Sex of the respondent 1 = male 0 = female Total	 65 (41.1) 93 (58.9) 158 (100)	 80 (48.2) 86 (51.8) 186 (100)	 71 (49.3) 73 (50.7) 144 (100)	 2.65 (0.265)
Any formal schooling? 1 = yes 0 = no Total	 82 (51.9) 76 (48.1) 158 (100)	 96 (57.8) 70 (42.2) 166 (100)	 92 (63.9) 52 (36.1) 144 (100)	 4.61 (0.100)
Marital status of the respondent 1 = married 0 = never married Total	 128 (81.0) 30 (19.0) 158 (100)	 146 (88.0) 20 (12.0) 166 (100)	 118 (81.9) 26 (18.1) 144 (100)	 3.27 (0.195)
Household food cost: Mean (S.D.)	788.65 (679.98)	1156.62 (1227.55)	906.72 (594.03)	20.10 (0.000)
Malaria expenditures to treat malaria: Mean (S.D.)	348.28 (874.24)	323.28 (700.06)	136.61 (423.34)	19.13 (0.000)
Malaria prevention exp.: Mean (S.D.)	16.14 (51.09)	20.24 (74.46)	15.38 (64.96)	3.70 (0.157)
Whether household owns a radio 1 = yes 0 = no Total	 113 (71.5) 45 (28.5) 158 (100)	 134 (80.7) 32 (19.3) 166 (100)	 100 (69.4) 44 (30.6) 144 (100)	 6.43 (0.040)
Whether household owns a bicycle 1 = yes 0 = no Total	 104 (65.8) 54 (34.2) 158 (100)	 99 (59.6) 67 (40.4) 166 (100)	 80 (55.6) 84 (44.4) 144 (100)	 3.34 (0.188)
Whether household owns a motorcycle 1 = yes 0 = no Total	 27 (17.1) 131 (82.9) 158 (100)	 30 (18.1) 136 (81.9) 166 (100)	 32 (22.2) 112 (77.8) 144 (100)	 1.34 (0.512)
Whether household owns a motorcar 1 = yes 0 = no Total	 19 (12.0) 139 (88.0) 158 (100)	 15 (9.0) 151 (91.0) 166 (100)	 9 (6.3) 135 (93.8) 144 (100)	 2.97 (0.227)
Occupational group 1: Unemployed/unskilled lab/house wife 1 = yes 0 = no Total	 6 (3.8) 152 (96.2) 158 (100)	 5 (3.0) 161 (97.0) 166 (100)	 7 (4.9) 137 (951.) 144 (100)	 3.05 (0.217)
Occupational group 2: Farmer 1 = yes 0 = no Total	 99 (62.7) 59 (37.3) 158 (100)	 106 (63.9) 60 (36.1) 166 (100)	 66 (45.8) 78 (54.2) 144 (100)	 13.09 (0.001)
Occupational group 3: Skilled labourer/trading/pensioner 1 = yes 0 = no Total	 40 (25.3) 118 (74.7) 158 (100)	 37 (22.3) 129 (77.7) 166 (100)	 56 (38.9) 88 (81.1) 144 (100)	 11.67 (0.003)
Occupational group 4: Formally employed 1 = yes 0 = no Total	 7 (4.4) 151 (95.6) 158 (100)	 11 (6.6) 155 (93.4) 166 (100)	 9 (6.3) 135 (93.8) 144 (100)	 1.15 (0.562)
Occupational group 5: Professionals/mid & big time business 1 = yes 0 = no Total	 6 (3.8) 152 (96.2) 158 (100)	 9 (5.4) 157 (94.6) 166 (100)	 7 (4.9) 137 (95.1) 144 (100)	 0.78 (0.677)

7.2 Modelling of Actual WTP

The results of the actual WTP (purchase of ITNs) determined during the sale of ITNs are presented here. The first part compares the groups of respondents that bought nets versus those that did not. The last part presents the results of the econometric modelling of actual WTP.

7.2.1 Descriptive comparison of the actual WTP and respondents in the three different elicitation method groups

Few of the respondents with who contact was made during the sales of ITNs actually bought the nets and the numbers were 31 (19.6%), 41(21.7%) and 35 (24.3%) of the BG, BWFU and SH groups respectively (Table 7.2). There was no statistically significant difference in actual WTP across the three question formats (chi-square: 2.13 (p=0.346)).

In all question formats, household heads, males, older people and those that had formal education were more likely to buy nets. Being married or single had no effect on purchase of an ITN in the bidding game and the structured haggling groups, but did in the BWFU group.

Other socio-economic variables differed between the respondents who purchased a net versus those who did not for some question formats. Only the differences in status in the household and sex were statistically significant in the bidding game group, while the number of household residents and whether the respondent had formal education were significant in the BWFU group.

Higher stated WTP amounts in the first survey, together with higher expenditure on some key items (defined in chapter 5), were statistically significantly associated with higher probabilities of buying ITNs in all question formats. Respondents that had more malaria attacks together with those that had higher food costs were significantly associated with actual purchase of the ITNs in the BWFU group.

The statistically significant findings on the effects of occupational groups on actual WTP were varied. Farmers (occupational group 2) were less likely to purchase a net in the structured haggling group. Also, the regular wage earners (occupational group 4) were more willing to purchase a net in the structured haggling group. Being in occupational group 3 (skilled labourers/small trading/pensioners) increased the probability of purchasing of a net in the BWFU group.

People living closer to the net selling point were more likely to purchase a net in the bidding game and structured haggling groups. The number of minutes it took the respondents to get to the selling point using their normal means of transport was used as the proxy for distance.

Similarly, people with higher food costs bought more nets in the BWFU group. Also, people with more household assets bought more nets. However, the differences in ownership of a grinding machine were statistically significant in the bidding game and BWFU groups, while that from ownership of motorcycle was statistically significant in only the BWFU group. Ownership of bicycle and motorcar had no effect on the actual decision to buy a net.

Table 7.2: Comparisons of socio-economic status of two groups who bought and did not buy ITNs

	BG		BWFU		SH	
Variables	Didn't buy N=127 Mean (SD)	Bought N=31 Mean (SD)	Didn't buy N=125 Mean (SD)	Bought N=41 Mean (SD)	Didn't buy N=109 Mean (SD)	Bought N=35 Mean (SD)
Status in the household	.61 (.49)***	.87 (.34)	.85 (.36)	.90 (.30)	.73 (.44)	.74 (.44)
Number of household residents	3.84 (2.32)	4.42 (2.77)	4.09 (2.69)**	4.98 (2.42)	3.76 (1.89)	4.11 (2.17)
Sex	.37 (.48)**	.58 (.50)	.45 (.50)	.59 (.50)	.47 (.50)	.57 (.50)
Age	48.97 (14.35)	52.03 (13.86)	50.58 (15.63)	50.59 (13.29)	49.75 (15.30)	52.69 (16.25)
Attended school	.50 (.50)	.58 (.50)	.54 (.50)*	.71 (.46)	.61 (.49)	.74 (.44)
Marital status	.81 (.39)	.81 (.39)	.86 (.35)	.95 (.22)	.82 (.39)	.83 (.38)
Occupational group 1: Unemployed/unskilled lab/house wife	.04 (.20)	.03 (.18)	.04 (.20)	.00 (.00)	.04 (.19)	.09 (.28)
Occupational group 2: Farmer	.65 (.48)	.52 (.51)	.66 (.47)	.56 (.50)	.52 (.50)***	.26 (.44)
Occupational group 3: Skilled labourer/trading/ pensioner	.24 (.43)	.32 (.48)	.19 (.40)*	.32 (.47)	.37 (.48)	.46 (.51)
Occupational group 4: Formally employed	.03 (.18)	.10 (.30)	.08 (.27)	.02 (.16)	.03 (.16)***	.17 (.38)
Occupational group 5: Professionals/mid & big time business	.04 (.20)	.03 (.18)	.04 (.20)	.10 (.30)	.05 (.21)	.06 (.24)
Actual incidence of malaria	.41 (.61)	.61 (.72)	.72 (.93)**	1.07 (1.13)	.36 (.70)	.40 (.65)
WTP self	139.45 (109.07)***	259.68 (140.49)	177.52 (118.39)***	262.20 (118.73)	171.70 (143.99)***	371.71 (211.94)
Total expenditure	3274.45 (7005.53)**	5545.16 (10186.10)	4968.88 (13113.0)**	5788.05 (13009.77)	2835.79 (5183.43)**	12256.29 (25362.42)
Sales distance	17.47 (12.14)***	10.65 (11.22)	15.62 (15.21)	9.83 (7.34)	16.58 (12.15)**	12.14 (9.13)
Food cost	786.29 (727.20)	798.23 (448.12)	1034.02 (1130.0)***	1530.39 (1437.79)	872.69 (603.25)	1011.71 (559.92)
Ownership of radio	.71 (.46)	.74 (.44)	.76 (.43)**	.95 (.22)	.66 (.48)	.80 (.41)
Ownership of bicycle	.66 (.48)	.65 (.49)	.58 (.49)	.63 (.49)	.56 (.50)	.54 (.51)
Ownership of grinding machine	.06 (.24)*	.16 (.37)	.10 (.31)**	.24 (.43)	.15 (.36)	.20 (.41)
Ownership of motorcycle	.14 (.35)**	.29 (.46)	.16 (.37)	.24 (.43)	.20 (.40)	.29 (.46)
Ownership of car	.13 (.33)	.10 (.30)	.08 (.27)	.12 (.33)	.05 (.21)	.11 (.32)

Significance of differences between those that bought and that did not buy: * <0.10, **<0.05, ***<0.01

7.2.2 Logistic regression of actual WTP with explanatory variables

The distance from the sales location to individual households was the only variable that was statistically significant in all the three question formats' groups (Table 7.3). It was highly significant at $p<0.01$ in both the bidding and BWFU groups, and $p<0.05$ in the haggling group. The coefficient also had the hypothesised negative sign showing that the farther the net selling point to the respondents, the less likely it is that they would buy a net.

The marginal effects of the statistically significant variables are used to explain their effects on actual WTP, instead of log odds of the coefficients as marginal effects is more informative, as they are probabilities of an event occurring. Thus, a unit reduction in the number of minutes it took the respondents to get from their homes to the sales spot will increase the probability of the respondent to actually purchase a net by 0.01 in all question formats.

The logarithm of total expenditures made in the respondents' households in the period between determining the stated and actual WTP had a small but positive effect on WTP in both the bidding game and structured haggling groups, with the marginal effects of 0.001 in both groups. The positive signs of the coefficients are indicative that higher expenditures could mean that the households had higher income and thus, were better able to actually buy the ITNs.

In the bidding game group, other statistically significant variables that explained actual WTP were household heads, people who had some formal education and those that owned a motorcycle and their marginal effects on actual WTP were 0.14, 0.10 and 0.20 respectively. Respondents in these categories were more likely to buy a net and all the coefficients had the hypothesised positive signs. Here, if a household head responded, there would be 0.14 increase in the probability of actual purchase of a net and increasing the ownership of motorcycle by a unit would increase the probability of actual purchase of a net by 0.20.

The possession of a radio by a household, a recent episode of malaria in the household, and a unit increase in age of the respondent would increase the probability of net purchase by 0.16, 0.06 and 0.005 respectively, in the BWFU group.

The ownership of a car in the structured haggling group would increase the probability of purchasing a net by 0.33. However, ownership of a grinding machine, though being statistically significant in this group had the unexpected sign, and its possession would decrease the probability of purchasing a net by 0.21. Also, being a farmer would decrease the probability of purchasing a net by 0.34.

Tests of model diagnostics showed that the reduced models for the three question formats were all statistically significant ($p<0.01$) and were able to predict more than 75% of the observations.

The likelihood ratio test showed that the full and reduced models were the same in the three question formats' groups. The probability of purchasing a net was highest in the structured haggling group (0.20), followed by the BWFU (0.17) and the bidding game group (0.14).

Table 7.3 Logistic regression of actual WTP with explanatory variables

Dependent variable: ActualWTP 1 = purchased, 0 = did not purchase

	BG		BWFU		SH	
Variables	Full Model Coeff. (SE)	Reduced model Coeff.(SE)	Full Model Coeff. (SE)	Reduced model Coeff.(SE)	Full Model Coeff. (SE)	Reduced model Coeff. (SE)
Status in the household	1.12 (.76)	1.26 (.61)**	.49 (.94)		-.10 (76)	
Number of household residents	-.003 (.10)		-.04 (.10)		-.01(.14)	
Sex	-.49 (.67)		-.06(.53)		-.04 (.74)	
Age	.01 (.02)		.03 (.02)	.03 (.02)*	.03 (.02)	.02 (.02)
School	.71 (.61)	.80 (.47)*	.58 (.56)	.68 (.51)	.20 (.69)	
Marital status	-.27 (.68)		1.40 (.87)	1.28 (.86)	-.21 (.69)	
Actual incidence of malaria	.48 (.37)	.43 (.35)	.39 (.24)*	.41 (.22)*	.17 (.39)	
Food cost	.0002 (.0004)		.0003 (.0002)	.0002 (.0001)	.0002 (.0005)	
Total expenditure	.0001 (.00003)*	.0001 (.00003)*	-.00002 (.00002)		.0001 (.00003)**	.0001 (.00003)**
Sales distance	-.06 (.03)**	-.06 (.02)***	-.08 (.03)***	-.08 (.03)***	-.06 (.03)**	-.06 (.03)**
Ownership of radio	.021(.59)		1.49 (.82)*	1.53 (.80)*	.95 (.70)	.61 (.59)
Ownership of bicycle	-.32 (.55)		-.47 (.47)	-.53 (.46)	-.53 (.61)	
Ownership of grinding machine	.05 (.78)		.76 (.69)		-2.04 (.92)**	-2.00 (.90)**
Ownership of motorcycle	.99 (.64)	1.14 (.60)*	.79 (.62)	.71 (.59)	-.45 (.64)	
Ownership of car	-1.26 (1.00)	-1.14 (.87)	-.25 (.83)		1.81 (.99)*	1.54 (.90)*
Occupational group 2: Farmer	-.39 (1.37)		1.64 (1.80)	1.27 (1.74)	-2.31 (1.10)**	-2.27 (1.02)**
Occupational group 3: Skilled labourer/trading/ pensioner	-.44 (1.43)		2.15 (1.79)	1.90 (1.75)	-1.25 (1.09)	-1.24 (.99)
Occupational group 4: Formally employed	.34 (1.67)		.01 (2.13)	-.42 (2.06)	1.30 (1.32)	1.30 (1.25)
Occupational group 5: Professionals/mid & big time business	-.95 (1.85)		2.46 (1.97)	1.95 (1.90)	-1.39 (1.48)	-1.11 (1.31)
Constant	-2.03 (1.75)	-2.46 (.73)***	-7.37 (2.65)***	-6.67 (2.45)***	-1.08 (1.67)	-.69 (1.29)
Number of cases	158	158	166	166	144	144
Chi2	28.57	26.32	39.71	37.11	38.237	36.05
P value	.0730	.0004	.0036	.0004	.0055	.0001
Pseudo R2	.183	.169	.216	.202	.242	.228
Log likelihood	-63.497	-64.847	-72.080	-73.377	-59.896	-60.989
Correct predictions	82.69%	82.80%	76.07%	75.46%	81.56%	80.14%
Model's marginal effect	0.14		0.17		0.20	
Likelihood-ratio test chi2 (p value)	6.45 (0.89)		4.36 (0.50)		13.06 (0.16)	

Significance of parameters * <0.10, **<0.05, ***<0`.01

7.3 Criterion Validity: (Correlation of Stated and Actual WTP).

More than 75% of the respondents in all three question formats' groups had positive stated willingness to pay for the ITNs, with the highest proportion of 80.5% coming from the BWFU group as determined in the first survey (see Table 6.5 in chapter 6). An analysis of the positive WTP responses showed that very small proportions were willing to pay the actual price at which the nets were offered for sale (350 Naira). The lowest proportion was in the BG, where only 19/261(7.3%) of the respondents stated amounts that were enough to cover the price of the net. The proportion was higher with the SH and BWFU groups with 36/273(13.2%) and 36/267 (13.5%) respectively.

Of the respondents whose stated WTP amounts were at least equal to the sale price of the ITNs and that were available for determination of actual WTP, 11 (73.33%), 13(61.91%) and 26 (74.29%) of the BG, BWFU and SH groups actually bought the nets (Table 7.4 – third column). These proportions are also the positive predictive validity of the three question formats. Table 7.4 (third column total) also shows that more sales were recorded with the BWFU group since many of those who stated they were unwilling to pay in the first survey actually bought nets.

It should be noted that the sample size in the first column of Table 7.4 exceeded that in Table 7.1 because there were 3, 1 and 6 original respondents in the first survey from the BG, BWFU and SH groups that bought the nets, but refused to be interviewed. Thus, because information on their earlier stated WTP was available, they were included in testing criterion validity but excluded from other analysis that required detailed demographic and socio-economic information.

The phi coefficient as the test of criterion validity showed that the SH was the most valid followed by the BG and the BWFU (Table 7.5). The phi correlation coefficient was highest with SH with a value of 0.61 (95% confidence interval 0.50 – 0.71). It was followed by the BG with a phi coefficient of 0.42 (95% confidence interval 0.29 – 0.54), while the BWFU was the least with a phi coefficient of 0.32 (95% confidence interval of 0.20 – 0.44). All phi coefficients were statistically significant ($p<0.05$).

However, the confidence intervals of the SH and BG overlapped, so the SH could not be said to be clearly better than the BG, though it was better than the BWFU. Similarly, the confidence intervals of the BG and BWFU overlapped suggesting that that they are not likely to be different from each other.

The positive predictive validity of the SH and bidding game were close to each other though the SH was slightly higher and the BWFU the lowest. However, there was no statistical difference in predictions across the three question formats ($p>0.05$).

Table 7.4 Comparison of stated and actual WTP of those seen in the second survey

	Bought a net			Did not buy a net			Total (a, b, c)		
	BG	*BWFU*	*SH*	*BG*	*BWFU*	*SH*	BG	BWFU	SH
Stated ≥350	11	13	26	4	8	9	a. 15	21	35
Stated WTP<350	22	29	13	124	117	102	b. 146	146	115
Total(a, b, c)	a. 33	42	39	b. 128	125	111	c. 161	167	150

Column total a: total number of respondents that bought nets;

Column total b: total number of respondents that did not buy nets

Column total c: total number of respondents seen for determination of actual WTP

Row total a: total number of respondents that were hypothetically WTP;

Row total b: total number of respondents that were hypothetically unwilling to pay

Row total c: total number of respondents seen for determination of actual WTP

Table 7.5 Comparison of criterion validity of the question formats

	BG	BWFU	SH
Positive predictive validity	73.33%	61.91%	74.29%
Overall correct prediction	83.85%	77.84%	85.33%
Phi coefficient (95% C.I.)	0.42 (0.29 - 0.54)	0.32 (0.20 - 0.44)	0.61 (0.50 - 0.71)
P value for phi coefficient	<0.05	<0.05	<0.05

7.4 Factors causing a divergence between Stated and Actual WTP

The first part of this section provides the descriptions of the respondents with and without divergences (i.e. changed their mind) between stated WTP in the first survey and actual purchase of nets. The three groups of respondents were classified as those with an increase in actual WTP (i.e. stated WTP was lower than the price of a net but bought a net); with a decrease in actual WTP (i.e. stated WTP was equal or more than the price of a net, but did not buy a net); and the same preferences (i.e. WTP did not change). The comparison also shows whether socio-economic and other characteristics of the above three groups of respondents were statistically different.

Bivariate tests of the independent variable versus a dependent variable called Divergent WTP (with values 1 = decrease; 2 = increase; 3 = same) as described in Chapter 5 were performed. This was to see what variables were statistically significantly related with the direction of divergences in actual WTP over stated WTP. Further bivariate tests of the effect of the special change variables on divergences in WTP are also presented.

7.4.1 Comparison of the causes of divergences in WTP in the three question format groups

The greatest proportions of respondents were WTP either increased or decreased were in the BWFU group (78.73% or 37/47), when compared to the other two groups (Table 7.6). The numerator is the total number of respondents with divergences in WTP, while the denominator is the total number of people that actually purchased ITNs. The proportion with divergent WTP in the bidding game was 69.4% or 25/36 and in the structured haggling group, it was 48.8% or 21/43.

The respondents were categorised into three groups to aid the understanding of causes of the divergences: those with decrease in WTP in the presence of divergence; those with increase in WTP in presence of divergence; and those whose WTP did not diverge between stated and actual WTP (same WTP). Most of the socio-economic and other variables did not significantly differ across the three groups of respondents.

The stated WTP amount in the first survey was the only variable that explained the divergences in WTP in the three question formats ($p<0.01$). Comparing those with a decrease in WTP and those with an increase in WTP, the higher the stated WTP in the first survey, the more likely that actual WTP would be lower. However, in the structured haggling group, the difference in mean WTP between those with a decrease in WTP and those where WTP did not diverge was negligible.

Comparing those with a decrease in WTP and those with an increase in WTP, respondents from households that incurred higher food costs were more likely to increase their actual WTP in the BWFU and the structured haggling groups. However, those with non-divergent WTP had the highest food costs in those two question format groups ($p<0.05$). Male respondents were less likely to have a divergent WTP in the BG and BWFU group ($p<0.05$).

The other variables that explained the different characteristics of the three groups of respondents were either statistically significant in a pair of question formats or just in a single question format. Most of these variables were found in the BWFU group. The marital status of the respondent was only statistically significant in the BWFU group ($p<0.10$).

If the respondent was a household head, there was a higher likelihood of the divergences in WTP leading to increases. Also, respondents whose households had recent attacks of malaria were more likely to increase their WTP in BWFU group ($p<0.10$ and $p<0.01$ respectively).

In the bidding game, respondents living nearer the places the nets were sold were more likely to increase their WTP ($p<0.01$). Also, in the bidding game, people that increased their actual WTP were more likely to belong to households with a radio ($p<0.05$), while those with the least total expenditures on unplanned items were more likely to increase their WTP in the BWFU group ($p<.10$).

Table 7.6 Distribution of the directions of the divergences in WTP

	BG			BWFU			SH		
Variables	Decrease N = 4 Mean (SD)	Increase N = 21 Mean (SD)	Same N= 11 Mean (SD)	Decrease N = 8 Mean (SD)	Increase N= 29 Mean (SD)	Same N = 10 Mean (SD)	Decrease N = 9 Mean (SD)	Increase N = 12 Mean (SD)	Same N = 22 Mean (SD)
Status in the household	1.00 (0)	.76 (.44)	1.00 (.00)	.63 (.52)	.86 (.35)	.1 (.0)*	.78 (.44)	.67 (.49)	.77 (.43)
Household residents	5.25 (3.30)	4.19 (2.06)	5.00 (3.77)	4.00 (2.62)	4.83 (2.73)	5.20 (1.55)	4.33 (2.06)	3.17 (1.70)	4.50 (2.24)
Sex	.75 (.50)	.38 (.50)	.91 (.30) **	.50 (.53)	.45 (.51)	.90 (.32) **	.56 (.53)	.42 (.51)	.64 (.49)
Age	57.25 (8.62)	48.19 (13.95)	58.27 (11.31)	40.00 (12.74)	51.17 (13.40)	48.80 (13.97)	45.00 (12.68)	51.50 (17.85)	53.18 (16.09)
Attended school	.50 (.58)	.57 (.51)	.64 (.50)	.63 (.52)	.66 (.48)	.90 (.32)	.78 (.44)	.83 (.39)	.73 (.46)
Marital status	1.0 (.00)	.76 (.44)	.91 (.30)	.75 (.46)	1.0 (.00)	.80 (.42) **	.89 (.33)	.75 (.45)	.86 (.37)
Occupational group 1: Unemployed/unskilled lab/house wife	.00 (.00)	.1 (.22)	0 (0)	.00 (.00)	.00 (.00)	.00(.00)	.00 (.00)	.17 (.39)	.1 (.21)
Occupational group 2: Farmer	.50 (.58)	.57 (.51)	.36 (.50)	.50 (.53)	.62 (.49)	.50 (.53)	.44 (.53)	.17 (.39)	.27 (.46)
Occupational group 3: Skilled labourer/trading / pensioner	.50 (.58)	.29 (.46)	.36 (.50)	.38 (.52)	.28 (.45)	.30 (.48)	.44 (.53)	.50 (.52)	.45 (.51)
Occupational group 4: Formally employed	.00 (.00)	.1 (.30)	.18 (.40)	.00 (.00)	.1 (.19)	.00(00)	.11 (.33)	.17 (.39)	.18 (.39)
Occupational group 5: Professionals/mid & big time business	.00 (.00)	.00 (.00)	.1 (.30)	.13 (.35)	.1 (.23)	.20(.42)	.00 (.00)	.00 (.00)	00 (.00)
Actual malaria incidence	.50 (1.0)	.57 (.68)	.73 (.79)	.38 (.74)	1.14 (1.22)	.70 (.67) *	.44 (.73)	.58 (.79)	.27 (.55)
Stated WTP	462.5 (62.92)	190.48 (94.37)	386.36 (118.51) ***	425.00 (26.73)	205.17 (81.66)	390.00 (69.2) ***	488.89 (116.67)	156.67 (101.74)	497.73 (149.98) ***
Expenditure on unplanned items	3300.0 (3311.6)	4685.7 (10606.8)	6863.64 (9168.13)	3387.5 (7140.1)	1316.9 (519.4)	4565.00 (5242.35) *	1761.11 (2120.9)	6801.7 (19930.0)	15788.64 (28216.5) **
Sales distance	16.25 (8.54)	9.86 (10.69)	11.64 (12.26) ***	11.25 (14.17)	10.52 (8.20)	9.98 (8.79)	13.11 (8.78)	15.67 (12.56)	10.32 (6.43)
Food cost	671.25 (378.56)	732.38 (441.20)	933.18 (430.12)	821.3 (391.9)	1316.9 (519.4)	2175.00 (2780.0) **	640.56 (290.54)	690.00 (253.66)	1151.36(603.92) **
Radio	.25 (.50)	.86 (.36)	.55 (.52) **	1.0 (.00)	.93 (.26)	1.0(.00)	.67 (.50)	.83 (.39)	.77 (.43)
Bicycle	.50 (.58)	.52 (.51)	.82 (.40)	.75 (.46)	.66 (.48)	.50 (.53)	.67 (.50)	.58 (.51)	.50 (.51)
Grinding machine	.00 (.00)	.1 (.30)	.27 (.47)	.25 (.46)	.24 (.44)	.20 (.42)	.22 (.44)	.25 (.45)	.18 (.39)
Motorcycle	.00 (.00)	.19 (.40)	.45 (.52)	.13 (.35)	.25 (.46)	.20 (.42)	.11 (.33)	.33 (.49)	.27 (.46)
Car	.00 (.00)	.1 (.22)	.27 (.47)	.25 (.46)	.14 (.35)	.1 (.32)	.00 (.00)	.00 (.00)	.5 (.21)

Significance of differences between three divergent WTP groups using the Kruskal-Wallis test: * <.10, **<.05, ***<.01

Only stated WTP explained divergences in WTP across the three question formats (Table 7.7), using non-parametric tests to determine the socio-economic variables that were bivariately significantly related to divergences in WTP. Food cost was statistically significant in the BWFU and SH groups, while sex was statistically significant in the BG and BWFU. The other variables were only statistically significant in individual question formats. Ownership of a radio in the BG, unplanned expenditures in SH bidding game, and the status of the respondent in the household, marital status, together with actual occurrence of malaria in the BWFU group explained divergences in WTP.

Table 7.7 Statistically significant socio-economic and demographic factors explaining divergences in WTP

	BG Chi-square (p values)	BWFU Chi-square (p values)	SH Chi-square (p values)
Status in the household		4.90 (0.086)	
Sex	8.55 (0.014)	6.12 (0.047)	
Marital status		7.03 (0.030)	
Actual malaria incidence		4.82 (0.090)	
Stated WTP	22.13 (0.000)	33.78 (0.000)	25.63 (0.000)
Food cost		6.86 (0.032)	7.53 (0.023)
Expenditure on unplanned items			6.50 (0.039)
Ownership of a radio	7.29 (0.026)		

7.4.2 Effect on divergent WTP by the special change factors

As described in the chapter on methods (section 5.6.5), chi-square analysis through a cross-tabulation of divergent WTP and these special change factors was used to investigate whether there were statistically significant relationships. These special change factors were also explained in detail in section 5.6.5. Table 7.8 shows that the total number of respondents for the special change factors varied because the questions were designed differently as already described.

Multiple expenditure, the distance of the respondents' household to the point of sale of the nets, "time to think effect" and multiple interviews were measured using just one question to all the respondents. The rest of the variables were determined using two questions. The first question determined whether the variable being measured had any effect on WTP. If the answer to the first question was in the affirmative, then a follow-up question was asked and their answers on the direction of change classified into decreased, increased and no effect.

Two of the variables explained divergences in WTP in the BWFU group, while one of the variables explained the divergences in the bidding group and none in the structured haggling group respectively. The external information the respondents received and the time the respondents had to think about their WTP accounted for divergences in the BWFU ($p<0.05$). On the other hand, the attitude of the community leaders led to divergences in WTP in the bidding group ($p<.05$). The potential effect of these three variables in the two question format groups led to an increase in WTP.

Table 7.8 Result of cross-tabulation of divergences between stated and actual WTP with special change factors

Change Factor	BG		BWFU		SH	
	n	(%)	n	(%)	n	(%)
Income	Decreased = 4	(16)	Decreased = 12	(40.0)	Decreased = 19	(65.5)
	Increased = 9	(36)	Increased = 5	(16.7)	Increased = 4	(13.8)
	The same = 12	(48)	The same = 13	(43.3)	The same = 6	(20.7)
	Total = 25	(100)	Total = 30	(100)	Total = 29	(100)
	(Chi-square = 2.81; p=0.59)		(Chi-square = 6.95; p=0.139)		(Chi-square = 1.98; p=0.74)	
Time to think	Decreased = 2	(5.6)	Decreased = 7	(14.9)	Decreased = 4	(9.3)
	Increased = 3	(8.3)	Increased = 21	(44.7)	Increased = 13	(30.2)
	The same = 31	(86.1)	The same = 19	(40.4)	The same = 26	(60.5)
	Total = 36	(100)	Total = 47	(100)	Total = 43	(100)
	(Chi-square = 5.87; p=0.21)		(Chi-square = 10.88; p=0.03)		(Chi-square = 5.74 p=0.22)	
Multiple interviews	Decreased = 0	(0)	Decreased = 2	(4.3)	Decreased = 0	(0)
	Increased = 7	(19.4)	Increased = 22	(46.8)	Increased = 11	(25.6)
	The same = 29	(80.6)	The same = 23	(48.9)	The same = 32	(74.4)
	Total = 36	(100)	Total = 47	(100)	Total = 43	(100)
	(Chi-square = 2.752; p=0.25)		(Chi-square = 2.28; p=0.68)		(Chi-square = 1.40 p=0.50)	
Attitude of community leaders	Decreased = 0	(0)	Decreased = 3	(17.6)	Decreased = 1	(6.3)
	Increased = 9	(64.3)	Increased = 3	(17.6)	Increased = 11	(68.7)
	The same = 5	(35.7)	The same = 11	(64.7)	The same = 4	(25.0)
	Total = 14	(100)	Total = 17	(100)	Total = 16	(100)
	(Chi-square = 8.19; p=0.02)		(Chi-square = 2.94; p=0.57)		(Chi-square = 3.39 p=0.50)	
External information	Decreased = 1	(4.5)	Decreased = 1	(2.8)	Decreased = 0	(0)
	Increased = 11	(50.0)	Increased = 26	(72.2)	Increased = 18	(78.3)
	The same = 10	(45.5)	The same = 9	(25.0)	The same = 5	(21.7)
	Total = 22	(100)	Total = 36	(100)	Total = 23	(100)
	(Chi-square = 2.80; p=0.59)		(Chi-square = 10.39; p=0.03)		(Chi-square = 0.65 p=0.72)	
Neighbours effect	Decreased = 1	(10.0)	Decreased = 0	(0)	Decreased = 2	(25.0)
	Increased = 1	(10.0)	Increased = 0	(0)	Increased = 1	(12.5)
	The same = 8	(80.0)	The same = 8	(100)	The same = 5	(62.5)
	Total = 10	(100)	Total = 8	(100)	Total = 8	(100)
	(Chi-square = 2.56; p=0.63)		No chi-square calculated		(Chi-square = 4.13 p=0.39)	

Note: Divergences measured as 1 = decrease; 2 = increase and 3 = Same.

7.5 Discussion

This discussion starts by examining the factors that explained actual WTP. The actual criterion validity results are then discussed. The discussion thereafter explores why stated and actual WTP could diverge, together with the implications of divergences in WTP for asserting the criterion validity of CVM, together with that of the different question formats. Conclusions are given in the final section.

7.5.1 Actual purchase of ITNs

The proportions of the respondents that actually bought the nets exceeded those that stated willingness to pay amounts enough to cover the price of the nets. The respondents did not know the actual price of the nets when their stated WTP was determined, so it could be questioned whether the increase of actual WTP (purchase of nets) over stated WTP represented strategic behaviour. When these proportions of stated and actual WTP are scrutinised further across the three question formats, it is seen that while 15 people were hypothetically willing to pay the sale price in the bidding game group, 33 people actually bought the net, an increase of more than 100%. The increase was even more with the BWFU, but less in the structured haggling group.

Most people who did not buy ITNs were from the lower socio-economic groups and this is understandable from the factors that affect demand. Thus, one could argue that the ITNs are not inferior goods in this population, since demand for them rose with income, using food cost as a proxy measure. Nevertheless, it could be argued that the income range one finds in rural villages may be too low, to inform whether the ITNs are inferior goods.

However, these findings reveal the positive effects of income on actual WTP. This is confirmed by the effect of the household assets, where generally people with more household assets were more likely to buy a net than those without. The relationship is reinforced by the result found in the first survey where stated WTP was determined and this is a piece of evidence that supports the validity of the contingent valuation technique.

Subsistence farmers who formed a major part of the lower socio-economic group, indeed, bought fewer nets. This is understandable because the farmers income is quite seasonal, and many a times especially during the planting seasons, farmers may not have enough money to spare for nets. Even during the harvesting season (when the nets were sold), they may also not have enough money if they had a poor harvest or are still paying off debts accumulated during the planting season. The occupational groups that earned regular monetary income like small traders and skilled labourers (occupational group 3) together with people employed in formal regular wage-earning sectors (occupational group 4) were more likely to purchase the nets. However, this was not necessarily the case with formal regular wage-earning sectors (occupational group 4) in the BWFU group.

Apart from the socio-economic factors, a direct factor that influenced some respondents to buy the nets was their immediate malaria experience (though only statistically significant in the BWFU group). As people get more malaria, their desire to protect themselves with ITNs increased. This desire was revealed as actual purchases of the nets. It could thus be argued that net coverage would be higher in areas with intense malaria transmission like in most parts of Nigeria. However, at the individual or household level, those that are more vulnerable to malaria such as households with pregnant women, children under five years, immuno-suppressed people etc., would form the majority of actual purchasers of the nets.

The further the people lived from the sales points, the less willing they were to actually buy the nets. This was a consistent finding in the three question formats and was confirmed in the logistic regression analysis. Acton (1975) showed that health care utilisation declines as travel time to a medical facility increases. Phelps (1992) from analysing the effect of health insurance policy on demand showed how time acts as a cost of care. In fact, almost all of the studies on the demand for health care in developing countries cited travel time as an important and significant determinant of demand for medical care (Gertler and van der Gaag, 1990). While these cited studies were for curative care, it is hypothesised that same factors would affect the demand for ITNs, as was found in this study. Thus, efforts must be made in designing and implementing ITNs sales so that people do not incur considerable time and travel costs.

Other factors that affected actual purchase decisions (though statistically significant in either one or two of the groups) also provided insights into the demand for ITNs. They also showed how the demand could be shifted to increase sales of ITNs. For example, from the bidding game group, it was shown that those with formal education are more likely to buy the nets. Thus, increasing the literacy rate would be a medium to long-term application of the finding to improve malaria control with ITNs as people could potentially become more aware about the scientific causes of the disease and how best to control it. However, in the short-term, the use of mass health education about malaria, as an intrinsic component of ITNs distribution strategies, could be pursued.

Since the bivariate and multivariate results show that lower socio-economic groups would be overlooked by ITNs distribution systems that are based on full cost-recovery, means of protecting them is needed. Thus, targeted subsidies or exemptions may have to be instituted if equity is to be maintained in ITNs sales. However, in the event that such options are not possible because of provider or government policies, then extended periods for instalment payments is an alternative that could be built into payment mechanisms. Farmers, who were less likely to purchase nets, also require financial protection from adverse effects of cash-and-carry or short instalment payment systems. Either of the above two techniques could be used. Farmers' co-operative societies could be initiated. These would allow for the pooling of funds over time by farmers that wish to buy nets.

7.5.2 Criterion validity

The confidence intervals generated around the phi coefficient point to the fact that the structured haggling and the bidding game methods could both be used with a good measure of confidence in valuing benefits and for market research. The confidence intervals of these two methods also overlapped, showing that they were quite similar. The findings also signal that WTP determination through contingent valuation would generate valid results.

Further comparison of stated and actual WTP using the positive predictive validity and overall correct prediction confirmed the pattern obtained with the phi coefficient. Positive predictive validity was high in the three groups. However, there was no statistical difference across the three question format groups ($p>0.05$).

Criterion validity of the bidding game and structured haggling technique were more related using positive predictive validity than with phi coefficient. This is because, only a sub-sample of respondents are considered in the test of positive predictive validity, while the phi coefficient uses the entire sample and thus gives a more robust indication of criterion validity.

Thus, using the phi coefficient as the basis for judging the performance of the question formats, the overlapping confidence intervals showed that the structured haggling and the bidding game were similar. Also, the confidence intervals of the bidding game and BWFU overlapped. The fact that the phi coefficient of the structured haggling method was highest, followed by the bidding game and lastly the binary with follow-up method could be used to rank the performance of the question formats. This is because, the higher the correlation, the more valid the test is for this particular criterion (Streiner and Norman, 1997). Therefore, it is concluded that the structured haggling had better criterion validity than the BWFU, but that the bidding game was similar to both question formats because of the overlapping confidence intervals.

The performance of the question formats in terms of criterion validity can be related to the tests of content validity presented in Chapter 6. Better content valid question formats will increase the probability of a stated WTP approximating the actual WTP. The three question formats were content-valid to varying extents, thus probably leading to the similar criterion validity. Hence, the positive predictive validity of the bidding and structured haggling question formats was high and that for the BWFU medium. It is possible that lower criterion validity would have been found if question formats like the open-ended and dichotomous choice methods had been applied in the study area.

The methods for eliciting the amount of money people are willing to pay for goods and services need to be relevant to the area in which they will be used, so that responses will be valid. It is recommended that future studies should establish and improve the content validity of question formats in areas and contexts where they will be used before administering questionnaires to respondents, in order to improve the validity of the findings.

It is also important to further develop strategies to improve the criterion validity of the question formats, together with ways to clearly demonstrate whether or not differences exist in their predictive abilities. The use of bigger sub-samples that state WTP amounts that reach the sale price of the nets, by using the median stated WTP as that price, is a possible strategy for improving the power of the tests. It was not possible to implement such a price regimen in this study because the National Malaria Control Programme that supplied the nets refused to lower the price.

One could also compare question formats with little or no content validity in the study area versus those with some content validity in that area to determine whether the choice of question format is important for the criterion validity of CVM. In the present study setting, this would have entailed comparing the structured haggling technique with the open-ended or dichotomous choice formats. This was not done in this thesis because the idea was to use question formats with good levels of content validity (also improving the content validity of BWFU), so that all the results generated would be valid to certain acceptable extent and thus, to be found credible by policy makers.

All in all, the findings give the confidence that the contingent valuation method is a valid and justifiable tool for determining peoples' valuation of goods and services. It was not clearly proven that one question format had more criterion validity than others, though the structured haggling technique had higher phi coefficient scores than the other two and was clearly better than the BWFU. Nevertheless, it is necessary to validate the findings in other areas where haggling is the common price-taking mechanism. It is equally important to determine how the validity of the structured haggling technique and other WTP question formats could be improved.

7.5.3 Divergences between stated and actual WTP

Having a divergent WTP should not be regarded as a worrisome phenomenon, if the reasons behind it could be explained. Changes constantly happen, as response to changing circumstances. Thus, the failure to do what one says s/he was going to do does not necessarily mean poor validity, but may be a justifiable response to changes in determinants of purchase of the ITNs by the households.

Thus, in the context of contingent valuation, not all divergent WTP may be due to bias but could be due genuine reasons. It then behoves the researchers to determine whether genuine reasons for divergences in WTP exist, and to characterise the factors that could cause them. In real markets, both buyers or consumers could decide to change the amounts they had agreed to pay for or sell goods and services depending on some factors that can lead to the change. However, "by reducing the time period between the elicitation of the WTP bids to one or two days, the possibility of changes in other potential determinants of WTP is avoided" (Whittington et al., 1992).

It was seen that in the three question formats, there were few variables that were significantly associated with increased, decreased or constant WTP. Some of the effects of the variables on the direction of the divergences in WTP accorded with theoretical expectations. For instance, respondents from a household with a recent malaria attack in the BWFU group increased their actual WTP because the recent experience would have impressed on them the need to acquire the nets to prevent malaria. Also, respondents from households with lower unplanned expenditures were more likely to increase their actual WTP, when it diverged from stated WTP in the BWFU group. This may imply that they had free money to spare on the ITNs, unlike other households that had higher expenditures on items they did not plan for in between the surveys.

The positive effect of the ownership of a radio on actual WTP in the bidding group could be either that it signalled a higher socio-economic class or because the people got some positive external information through the radio that led them to increase their actual WTP. However, a counter argument is that having a radio could lead to a decrease in actual WTP if people receive adverse information about ITNs from their radio. Nevertheless, the ownership of a radio was also statistically significantly related to divergent WTP in bivariate analysis in the bidding game group showing that possibly increasing the information about the good being valued through other mass media channels could be used to influence the outcome of WTP valuations.

Furthermore, as income increases (using food cost as the proxy) in the BWFU and structured haggling groups, the people who changed their minds were more likely to increase their WTP. This variable was also statistically significantly related to divergent WTP in bivariate analysis in the BWFU and structured haggling groups. Thus, the implication is that richer households were more likely to actually purchase the nets, regardless of what they stated in the survey.

It is not clear why people that stated higher WTP amounts in the first survey were more likely to decrease their actual WTP. The variable was also statistically significantly related to divergent WTP across the three groups. However, a possible explanation is that these people from the beginning were more sensitive to the need for the nets, but due to the influence of other factors, they changed their minds. However, this signals that in future studies, more divergences in WTP could be expected from those that state higher amounts. Nevertheless, more in-depth qualitative and quantitative studies would be needed to unravel why stating higher WTP leads to lower actual WTP, as this is a piece of evidence that challenges the validity of WTP elicited using CVM.

The analysis of the variables that were included in the survey conducted alongside the sales of the nets showed that there were other factors that could explain the divergences in WTP and possibly the behaviours of other explanatory variables. Giving respondents time to think about their stated WTP was a significant variable in the BWFU group and showed that it lead to increases in WTP in this study. However, it could possibly work in both directions. For some respondents, it may increase their WTP because they could have better considered their household income and expenditures or could have learnt more about the ITNs, and thus better appreciate its benefits. For other respondents, they could decide to lower their WTP if they considered the ITNs less worthy than they did in the first survey. Allocation decisions in decreasing WTP with given time to think may well be the result of a complex decision process involving the contributions of several household members to the household's monetary and non-monetary sources (Whittington et al., 1992).

The significant effect of the attitude of community leaders on actual WTP that was found in the bidding game's group, shows that these leaders could influence opinion and outcome of research and other programmes in their communities. One hypothesis would be that if the community leaders support a programme and convey this opinion to their community members, these ordinary members would also act positively to the programme and vice versa. Thus, in this study where the community leaders were involved through informal meetings and participated in the focus group discussions, could have made them to support the programme. This could lead to some of the households to increase their actual WTP.

The interpretation of effect of external information on divergent WTP is similar to that given for the possession of a radio and it had a very significant effect on divergences in WTP in the BWFU group. Thus, what respondents heard from other people or obtained from elsewhere had an important influence on actual WTP in that question format's group. The effect could be to increase actual WTP over stated WTP if the information is positive about ITNs as was found in this study. Thus, an awareness of the nature of information filtering to the respondents of a CVM study about the resource in question is necessary for better interpretation of the WTP values.

As the results presented in this chapter showed, there were genuine reasons people changed their minds, and thus not all WTP divergences could be said to be due to strategic or hypothetical bias. Therefore, "in order to better understand why hypothetical and real CV valuations may differ, it is important to analyse the motivation in individual answers to hypothetical questions and develop a theory of how individuals respond to hypothetical questions" (Blumenschein et al., 1998). These divergences would also inform discussions on how to interpret criterion validity estimates, especially in the comparison of different question formats.

7.5.4 Conclusion

In light of the findings about the causes of divergences in WTP, the criterion validity results regarding the performance of the different question formats need to be interpreted with caution. It was seen that the lowest proportion of respondents with divergences in WTP were found in the structured haggling group, while the greatest proportion was found in the BWFU that had the lowest criterion validity scores. However, the BWFU group had the most reasons out of the three question formats to explain why the divergences occurred. Therefore, lower criterion validity scores for BWFU would not necessarily imply that the WTP valuation was less valid, if the causes could be explained.

Thus, the BWFU could not confidently be categorised as the least criterion valid question format out of the three without further testing, as low criterion validity estimates do not mean that the problem is with the format, but with changes in determinants of demand. Also, there were no significant differences between the methods in positive predictive validity of the three question formats.

The differences in criterion validity and the divergences in WTP could have been demonstrated better in this study if a higher number of respondents were hypothetically willing to pay up to the sale price of the net. Selling at the median WTP or very close to the median is advocated for future studies in order to have large enough sub-samples that could be used to clearly determine criterion validity, especially when comparing different question formats.

Finally, studies to classify individual divergent responses into strategic and non-strategic divergent WTP are needed and a triangulation of qualitative and quantitative research methods used to achieve this. It would also be useful if future studies could further determine divergences due to hypothetical bias. These insights could then be used as the basis for determining the criterion validity of CVM and different question formats in different contexts.

CHAPTER 8: RELIABILITY

The chapter presents and discusses the reliability of the contingent valuation method (CVM), together with the relative reliability of WTP estimates from three question formats. The results further show whether better content validity of question formats translates into better reliability of CVM. The first section presents the results of the inter-rater reliability followed by the results of the test-retest reliability, together with the determinants of lower than expected test-retest reliability coefficients. The results of evaluation of factors that could cause a divergence between the stated WTP in the first survey and the retest are provided in the third section. This will show whether lower than expected test-retest reliability coefficients meant that CVM or different question formats were unreliable. The reliability coefficients of the socio-economic variables are also presented to show the extent to which other variables measured at the same time could vary and the influence of possible rater errors in the interpretation of reliability coefficients. The discussion centres on the level of confidence with which the results should be accepted, together with their generalisability.

8.1 Inter-rater reliability

This section presents, the results of the reliability of the elicited WTP amounts when the questioning were conducted by two different people on the same respondents within a one to three day interval. Descriptive statistics and reliability coefficients on socio-economic variables are first presented, to prove a perspective for understanding the WTP reliability coefficients. Then, the descriptive statistics and inter-rater reliability coefficients for the WTP estimates are then presented.

8.1.1 Socio-economic data of the respondents

A total of 109, 110 and 102 respondents were interviewed in the bidding, BWFU and structured haggling groups respectively by two interviewers (see Appendix 15). The majority of the respondents in the three question formats were married. Females formed a slightly higher proportion of the respondents in all question formats. Also, slightly higher proportions of the respondents had some formal education with the average number of years of formal education being 3, 4 and 5 in the bidding, BWFU and structured haggling groups respectively. The average number of household residents was 4 people in all the question formats. Most of the respondents' households possessed

a radio and bicycle, but minorities had a grinding machine, motorcycle and motorcar in the three question formats.

Table 8.1 shows the inter-rater correlation coefficients of the variables measured by the raters. All the coefficients were statistically significant at the 1% level. Gender was perfectly correlated between the pairs of interviewers in the three question formats. The age of the respondent, status in the household, whether the respondent had formal education and number of household residents were also highly correlated. However, while the number of years of schooling was highly correlated between raters in the bidding game (BG) and BWFU, at 0.81 and 0.85 respectively, it was poorly correlated in the structured haggling (SH) at 0.36. The inter-rater reliability correlation of the variables measuring the presence of assets was mixed, particularly for ownership of a car, grinding machine and radio, but there were no particularly discernible patterns across the question formats.

Table 8.1 Inter-rater reliability coefficients (demographic factors and assets)

	BG	BWFU	SH
Status	0.76	0.68	0.65
No. of residents	0.76	0.68	0.61
Gender	1.00	1.00	1.00
Age	0.75	0.87	0.91
Attended school	0.62	0.82	0.63
Years of schooling	0.81	0.85	0.36
Marital status	0.91	0.86	0.82
Own radio	0.55	0.45	0.72
Own bicycle	0.66	0.72	0.76
Own grinding machine	0.45	0.38	0.64
Own motorcycle	0.71	0.76	0.77
Own motorcar	0.68	0.30	0.68

All correlation coefficients were statistically significant at the 0.01 level (2-tailed).

8.1.2 Willingness to pay estimates

Table 8.2 shows that the respondents' decisions of whether or not to pay anything for ITNs were statistically significantly different in the BG (p=0.02). However, it was not statistically significantly different in the BWFU group (p=0.48) or in the SH (p=0.10). There was no statistically significantly difference in the mean WTP for own nets elicited by the pairs of interviewers in the three groups; BG (p=0.26), BWFU (p=0.71) and SH (chi-square 0.13; p=0.72). However, there were significant differences in the WTP for others in the BG (p=0.01) and BWFU (p=0.03), but not SH (p=0.18). There was also a statistically significant difference in the BG (p=0.01) and BWFU (p=0.06) for WTP for re-treatment although not so for the SH (p=0.18).

Table 8.2 Willingness to pay estimates for inter-rater reliability

	BG		BWFU		SH	
Variables	1st Interviewer n (%)	2nd Interviewer n (%)	1st Interviewer n (%)	2nd Interviewer n (%)	1st Interviewer n (%)	2nd Interviewer n (%)
Whether willing to pay						
1 = Yes	76 (69.7)	83 (76.1)	85 (77.3)	87 (79.1)	81 (78.6)	84 (82.4)
0 = No	33 (30.3)	26 (23.9)	25 (22.7)	23 (20.9)	22 (21.4)	19 (17.6)
Total	109 (100)	109 (100)	110 (100)	110 (100)	103 (100)	103 (100)
Chi-square (p value)	5.4 (0.02)		0.50 (0.48)		2.7 (0.10)	
WTP for own nets						
Mean	161.10	170.00	193.68	192.91	192.28	188.74
(S.D)	(145.52)	(146.36)	(139.94)	(138.51)	(166.91)	(153.11)
Median	200.00	200.00	200.00	200.00	155.00	180.00
Chi-square (p value)	1.29 (0.26)		0.14 (0.71)		0.13 (0.72)	
WTP for others						
Mean	93.59	57.43	92.59	71.68	89.40	120.06
(S.D)	(130.00)	(108.63)	(127.49)	(123.39)	(136.40)	(156.24)
Chi-square (p value)	7.11 (0.01)		4.80 (0.03)		1.81 (0.18)	
WTP for re-treatment						
Mean	41.33	46.42	46.19	46.63	48.15	52.69
(S.D)	(30.69)	(33.73)	(21.05)	(22.99)	(28.76)	(30.57)
Median	30.00	50.00	50.00	50.00	50.00	50.00
Chi-square (p value)	6.43 (0.01)		3.57 (0.06)		1.80 (0.18)	

The reliability coefficient on the decision whether or not to pay anything for ITNs was high[20] in all three groups ($p<0.01$) (Table 8.3). The table also shows that the highest coefficient of 0.82 was in the structured haggling group, followed by 0.80 in the bidding game and 0.79 in the BWFU group, but all the confidence intervals over-lap. There was a similar high degree of correlation in the WTP for the respondents' own nets at 0.77, 0.75 and 0.74 in the bidding game, BWFU and structured haggling groups respectively (with overlapping confidence intervals). However, the coefficients for other household members' nets were low to moderate; 0.58 (SH), 0.54 (BG) and 0.48 (BWFU) (with overlapping confidence intervals). The coefficients for re-treating nets were moderate to high, with the bidding game having the highest at 0.72, followed by the haggling group (0.59) and BWFU group (0.51) (with overlapping confidence intervals). All the inter-rater correlation coefficients were statistically significant ($p<0.01$).

Table 8.3 WTP Inter-rater reliability coefficients

	BG	BWFU	SH
Whether WTP (95% C.I.)	0.80 (0.72 - 0.81)	0.79 (0.72 - 0.81)	0.82 (0.74 - 0.87)
WTP for own net (95% C.I.)	0.77 (0.72 - 0.86)	0.75 (0.64 - 0.81)	0.74 (0.63 – 0.82)
WTP for others (95% C.I.)	0.54 (0.40 – 0.66)	0.48 (0.32 – 0.62)	0.58 (0.46 – 0.71)
WTP for re-treatment (95% C.I.)	0.72 (0.62 – 0.80)	0.51 (0.45 – 0.70)	0.59 (0.46 – 0.71)

All correlation coefficients were statistically significant at the 0.01 level (2-tailed).

[20] Munro (1997) classified the strengths of the correlation coefficient (r): 0 - 25 (little if any); 0.26 - 0.49 (low); 0.50 - 0.69 (moderate); 0.70 - 0.89 (high); and 0.90 - 1.00 (very high)

8.2 Test-retest reliability

The results of the test-retest reliability of WTP, together with factors causing a divergence between stated WTP in the first survey and the retest are presented in this section. The descriptive statistics and reliability of the demographic factors are first presented and then a similar presentation is made for the WTP estimates, before the possible causes the lower than expected test-retest coefficients are presented. To recapitulate, the retest was conducted one month after the end of the first survey.

8.2.1 Demographic data of the respondents

Descriptive statistics

A total of 146, 161 and 139 of the original respondents were re-interviewed in the bidding game, BWFU and structured haggling groups respectively by the same interviewers (see Appendix 16). Out of the people seen during the retest, a total of 12, 5 and 5 respondents in the bidding, BWFU and haggling groups respectively were either not available or refused to be interviewed. The majority of the respondents that were re-interviewed were household heads, females, married and with the average age been about 50 years in all elicitation method groups. Their major household asset possessions were radio and bicycle, while only a few households had a grinding machine, motorcycle or motorcar.

Table 8.4 shows that correlation coefficients for the same variables were all statistically significant at the 1% level. Gender was perfectly correlated in each group. All the other demographic variables showed test re-test reliabilities of more than 0.60, with the exception of the status of the respondent in the household (i.e. whether the head or a representative) with a coefficient of only 0.45 in the bidding game. The findings on the assets were again mixed depending on the question format, with the bidding game having the lowest coefficients and the BWFU have the largest number of higher coefficients.

Table 8.4 Test re-test reliability coefficients (demographic factors and assets)

	BG	BWFU	SH
Status	0.45	0.76	0.80
Sex	1.00	1.00	1.00
Age	0.69	0.76	0.91
Attended school	0.62	0.81	0.62
Marital status	0.61	0.72	0.83
Own radio	0.37	0.65	0.44
Own bicycle	0.50	0.63	0.65
Own grinding machine	0.37	0.61	0.44
Own motorcycle	0.64	0.71	0.69
Own motorcar	0.49	0.56	0.52

All correlation coefficients were statistically significant at the 0.01 level (2-tailed).

8.2.2 Willingness to pay estimates

The mean WTP amounts elicited in the first and re-test surveys were much closer for WTP for own nets in the bidding game and structured haggling groups, while the converse was true for the BWFU using WTP for other household members' nets (Table 8.5). The median WTP in both surveys were the same in the bidding game and structured haggling groups for WTP for ITNs, but different in the BWFU group.

The non-parametric comparison of means showed that the respondents' WTP for own nets were slightly significant in the bidding game ($p=0.059$) and structured haggling ($p=0.051$), but significantly different in the BWFU group ($p=0.01$) (Table 8.5). Conversely, the mean WTP for other peoples' nets in the two periods was statistically significant in the bidding game ($p=0.000$), but not in the BWFU ($p=0.60$) and structured haggling ($p=0.66$).

Table 8.5 Willingness to pay estimates for test-retest reliability

	BG		BWFU		SH	
Variables	1st Interview	2nd Interview	1st Interview	2nd Interview	1st Interview	2nd Interview
WTP for own net						
Mean	158.29	167.53	190.03	220.92	190.07	194.64
(S.D)	(109.09)	(133.86)	(107.47)	(135.36)	(126.97)	(145.72)
Median	200.00	200.00	200.00	250.00	200.00	200.00
Chi-square (p value)	3.57 (0.059)		18.68 (0.01)		3.81 (0.051)	
WTP for others						
Mean	78.70	47.53	101.89	101.75	108.35	73.13
(S.D)	(99.65)	(91.22)	(121.94)	(129.39)	(168.74)	(126.94)
Chi-square (p value)	17.89 (0.00)		0.28 (0.60)		0.19 (0.66)	

The correlation coefficients were 0.56, 0.51 and 0.41 for WTP for own nets in the structured haggling, bidding game and BWFU groups respectively (Table 8.6). The test-retest coefficients for other household members' nets were lower than for own nets, except for BWFU where it was higher for others net. All the test re-test reliability coefficients were statistically significant ($p<0.01$). The confidence intervals overlapped for WTP for own nets amongst the question formats, but WTP for others was higher in the BWFU group.

Table 8.6 WTP Test re-test reliability coefficients

	Bidding	BWFU	Haggling
WTP for own net	0.51	0.41	0.56
(95 C.I.)	(0.40 – 0.62)	(0.28 – 0.53)	(0.41 – 0.65)
WTP for others	0.34	0.52	0.33
(95% C.I.)	(0.20 – 0.47)	(0.48 – 0.68)	(0.20 – 0.47)

All correlation coefficients were statistically significant at the 0.01 level (2-tailed).

8.3 Factors causing a divergence between WTP at first survey and at retest

Only stated WTP for own nets was used in this analysis, but the results could be extrapolated to explain the test-retest coefficients of WTP for others. This is because, most of the changes that had occurred to cause some people to change the WTP for their own ITNs, could also cause them to change their WTP for other household members ITNs. The factors were categorised into the fixed socio-economic and demographic factors, and the special dynamic factors that could lead to change in WTP over time.

The first part of this section provides the descriptions of the respondents whose stated WTP for own nets diverged between the first survey and retest. This comparison is accompanied by analysis of whether the respondents whose WTP diverged and those where it did not diverge, were statistically different. In the third part, the result exploring the direction of the divergences is presented. In the last part, chi-square tests of the effect of some special factors that could explain divergences between stated and actual WTP are presented.

8.3.1 Causes of divergences in WTP in first survey and the retest

Table 8.7 shows that 54 (34.2%), 94 (56.6%) and 51 (35.4%) of the respondents in the BG, BWFU and SH respectively had a divergent stated WTP (changed their minds) between the first survey and the retest. The three question formats had similar characteristics in thirteen out of the twenty-one demographic variables measured for divergences in WTP (Appendix 17).

Few of the differences between those whose WTP diverged and those who did not were statistically significant. Average age was higher in the group that did not change their mind compared with those who did, but the difference was only statistically significant in the bidding game and BWFU groups. Also, occupational group 1 (unemployed, housewives and unskilled labourers) and stated WTP were statistically significant in the bidding game, while it was occupational group 4 (regular income earning occupations) in the BWFU group that significantly explained divergences in WTP. In the structured haggling group, the only variable that was significantly different between the WTP divergent groups was people occupational group 4.

For all question formats, farmers (occupational group 2), professionals and medium/big time business people (occupational group 5) were less likely to change their minds. Those that had some formal schooling, people that earn regular cash income, for example skilled workers, small time traders and those in formal wage-earning occupations were more likely to change their minds. Similarly, respondents from households that had more malaria attacks in between the first survey and the retest, those with higher stated WTP amounts and those that owned radios were more likely to change their minds. In the three question formats, representatives of households and the household heads were equally likely to either change their minds or maintain their original position.

The remaining eight variables were similar in some pairs of question formats. Females were more likely to change their minds in the bidding game group, while it was males in the structured haggling group. Sex had no effect on divergences of WTP in the BWFU group. Unemployed people and those that had higher food costs were more likely to change their minds in the bidding game and structured haggling group, while this was not the case for the BWFU group. Similarly, those with greater unplanned expenditures and those that spent more time getting to the sales spots were more likely to change their minds in the bidding game and BWFU groups, while the reverse was the case in the structured haggling group. The findings with respect to assets like ownership of bicycle, motorcycle and motorcar were mixed across the three question formats. Those with bicycles were less likely to change their minds in the bidding and BWFU groups, while the converse was true in the structured haggling group.

Table 8.7 Direction of the divergences of stated WTP in the first survey compared to the stated WTP in the retest (1 = increased WTP; 0 = decreased WTP)

	BG		BWFU		SH	
Variables	Increased WTP N=35 Mean (SD)	Decreased WTP N= 19 Mean (SD)	Increased WTP N=25 Mean (SD)	Decreased WTP N= 69 Mean (SD)	Increased WTP N=35 Mean (SD)	Decreased WTP N= 16 Mean (SD)
Status in the household	.66 (.48)	.56 (.51)	.81 (.39)*	.96 (.20)	.76 (.43)	.75 (.45)
Household residents	3.89 (1.95)	3.94 (1.95)	4.46 (2.30)	3.96 (3.28)	3.71 (1.82)	4.25 (1.65)
Sex	.26 (.44)*	.50 (.51)	.45 (.50)	.63 (.49)	.59 (.50)	.50 (.52)
Age	46.34(13.31)	44.06(13.14)	48.83 (13.9)	48.38(20.67)	49.06(18.01)	49.50(13.73)
Attended school	.51 (.51)	.67 (.49)	.65 (.48)*	.46 (.51)	.74 (.45)	.69 (.48)
Occupational group 1	.09 (.28)	.06 (.24)	.00 (.00)*	.04 (.20)	.06 (.24)	.00 (.00)
Occupational group 2	.57 (.50)	.56 (.51)	.58 (.50)	.75 (.44)	.38 (.49)	.63 (.50)
Occupational group 3	.23 (.43)	.39 (.50)	.25 (.43)**	.21 (.41)	.47 (.51)	.31 (.48)
Occupational group 4	.09 (.28)	.00 (.00)	.14 (.35)	.00 (.00)	.09 (.29)	.06 (.25)
Occupational group 5	.03 (.17)	.00 (.00)	.06 (.24)	.00 (.00)	.00 (.00)	.00 (.00)
Actual malaria incidence	.54 (.61)	.44 (.70)	1.0 (1.15)**	.50 (.66)	.56 (.75)	.25 (.58)
WTP self	170.00 (99.41)*	261.1(130.1)	183.3 (91.43)***	275.0 (110.3)	177.7 (92.90)	368.1 (213.1)
Total expenditure	4017.1 (8688.79)	3855.56 (4753.73)	5574.64 (15379.76)	6446.25 (14734.44)	4720.29 (12590.76)	2409.37 (3310.30)
Sales distance	14.14 (11.22)*	21.22 (14.59)	13.44 (12.47)	15.71 (15.01)	12.94 (11.03)	13.19 (7.38)
Food cost	886.0 (1201.4)	722.22 (292.1)	1116.0 (678.79)***	736.04 (505.69)	1005.44 (767.86)***	793.75 (522.20)
Radio	.83 (.38)	.72 (.46)	.88 (.32)	.75 (.44)	.79 (.41)	.63 (.50)
Bicycle	.57 (.50)	.72 (.46)	.61 (.49)	.50 (.51)	.62 (.49)	.63 (.50)
Grinding mach.	.09 (.28)	.06 (.24)	.16 (.37)	.12 (.34)	.15 (.36)	.13 (.340
Motorcycle	.17 (.38)	.06 (.24)	.22 (.42)*	.04 (.20)	.35 (.49)*	.13 (.34)
Car	.11 (.32)	.11 (.32)	.12 (.32)	.08 (.28)	.03 (.17)	.00 (.00)

Significance of differences between groups using the Kruskal-Wallis test: * <0.10, **<0.05, ***<0.01

8.3.2 The direction of the divergences caused by the socio-economic factors

There were different findings in the three groups concerning the effects of different variables on the direction of the divergences (Table 8.7). Only the statistically significant differences are presented. In the bidding game group, males were more likely to reduce their WTP when they change their minds. In that group too, people who stated higher WTP amounts, together with those that live at greater distances to the place where the sales were held, decreased their stated WTP in the retest.

In the BWFU group, household heads and the unemployed/housewives reduced their WTP when they changed their mind. Also, the respondents that had higher stated WTP were more likely to have reduced WTP in the retest. However, skilled labourers, people with formal education, with increased malaria attacks, with higher food costs and who owned a grinding machine were more likely to increase their WTP at retest. There was no statistically significant variable in the structured haggling group that explained the direction of the divergences.

8.3.3 Effect on divergent WTP by the special change factors

Table 8.8 shows the descriptive results of special variables added in the second survey to explain the reasons for divergences between WTP in the first survey and at retest, and how they were measured. The detailed definitions and theoretical expectations from these variables were provided in Chapter 5. The total numbers of respondents for all the variables are not uniform because of the way they were measured.

All the variables were statistically significantly different across the question formats. According to 46.4%, 35.1% and 28.2% of the respondents in the bidding game, BWFU and structured haggling groups respectively, changes in income had no effect on their WTP. However, the reported income changes had some effect for other respondents, and while the income changes decreased the WTP for some, the changes increased that of others.

The majority of respondents reported that the period of time in between the first survey and actual sale of nets had no effect on their WTP, with just few people declaring that it made them to either increase or decrease their WTP. The smallest effect on people changing their WTP, was that of the effect of someone's neighbours either buying or not buying ITNs.

Table 8.8 Additional variables measured only in the second survey to explain the causes of divergences in WTP (Special change factors)

Variables	Definition and measurement	BG	BWFU	SH	X^2 (p value)
Household Income effect	Effect of change in income				
	0 = Decreased WTP	45 (40.2)	64 (56.1)	45 (52.9)	6.87 (0.032)
	1 = Increased WTP	15 (13.4)	10 (8.7)	16 (18.8)	
	2 = No effect	52 (46.4)	40 (35.1)	24 (28.2)	
	Total	112 (100)	114 (100)	85 (100)	
Time to think effect	The effect time the respondents thought about their WTP				
	0 = Decreased WTP	18 (11.4)	23 (13.9)	14 (9.7)	25.06
	1 = Increased WTP	12 (7.6)	51 (30.7)	27 (18.8)	(0.000)
	2 = No effect	128 (81.0)	92 (55.4)	103 (71.5)	
	Total	158 (100)	166 (100)	144 (100)	
Multiple interviews	The effect of multiple interviews				
	0 = Decreased WTP	14 (8.9)	10 (6.0)	1 (0.7)	
	1 = Increased WTP	20 (12.7)	57 (34.3)	24 (16.7)	23.16
	2 = No effect	124 (78.5)	99 (59.6)	119 (82.6)	(0.000)
	Total	158 (100)	166 (100)	144 (100)	
Community leaders attitude	Community leaders effect				
	0 = Decreased WTP	1 (3.3)	4 (10.3)	4 (7.5)	
	1 = Increased WTP	13 (43.3)	11 (28.2)	32 (60.4)	8.57 (0.014)
	2 = No effect	16 (53.3)	24 (61.5)	17 (32.1)	
	Total	30 (100)	39 (100)	53 (100)	
Information	The effect of external information the respondents got about the ITNs				
	0 = Decreased WTP	6 (9.7)	2 (2.1)	1 (1.4)	12.35
	1 = Increased WTP	20(3.2)	49 (51.6)	48 (68.6)	(0.002)
	2 = No effect	36(58.1)	44 (46.3)	21 (30.0)	
	Total	62(100)	95 (100)	70 (100)	
Neighbours effect	The influence of neighbours buying or not buying ITNs				
	0 = Decreased WTP	2 (7.4)	1 (3.3)	5 (26.3)	8.31 (0.016)
	1 = Increased WTP	2 (7.4)	1 (3.3)	2 (10.5)	
	2 = No effect	23 (85.2)	28 (93.4)	12 (63.2)	
	Total	27 (100)	30 (100)	19 (100)	

As described in Chapter 5, chi-square analysis through a cross-tabulation of divergent WTP (without direction of divergences explored) and these special change factors were used to investigate whether there were statistical significant relationships. Also, further cross-tabulation of divergent WTP2 (1 = increase in WTP and 0 = decrease in WTP where there are divergences) with these factors showed the extent to which they explained the divergences in WTP.

Table 8.9 shows that the total number of respondents for the special change factors varied because the questions were designed differently. In the case of income, all the respondents were first asked whether there was any change in their household income since this could lead to a divergence in WTP. Then, for those whose income changed, they were then asked whether this affected their WTP, and the options were that it either had an effect or no effect. Thus the total number presented was that of the respondents who stated a change in household income. The same type of questioning was used to determine the effect of the attitude of community leaders, external information and neighbours influence. The effect of time to think and multiple interviews were not broken down but determined as a straight effect or no effect option.

The effect of subjecting the respondents to more than one interview (to determine inter-rater reliability and test-retest reliability) was a statistically significant factor explaining divergences in WTP across the three question formats' groups. The significance levels were $p<0.05$ for the BWFU and structured haggling groups and $p<0.10$ for the bidding game group. The changes in household income between the two surveys also significantly explained why divergences occurred in the bidding game and BWFU groups, while the time they had to think about the ITNs in between the surveys was statistically significant in the BWFU and structured haggling groups. Similarly, the external information the respondents received about the ITNs explained the divergences in the BWFU group ($p<0.01$).

Further testing of means and chi-square tests of these factors to reveal the directions of change revealed some interesting results. The one-month period they had to think about their WTP led to more of an increase than decrease in WTP in the three groups. However the result was statistically significant in the BWFU group (Chi-square = 3.83; $p=0.04$) but not in the bidding game and structured haggling groups. Income changes led to more of a decrease in the BWFU and structured haggling groups and the converse in the bidding game group. However, the differences were not statistically significant.

The fact that the respondents were interviewed more than once led to increases in the stated WTP in the retest over that elicited in the first survey in the BWFU (Chi-square = 12.19; p=0.000) and structured haggling groups (Chi-square = 2.23; p=0.12), but no directional change in the bidding game group. The attitude of the community leaders led to WTP increases in the bidding game and BWFU group but the converse was true in the structured haggling group. However, this change was only statistically significant in the bidding game group (Chi-square = 6.74; p=0.021). Finally, the external information the respondents received led to increases in WTP in the bidding game and BWFU group, but no change in the structured haggling group. This change was only statistically significant in the BWFU group (Chi-square = 14.50; p=0.000).

Table 8.9 Result of cross-tabulation of divergences in stated WTP with special change factors

Change Factor	BG	BWFU	SH
Income	No = 52 Yes = 60 Total = 112 (Chi-square = 4.12; p=0.03)	No = 40 Yes = 74 Total = 114 (Chi-square = 3.00; p=0.06)	No = 24 Yes = 60 Total = 84 (Chi-square = 0.864; p=0.25)
Time to think	No = 128 Yes = 30 Total = 158 (Chi-square = 0.01; p=0.55)	No = 92 Yes = 74 Total = 166 (Chi-square = 3.04; p=0.06)	No = 103 Yes = 41 Total = 144 (Chi-square = 4.48; p=0.03)
Multiple interviews	No = 124 Yes = 34 Total = 158 (Chi-square = 3.20; p=0.06)	No = 99 Yes = 67 Total = 166 (Chi-square = 4.24; p=0.03)	No = 119 Yes = 25 Total = 144 (Chi-square = 3.64; p=0.05)
Attitude of community leaders	No = 16 Yes = 14 Total = 30 (Chi-square = 0.48; p=0.37)	No = 24 Yes = 15 Total = 39 (Chi-square = 0.90; p=0.27)	No = 17 Yes = 36 Total = 53 (Chi-square = 0.050; p=0.53)
External information	No = 36 Yes = 26 Total = 62 (Chi-square = 0.244; p=0.41)	No = 44 Yes = 51 Total = 95 (Chi-square = 15.65; p=0.00)	No = 12 Yes = 24 Total = 36 (Chi-square = 1.928; p=0.38)
Neighbours effect	No = 23 Yes = 4 Total = 27 (Chi-square = 0.482; p=0.46)	No = 28 Yes = 2 Total = 30 (Chi-square = 0.652; p=0.58)	No = 12 Yes = 7 Total = 19 (Chi-square = 0.326; p=0.47)

8.4 Discussion

The discussion of the implications of the findings for asserting the relative inter-rater and test-retest reliability of the different question formats is presented. These results are also used to infer the reliability of the CVM in the study setting. This then leads to a discussion about the level of confidence with which the results could be accepted by first discussing the empirical causes of lower than expected test-retest reliability coefficients that has a bearing on the acceptability of the reliability coefficients. There is a discussion in the next section on why reliability may vary depending on how accurately the questions are asked, answered and recorded, and effect of eliciting WTP for own use and WTP for other household members from a respondent. The final part provides a conclusion on generalisability of the findings and how to better determine reliability of the CVM in future studies.

8.4.1 Which question format was more reliable and why: what standards to use?

The three question formats were reasonably reliable in the study setting, as most of the inter-rater and test-retest reliability coefficients showed at least moderate correlation coefficients. The strongest evidence was the inter-rater reliability using respondents' WTP for their own nets and the decision of whether or not to pay anything for ITNs, where all correlation coefficients were high. Thus, one could argue that the CVM as a whole was a reliable technique for eliciting WTP.

The reliability coefficients are comparable to, or higher than, those generated by previous studies in the health sector like O'Brien and Viramontes (1994), Flowers et al. (1997) and Thompson et al. (1984) as already presented in Chapter 4. This evidence showed that the CVM could be potentially used to generate reliability estimates of WTP even in populations with low literacy and income levels. The inter-rater reliability coefficients in this thesis might have even been higher if the period between the two interviews had not been as long as five days as happened in some instances.

It was not easy to classify the overall performance of the three question formats because of the varied results. However, if the criterion was based on the respondents' WTP for own nets (because respondents are clearer about the value of a good for their own consumption), then it could be argued that the bidding game and structured haggling methods were generally better than the BWFU method. The bidding game had the best inter-rater reliability and the structured haggling method generated the best test-retest

data. However, all the confidence intervals overlapped, so the conclusion is that none of the question formats was the best or worst. However, the BWFU method was clearly better than the other two question formats in test-retest reliability of WTP for others since their confidence intervals were outside that estimated for reliability coefficient.

The case for the reliability of the CVM or of the three question formats could be weakened by the test-retest coefficients that were lower than the inter-rater reliability coefficients, because that signals that the estimates were possibly biased. However, a major conceptual difficulty in establishing test-retest reliability is in determining how much time should lapse between the first and second administrations (Fink, 1993). If too much time elapses, external events might influence responses for the second administration; if too little time passes, the respondents may remember and simply repeat their answers from the first (Fink, 1993). However, there could be genuine reasons for the lower coefficients, which may attest to the reliability of the CVM and that need to be investigated before reliability coefficients are judged as whether signifying reliability or not.

As Whitehead and Hoban (1999) stated, a CV instrument is still reliable if it measures significantly different WTP values across time periods, with differences in the appropriate direction, if any of the factors affecting WTP have changed over time. Therefore, it was important in this study to unravel the possible causes of lower than expected test-retest reliability coefficients using the inter-rater reliability coefficients as the best case scenario. The knowledge of these causes would contribute to determining the comparative reliability of different question formats and of the CV itself in the study setting.

The study of inter-rater and test-retest reliability could both have suffered from time effects, leading to changes in WTP elicited at the second survey, as respondents in both cases were re-interviewed after some days (inter-rater) and weeks (test-retest). In fact Whittington et al. (1992) found that 16% - 20% of people who were given five to ten minutes to revise their WTP bids did so immediately and suggested that the short term reliability of CV studies through WTP could be questioned. However, it is most likely that the magnitude of change or the influence of factors that will make the respondents change their minds would usually be less in inter-rater than in test-retest reliability because of the increased time difference between administration of the two surveys.

8.4.2 What socio-economic and behavioural factors possibly caused lower test-retest reliability coefficients?

The test re-test coefficient may also not be as high as inter-rater reliability coefficient because respondents may have changed their mind as a result of participating in the first survey (Bowling, 1991). This has been termed 'reactivity' by Carmines and Zeller (1979) who defined it to mean a situation where the process of measuring a phenomenon induces a change in the phenomenon itself.

The fact that the re-test was accompanied by actual sale of nets could also have contributed in lowering the test-retest estimates in many ways. For instance, some people that really wanted to buy the nets but that previously stated a low WTP would readily pay the price of the net. However, they could have stuck to their earlier WTP if nets were not sold. Conversely, some people that were not interested in buying the nets could have just stated very low amounts so as to be left alone.

In an attempt to minimise the influence of net sales, respondents were informed in the WTP scenario that nets would be sold after one month of the first survey. Mitchell and Carson (1989) suggest that hypothetical nature of contingent valuation scenarios introduce a random error without creating any systematic distortion and therefore could probably increase rather than reduce reliability. However, because the second survey was not strictly a hypothetical survey, there could have been a systematic distortion of the respondents' WTP.

The investigation of causes of divergence of WTP at test and retest yielded many possible causes of the lower than expected test-retest reliability. In addition, an assessment of the test-retest reliability of socio-economic variables showed that they were lower than their inter-rater reliability values. These two sets of factors could thus be used to explain some of the reasons for the test-retest reliability coefficients computed.

It was seen that the lowest proportion of respondents with divergent WTP were found in the bidding game group, as about 33.6% had a different stated WTP in the retest from the WTP they stated in the first survey. Similarly, 56% and 35.4% of the respondents in the BWFU and structured haggling groups respectively also had stated WTP amounts in the retest that were divergent from the stated WTP amounts in the first survey. The proportions of divergent WTP for own ITNs especially in the BWFU where it exceeded 50% was quite high, but it was seen that it was also in the BWFU group that most factors explaining change in WTP were found. Thus, an arguably correct interpretation of the result would be that the question format was reliable since it could pick up the effect of factors that determine WTP.

In real markets, buyers or consumers can decide to change the amounts they had agreed to pay for or sell goods and services. However, "by reducing the time period between the elicitation of the WTP bids to one or two days, the possibility of changes in other potential determinants of WTP is avoided" (Whittington et al., 1992). Thus, in the context of contingent valuation, not all divergent WTP may be due to bias but could be due to genuine reasons. It then behoves the researchers to determine whether genuine reasons for divergences in WTP exist, and to characterise the factors that could cause them.

People who had some formal education were more likely to change their minds and thus have divergent WTP. This factor could be that these people may understand the issue better and so, either raised or decreased their stated WTP after contemplating more on the subject or digesting some more information. Similarly, increased malaria incidences led the victim households to more frequently change their minds. This is logical as people who suddenly had an attack in the intervening period between stated WTP in the first survey and the retest would be more likely to increase their WTP for the nets, in so far that they believe that ITNs can prevent malaria. However, there could be instances where the malaria episode would consume considerable household resources leading them to lower their WTP for ITNs.

It was also found that generally people that earned regular cash income were more likely to change their minds. This factor could be due to unstable availability of cash income, especially with regards to skilled workers and artisans. Even those on formal wage jobs may either be owed salaries for months, or maybe just get their owed salaries paid during the sale of nets. Thus, availability of cash income works both ways depending on whether or not the respondent had income or not during either during the determination of stated WTP, or during net sales.

Also, people in households incurring more food cost (potentially, richer households) in the BWFU group were less likely to change their minds. This could imply that richer households have more stable preferences, are more confident about their WTP amounts, or that the poor have many competing needs for money and so are more frequently re-allocating planned expenditures.

The effect of different occupational groups on divergent WTP when contrasted with those respondents whose WTP did not diverge, matched theoretical expectations. The latter group of people could be classified as belonging to the lower socio-economic groups, and in fact subsistence farmers were less likely to change their minds. One explanation is that from day one, people in lower socio-economic groups knew that their income situations were not going to change much within a month, and so stated WTP amounts within their known budget limits. In contrast, the regular money-earning people could have been speculating on their expected income within a month.

It is hard to explain why people with higher stated WTP amounts in the first survey were more likely to change their minds. However, a possible explanation is that these people from the beginning felt a greater need for the nets, but due to the influence of other factors, they changed their minds.

It was also seen that, the higher the unplanned expenditures in between the surveys, the higher the probabilities of divergent WTP occurring. This means, that as people make expenditures on items they did not plan for in between the surveys, then they would be most likely to change their minds. One would guess that these unexpected expenditures could possibly make them to reverse their WTP downwards, as many of them confirmed in open-ended responses on why they changed their minds.

Further determination of factors causing divergences in WTP through cross-tabulations of divergent WTP with some special change factors revealed more insights on causes of the phenomenon. Change in income, as signalled in the bivariate analysis, was a significant cause of divergent WTP in the bidding game and BWFU groups. It would also lead to an increase in stated WTP during the retest in the bidding game group, but lowered WTP in the structured haggling and BWFU group. However, these directional changes of WTP as a result of change in income were statistically insignificant in the three groups.

Giving respondents time to think about their stated WTP was a significant variable in the bidding game and structured haggling groups and could possibly work in both ways. While, some respondents may increase their WTP because they could have better considered their household income and expenditures or could have learnt more about the ITNs, others could decide to lower their WTP if they considered the ITNs less worthy than they did in the first survey.

The effect of multiple interviews on having divergent WTP could be increases in WTP during the retest over the WTP stated in the first survey. This was confirmed in the BWFU, where WTP increased due to the multiple interviews. Thus, the fact that another person had interviewed the respondent is expected to convey to them the seriousness of the matter, and may lead to an increase in WTP. Similarly, the significant effect of the attitude of community leaders on WTP elicited during the retest was found to be positive in the BWFU group, where some respondents said they decided to increase their WTP amounts because the community leaders were promoting the programme.

Dividing the direction of divergence into whether it led to an increase or decrease in WTP during the retest showed a myriad of causes in the bidding game and BWFU groups. However, some of the reasons for the directions of change that were found are not clear. For instance, it is not clear why males and those with higher stated WTP amounts in the first survey were more likely to lower their WTP in the retest when they change their minds. It was also not clear why the household heads, those that had some formal education, the unemployed/housewives and artisans lowered their WTP in the retest.

The time people had to think increased WTP in the three question formats, though the finding was statistically significant in the BWFU group alone. This is counter to an earlier study by Whittington et al (1992) in Nigeria, where they found that people who were given more time to think about a good under valuation lower their WTP. It is possible that the time made it possible to consult other household members or maybe reconsiders the ITNs and expected benefits or "that households used the time to reconsider their budget constraints and competing needs" (Whittington et al, 1992). Also, respondents that live at greater distances to the place where the sales were held decreased their WTP during the retest, because of the possible higher time and travel cost to be re-interviewed before been allowed to buy the net.

Thus there were genuine reasons people changed their minds, and thus not all lower than expected test-retest reliability coefficients imply that the CVM or the elicitation mechanisms used were not reliable. Although the test-retest reliability coefficient for WTP for own nets was lowest in BWFU, that method had a higher incidence of genuine reasons for divergence, followed by the bidding game where there were also genuine reasons for divergence. Test-retest reliability for WTP for own nets was best in the structured haggling group where there were fewest reasons for divergence.

Therefore, sub-optimal reliability may not mean that the respondents' preferences were unstable or that there is a problem with the question formats, but that there are legitimate changes in determinants of demand. The occurrence of these changes will increase in the length of time between test and retest. All in all, understanding the reliability coefficients and the message from them needs more studies. In other words, the "validity" of the reliability coefficients, needs to be established, and methods for studying the reasons for departures from the logically expected results means that studies need to be developed.

8.4.3 The effect of the reliability of socio-economic variables the acceptance of the WTP inter-rater and test-retest coefficients could be accepted as true.

The reliability coefficients for the demographic and asset variables and WTP were all higher for inter-rater than for test-retest reliability. However, if the test-retest coefficients had been high for the WTP variables but low for the socio-economic variables, some doubt would have been placed on the reliability of the WTP. This is because; the reliability of the socio-economic variables is expected to be high and stable and to act as a standard for knowing whether the interviewers properly conducted the valuations.

The findings that measures of socio-economic variables between interviewers or at different interviews by the same interviewer varied, may be indicating that the interviewers did not ask some of the questions properly or did not accurately record the answers given. It could also be that the respondents were tired of being interviewed and were just giving incorrect responses. However, these findings could also be showing that even demographic and asset measures can change with time.

Thus, the factors that led to lower than expected socio-economic and asset variables could be the same ones that led to lower than expected WTP reliability values. The low inter-rater and test-retest reliability of coefficients for some demographic factors and assets could cast doubt on the robustness of the reliability of the estimates of WTP themselves. This is because the apparent measurement errors in the other variables could have equally affected the reliability of the estimates of WTP. Nevertheless, the fact that the reliability coefficients of the WTP and socio-economic variables changed in the same direction arguably gives the confidence that the reliability coefficients are acceptable. The confidence intervals on the WTP reliability coefficients were also narrow, thereby increasing the confidence with which the result can be accepted.

The changes in both inter-rater and test-retest measures of the socio-economic variables signals the importance of establishing that the assets and other measures being used as proxy measures of socio-economic status are both inter-rater and test-retest reliable. It is also important to determine whether their direction of change is the same as WTP reliability coefficients, because if the reliability coefficients of these other variables and the WTP estimates are changing in the same direction, may be taken as confirmation of the reliability of the WTP estimates.

8.4.4 Does the respondents' valuation of WTP for own use or WTP for others affect the reliability of the estimates?

The nature of value attributed to a good has an impact on the reliability of estimates of WTP. The confidence intervals surrounding the reliability coefficients for respondents' WTP for their own nets were consistently higher than that for other peoples' nets, except in test-retest reliability in the BWFU group.

All the question formats had high levels of inter-rater reliability coefficients for WTP for personal nets. Also, in test-retest reliability, the bidding and haggling methods showed moderate levels of reliability for WTP for personal nets and the coefficients of WTP for other household members' nets were not as high as for personal nets in the bidding and haggling groups. However, the converse was true in this case for the BWFU method.

The inter-rater reliability coefficients for other household members' nets were not as high as for personal nets indicating possibly that respondents are either more knowledgeable about values for themselves than others, or that they thought less deeply about other people. Thus, they were less able to give similar answers when questioned a second time for other household members' nets.

The disparity in the reliability of WTP for personal use and that for other people was accentuated by the inter-rater reliability of WTP for net re-treatment where the personal use of the respondents was again involved. Higher levels of inter-rater reliability coefficients were found across the three groups than in the WTP for other household members' valuation. The data from the bidding group provides a very clear illustration of the higher coefficients in net re-treatment compared to WTP for others.

8.4.5 How generalisable are the results and how to improve next time?

This study provides evidence that the CVM could be used to generate reliable WTP estimates of the value attributed to a good for personal or household use in a mostly illiterate rural population in Nigeria. It is possible that such findings might be replicated in similar settings in other countries and in other parts of Nigeria. It is less clear that values attributed to consumption of a good by others is reliable and further work is

needed. This is particularly important in the case of interventions to reduce infectious diseases.

In order to improve the inter-rater reliability of the WTP estimates, better fieldwork strategies should be designed so that the maximum delay between the two interviews should not exceed two days. The field-workers also need to be trained thoroughly and checks instituted in the field to ensure that they administer the questions properly. The use of tape-recording of all the interviews could aid this quality control of the interviews. Finally, pilot inter-rater and intra-rater reliability estimation could be done before the actual survey. This will help correct all the errors due to the field-workers and ensure that the reliability estimates are only those of the respondents.

CHAPTER 9: DISCUSSION

The discussion is divided into eight sections, with the headings in the form of questions, so as to facilitate the provision of answers to the debates raised in the introductory chapter, and assess the extent to which the aims and objectives of the thesis have been achieved. Section 9.1 focuses on the validity and reliability of the question formats, which is then used to infer the validity and reliability of the CVM as a whole (Section 9.2). Following this, Section 9.3 discusses the implications of divergences in WTP for interpreting CVM estimates, and stating whether biases were observed in measurement; while Section 9.4 examines ways in which the process of eliciting WTP could be improved in future studies, based on the experiences gained in the thesis. The policy implications of the thesis for improving ITNs programmes and resource-allocation so that they are equitable are examined in Sections 9.5 and 9.6. Section 9.7 discusses the extent to which the CVM fully measured consumer surplus and what the thesis adds to the valuation of public goods. Lastly, Section 9.8 discusses the limitations of the thesis and highlights areas for improvement in future studies.

9.1 Which question format performed best?

It was difficult to select which question format performed best, as the results of the different tests for reliability and validity varied. Therefore the question shall be answered in three ways: to compare the results of absolute scores and statistical significance of tests conducted; to assess the impact of accounting for explained change over time in reinterpreting the results of criterion validity and test-retest reliability; and to consider whether one of the measurement tests (criterion validity) should take precedence over all other test results.

Table 9.1 displays the summary of performance of the question formats. The table is based on the absolute levels of the tests, as the differences among the question formats were statistically insignificant. The rank of 3, 2 and 1 were assigned to the best, medium and least performing question format on each test. However, in cases where there were ties in performance, the ranks were assigned so that the total for all tests would still be 6, thereby ensuring equal weighting of the tests. Adding up the ranks showed that the structured haggling technique (SH) had the highest score, followed by the bidding game (BG), while the binary-with follow-up technique (BWFU) had the least total score.

The ranking of content validity was based on the responses of the traders and the consumers, while that for construct validity for ITNs and their re-treatment were based on the number of statistically significant variables that had the correct hypothesised signs, from each question format. The confidence intervals of the phi coefficient for testing criterion validity between the bidding game and haggling technique overlapped; while those of the structured haggling technique and the BWFU did not. However, the confidence interval of the phi coefficient of the bidding game and BWFU overlapped. Therefore, the ranking conducted considering these ties accounted for the 2.5 assigned to the structured haggling technique. The bidding game should also have been assigned 2.5 because it was similar to the structured haggling technique; but because it overlapped with the BWFU, both were assigned 1.75 [(2.5 + 1) / 2)]. There were no statistically significant differences in inter-rater reliability and test-retest reliability for WTP own net, as all the confidence intervals overlapped. However, the BWFU produced a more statistically significant test-retest reliability for WTP for other peoples' nets than the other two question formats. The total ranks in Table 9.1 showed that the structured haggling technique was the best performing format, while the BWFU was the least performing question format.

Table 9.1 Summary of the performance of the question formats using the various validity and reliability tests

Test	BG	BWFU	SH
Content validity	1.5	1.5	3
Construct validity for ITNs	3	1	2
Construct validity for ITNs re-treatment	2	1	3
Criterion validity (own nets alone)	1.75	1.75	2.5
Inter-rater reliability for own net	2	2	2
Inter-rater reliability for other peoples' nets	2	2	2
Inter-rater reliability for ITNs re-treatment	2	2	2
Test-retest reliability for own nets	2	2	2
Test-retest reliability for other peoples' nets	1.5	3	1.5
Total score	17.75	16.25	20.00

The input of the various validity tests in Table 9.1 indicating the relative performance of the question formats produced mixed results. The structured haggling technique was the most content-valid. The bidding game was the most construct-valid of the three using estimates of WTP for ITNs, while the structured haggling technique was the most construct-valid using estimates of WTP for re-treatment. All three question formats were equally reliable for all the tests except for test-retest reliability of WTP for other people's nets where the BWFU was the best.

The point estimates and confidence interval estimates of phi correlation coefficient for WTP for respondents' own nets of the structured haggling technique ranged from medium to high, the bidding game from low to medium, while the BWFU was generally low (Munro, 1997). Although, the confidence intervals around the phi coefficients of the question formats formed the ranking of their performance in Table 9.1, the evidence on criterion validity was confounded by the findings of causes of divergences between stated and actual WTP across the question formats.

The BWFU with the lowest criterion validity had the highest genuine causes of the divergence, while the structured haggling technique had the lowest reasons for the divergence. In addition, there was no statistically significant difference in the positive predictive validity across the three question formats. Therefore, though the structured haggling had the highest levels of criterion validity, one could also argue that there were genuine reasons for the lower scores in the bidding game and BWFU groups that accorded with economic theory. Thus, it is inconclusive which question format had the best criterion validity.

However, the possibility of a similarity in criterion validity across the three question formats will not distort the overall performance of the formats based on the ranks in Table 9.1. This is because, even if all question formats were assigned an equal rank of 2, the structured haggling technique will still have the highest rank, while the BWFU the lowest total rank.

The overlapping of the confidence intervals of the three question formats for the three WTP inter-rater reliability estimates and test-retest reliability for respondents' own ITNs revealed that they were not easily distinguishable in test of reliability. However, the confidence intervals of the coefficients also overlapped for WTP for other household members' ITNs between the bidding game and structured haggling technique; but did not overlap with that of the BWFU, which was higher than those of the other two question formats.

There were, however, genuine reasons for the lowest test-retest reliability for own nets of the BWFU group compared to the other question formats, thereby questioning whether the result would have been the same if those causes of divergence between stated and actual WTP were held constant. However, a counter argument could be that the highest level of test-retest reliability for other household members' ITNs was in the BWFU group, thus questioning the extent to which the causes of divergences actually affected the performance of that question format.

Taking the absolute scores of the validity and reliability tests and causes of divergences into consideration, none of the question format was clearly the best in tests of criterion validity and test-retest reliability. Also, the proximity of the total scores in Table 9.1 between the question formats suggests that there were little differences in performance between them. Like all CVM studies, a possible limitation with this thesis is whether the scenario presented enough relevant information to the respondents, and whether the question formats were properly understood. The tests of content validity and pre-testing conducted before the survey suggested that that the scenario conveyed enough information and people understood the question formats. However, the difficulty that arises is that it is difficult to convey enough information in a short survey (Phillips et al., 1997).

The three groups of respondents used for the study were from the same villages; and their socio-economic and other characteristics revealed greater degrees of homogeneity rather than heterogeneity; and the study design was the same for the three groups. Therefore, it could be argued on one hand that the comparative reliability and validity of the question formats are robust, as they have no other explanation for the results apart from the effect of the question formats themselves. However, it could be also argued that since not all conditions for realism as presented in Chapter 1 were ensured, other factors could have been responsible for the findings. Also, it was not possible to offer the ITNs at the final prices elicited from each respondent in order to better determine criterion validity; as it would have caused problems if some people had paid more or less than others.

Despite the good performance of the structured haggling technique, it has its limitations. The principle limitation being that it is structured, and so does not allow for many haggles as different interviewers may have different haggling skills. Also, there are artificial boundaries set on the number of haggles in order to ensure uniformity of haggling and to reduce the interview effect on the final haggle. Furthermore, due to the numerous steps in the structured haggling technique, it required that the interviewers needed more time to master the steps. However, in order to improve its content validity and address its limitations, improvement in the structure of the structured haggling technique will form an essential part of future research.

Like most empirical studies, it was found that different question formats yielded similar criterion validity, and therefore the recommendations by NOAA (1993) for the preference of the dichotomous-choice format need to be reconsidered. Nevertheless, it could be argued that it is possible that most comparisons to date might not have used adequate sample sizes and appropriate data analytic techniques to properly illustrate the differences between diverse question formats. However, almost equal numbers of sales were recorded in the three question formats groups, and there were genuine reasons why WTP diverged in all the groups. Thus, as argued by Brown et al. (1995) in comparing close-ended and open-ended question formats, it may be possible that "the format could have mattered more for hypothetical than for actual payments".

The conclusions regarding the criterion validity of these three question formats in rural Nigeria are still not definite, as they could not be made from one study in a particular area, using one product. More studies in different areas, using ITNs and other goods are needed to validate these findings. This thesis also shows that some of the conclusions made from earlier studies could be erroneous due to weaknesses in data analysis, which thereby need to be standardised. Also, more content validity studies will be required to assist in modifying simpler methods like the BWFU and bidding game to accurately capture actual preferences.

Although it has been argued that improved content valid of the question formats should increase the reliability and validity of the CVM, further tests of criterion validity would be needed before conclusive evidences on the performance of different question formats in rural Nigeria could be made. Further testing will require larger sample sizes, and using the prices that most people have stated they would pay as the price to sell the nets, in tests of criterion validity. In so doing, perhaps, the differences in criterion validity would be drawn out. However, all things being equal, using content-valid question formats increases the realism of the WTP elicitation process, thereby decreasing the probability of the occurrence of hypothetical bias. Thus, one posits that the three question formats used in this thesis produced valid and reliable estimates of WTP, as they were all content valid in the study setting, although to varying levels.

However, the potential for biased estimates of WTP still exists even in the use of content-valid question formats. This stems from numerous factors including the hypothetical nature of contingent valuation; elements of the scenario; and researcher, interviewer and interviewee effects. The researcher could incorrectly design the scenario or use inappropriate survey and data analytic methods that would in turn lead to biases results. Interviewer effects refer to situations where the interviewers conduct the interviews poorly, whereas interviewee effects refer to incidences where the respondents would deliberately give false responses.

Finally, ways of having a better understanding of the reasons behind the values given by the respondents, together with understanding causes of divergences in WTP are needed in designing future studies, for eliciting more reliable and valid WTP for policy-making. The data analytic processes for comparing the reliability and validity of different question formats need to be standardised to increase the credibility of conclusions from different studies. Also, the use of different question formats within a study is an effective process for ensuring that the values of consumer surplus determined from consumers are reliable and valid. This is because, the various tests of reliable and validity could then be used to select the most accurate estimate as the value of consumer surplus.

9.2 Was the contingent valuation method both reliable and valid?

The CVM though developed in environmental economics for the valuation of environmental resources is steadily carving a niche for itself in the health care sector, where it is being increasingly used to measure benefits and understand the demand for goods and services. The estimates of WTP determined in this study could be used as a measure of benefit for ITNs and their re-treatment. Additionally, the findings could also be used to understand potential demand and optimal prices for the products. While the mean WTP is used as the measure of benefit, the median (as the price that at least 50% of the people are willing to pay) is preferred for price-setting and understanding demand (Weaver et al., 1996).

However, there remains the concern in literature about whether the estimates of WTP elicited through the CVM could be reliable and valid estimates in health care and in developing countries like Nigeria. The response to this concern was derived from the reliability and validity of the three CVM question formats used in this thesis. Therefore, the first part of this section deals with reliability of the CVM, while the second part is concerned with the validity of the CVM.

9.2.1 Reliability

The study showed that the contingent valuation method as a whole was a reliable technique for eliciting WTP for health care goods in Nigeria. This is supported by evidence presented in the inter-rater reliability determination using respondents' WTP for their own nets and the decision on WTP, where all the strengths of the correlation coefficients were high (Munro, 1997).

The case for stating that the CVM is reliable could be weakened by the lower test-retest coefficients. However, as was found in this thesis, there were genuine reasons for the lower than expected test-retest reliability in the valuations, using WTP for own net as a case study. Most of the causes of divergences acting in the way predicted by economic theory made some respondents to either increase or decrease their WTP in the retest. For instance, respondents whose households had malaria attacks in between the surveys increased their WTP. Thus, test-retest reliability coefficients that were lower than the inter-rater reliability coefficients arguable attested to the reliability of the CVM.

A low test-retest correlation may not indicate that the reliability of the test is low, but may, instead signify that the underlying theoretical concept has changed and "the longer the time interval between measurements, the more likely that the concept has changed". (Bowling, 1991). As argued by Whitehead and Hoban (1999), a CV instrument is reliable if it measures significantly different WTP values across time periods, with differences in the appropriate direction, if any of the factors affecting WTP have changed over time. Thus, "a naïve interpretation of test-retest correlation can drastically underestimate the degree of reliability in measurement over time by interpreting true changes as measurement instability" (Carmines and Zeller, 1979).

Further explanations for the lower test-retest reliability coefficients not explored in this thesis, include the possibility that some people that wanted to buy the nets but had previously stated a low WTP readily paid the price of the net when confronted by actual sales. In the absence of net sales, they could have stuck to their earlier WTP. Conversely, some people that were not interested in buying the nets could have simply stated very low amounts. In order to minimise the influence of net sales, the respondents were informed in the CVM scenario that nets would be sold after one month of the first survey, but it is still possible that the sales of ITNs introduced this bias.

Also, one can accept low reliability coefficients as translating to the reliability of the CVM, if it is found that the socio-economic and demographic variables were well measured, had acceptable levels of reliability and therefore the low coefficient was not due to gross measurement error of a survey setting. As was revealed in this thesis, the reliability coefficients of most of socio-economic and demographic variables were moderate to high, which thereby suggests that the CVM reliability coefficients were credible.

Overall, this thesis provides good evidence on the reliability of the CVM in the study setting. For instance, the inter-rater reliability coefficients for the decision on WTP and for WTP for own nets were more than 0.70 in the three question formats. The reliability of the CVM found in this thesis also appears strong relative to other studies in health care (Thompson et al., 1984; and O'Brien and Viramontes, 1984). However, those studies have been less elegantly designed and used small sample sizes, e.g. 20 people (O'Brien and Viramontes, 1994), or unrepresentative samples of the general population (Thompson et al., 1984).

Establishing reliability in CVM should be one of the priorities of CVM research. It could be argued that the most critical issue in the reliability of CVM is whether the same individual will elicit similar WTP values in repeated attempts. The next important issue is the extent to which two interviewers will elicit the same WTP estimate. This situation is especially important in surveys where some interviewers may not complete the interviews due to various reasons and the researchers do not want to lose the trend already established. However, it is also important for studies to gain an understanding of how reliable their estimates were within one survey, by determining inter-rater reliability.

9.2.2 Validity

The thesis shows that, overall, the CVM was valid in the study setting, considering the results of the content, construct and criterion validity tests. Criterion validity was the ultimate criterion for deciding on the validity of the CVM. Construct validity is usually investigated whenever no criterion or universe of content is accepted as entirely adequate to define the quality to be measured (Cronbach and Meehl, 1955).

However, several arguments could be advanced against the blanket acceptance of the validity of the CVM in this thesis. First, it could be argued that since the respondents were not actually faced by a purchase decision when determining stated WTP, they would have not stated their maximum WTP, and thus, the full value of the good was not determined. It is also doubtful whether the scenario completely conveyed the benefits of the nets to the respondents, in order for them to make a fully informed decision. These arguments are supported by the low value of phi coefficient in the BWFU and bidding game, together with considerable levels of divergences between stated and actual WTP. On the other hand, factors were identified which may be genuine causes of the divergences.

Further arguments against accepting that the CVM was valid could stem from the statistical insignificance of many of the variables that were hypothesised to explain WTP, or more strongly, the fact that some of the statistically significant variables had signs which differed from those hypothesised. Also, the use of different econometric models yielded different conclusions on the theoretical constructs of some of the variables. However, there may be logical reasons for the findings. One reason could be that the theoretical expectations were wrong. For instance, while it was hypothesised that married people should be more willing to pay than single people, the opposite was found in some of the econometric models. However, the set hypotheses could have been wrong for some of the variables. For instance, it is possible that single people are really more willing to pay for the nets since they are not worried about sharing their limited budget with a spouse. Researchers and rater errors in designing, administering and recording of responses could also lead to relationships, which were inconsistent with theoretical expectations.

Nevertheless, the arguments in favour of the validity of the CVM in terms of content validity was that both traders and consumers confirmed that the CVM could be used to elicit consumers' true valuation of goods and services. However, the extent to which the true values will be determined depend on the information given to the consumers before asking their WTP amounts, together with the extent to which the question format used mimics the usual price-taking behaviour in the markets.

A deeper understanding of content validity and the groups of people that could be used to test and improve the content validity of both the scenario and different question formats are needed. It may be necessary in future studies to present different versions of the scenario and question format (that have face validity to the researcher) to the people for them to assess and improve the study tools. The researchers would then have a wide array of tools from which to choose, or even combine elements from different versions of the same tool in order to have a better content-valid tool. Pre-testing could then be used to finalise the tools to ensure optimal content validity.

In terms of construct validity, the arguments for the validity of the CVM include the fact that the elicited estimates of WTP that were consistent with relationships predicted by demand theory. This was in line with the findings of most studies that have determined the construct validity of the technique. It was consistently found that WTP was significantly and positively associated with income (using food cost as the proxy). WTP also correlated with many other variables in manners suggested by demand theory and the hypotheses set in the study. Generally, the lower socio-economic groups were less likely to be willing to pay for the nets. Thus, "WTP is inevitably associated with ability to pay and such an association is good because it means that WTP is behaving in a way that would be expected a priori; thus confirming the internal validity of the technique" (Donaldson et al., 1997a).

The use of diagnostic testing to confirm that the appropriateness of the econometric models to affirm construct validity in this thesis, contrasts with most other studies in the health sector that were reviewed in Chapter 4. The majority of the reviewed studies attested to the construct validity of the CVM but used models without regression diagnostic tests. Therefore, one could not confidently accept the theoretical validity of the CVM in health-care as those models could have been spurious.

Nevertheless, due to some inconsistencies found across question formats and econometric models, improving results of construct validity will require more effective ways to: determine the hypothetical constructs of the variables; measure the variables; administer the questionnaires; and train and supervise the field workers. Also, asserting the construct validity of WTP may require the diagnostic testing of different econometric models.

The potential respondents for WTP valuation could be used through qualitative and/or quantitative studies to determine the theoretical expectations of certain variables, as a measure of construct validity. This is due to the variation of socio-economic and cultural characteristics in different settings. Thus, basing theoretical expectations of construct validity on the findings of previous studies conducted in different settings may lead to inaccurate conclusions on the construct validity of the CVM.

The CVM was also criterion-valid judging from the absolute values of positive predictive value and phi correlation coefficient in the three question formats; although the case for criterion validity of the CVM was arguably stronger in the structured haggling technique and bidding game than in the BWFU method. However, as further analysis of causes of divergences between stated and actual WTP illustrated, the case of low levels of criterion validity tests may not necessarily translate to lack of criterion validity (to be discussed in a later section of this chapter).

The data sets from earlier studies that compared stated and actual WTP (Loomis et al., (1997); Frykblom (1997); Johannesson et al., (1997, 1998, 1999); Blumenschein et al., (1998); Foster et al., (1997)) were not collected in ways that could be analysed using the phi coefficient. Those studies were also severely limited, especially regarding the good they used. Most of the studies only had one item of the good which was "sold" to the highest bidder. In addition, the studies did not examine goods that people really wanted to buy on the large scale, which was undertaken in this study.

The differences in settings used by studies in the literature contributed to differences in conclusions regarding criterion validity of the CVM. For instance, Blumenschein et al. (1998) found that the proportion of subjects who were hypothetically willing to pay for sunglasses significantly exceeded the number that actually purchased the sunglasses. Conversely, this thesis found that actual WTP exceeded stated WTP as found in an earlier study (Onwujekwe et al., 2001).

Thus, this thesis shows that there could be erroneous conclusions of the validity of the CVM if the processes and data analytic methods for content, construct and criterion validity are not thoroughly examined. Also, one could never confidently assert that CVM was criterion-valid without examining whether there were causes for divergences between actual and stated WTP. It is only when there are no causes of the divergences that the criterion validity estimates could be accepted, as the real test of validity of the CVM.

9.3 Do divergences in stated versus retest/actual WTP indicate the existence of bias in valuation?

The existence of lower than expected criterion validity or test-retest estimates does not necessarily mean that the CVM is less reliable or valid. This is because genuine causes of the divergences that were found in this study, actually support the reliability and validity of the technique. The CVM was in fact relatively sensitive in picking up these changes.

Unlike the conclusions of Blumenschein et al., (1998) that compared the divergences in WTP, and argued that the divergences provided evidence of hypothetical bias in CVM this study evaluated rather than assumed the reasons for divergence. The counter argument that comes from this thesis is that divergences in WTP do not necessarily mean the presence of bias.

The few studies that compared stated and actual WTP, or compared levels of stated WTP with another stated WTP at point of purchase (test-retest reliability), reported that there were divergences in many of the cases, thereby affirming the reservations of people regarding whether both WTP measures would converge. According to one school of thought, it is the very hypothetical nature of the contingent valuation context that is the cause of the divergence (Bateman and Turner, 1993).

Supporting the argument that we cannot conclude on the existence of hypothetical bias without determining why stated and actual WTP diverged was the findings of a study by Whittington et al. (1992) in Nigeria. Whittington et al. de-briefed their respondents and found that they all had very logical reasons for changing their minds and concluded that there was no evidence of strategic bias in their study. Similarly, Whitehead and Hoban (1999) in determining the test-retest reliability of WTP found that there were socio-economic factors that genuinely accounted for the divergence between stated and actual WTP.

Therefore, one can only be certain of the existence of hypothetical or strategic biases, by establishing that there were no logical reasons for people to provide different stated and actual WTP for the same good. This could only be done through investigating the causes of the divergence, using qualitative and/or quantitative research methods. The variables that will be responsive in detecting the causes of the change in the WTP need to be identified and included in the surveys. However, if people deliberately misrepresent their preferences for various reasons, it may be more difficult to ascertain the causes of divergence in WTP.

This type of study has not been conducted before, as such it "should be seen as hypothesis-generating rather than hypothesis-testing" (Olsen and Donaldson, 1998). The special variables included in this study to detect change helped to generate hypotheses on their influence on divergence of WTP. In addition, the testing of means helped generate hypotheses on the influence of several factors on the direction of divergence. This is because some factors could lead to an increase in actual WTP as compared to the stated WTP, while others will cause the opposite effect.

The implication for the CVM in investigating the causes of divergences in WTP is that unless we explicitly understand these factors, we cannot conclude that the failure of the two measures to converge is a sign of hypothetical bias or even strategic bias. This is attributed to the fact that in real life purchase decisions, people genuinely change their minds depending on the influences of external and demand-determining factors. The factors causing divergence in WTP and the direction of the divergence will vary from area to area, nature of the good used, and the context of the study. Nevertheless, further studies are necessary to generate more hypotheses on causes of divergence in WTP, while others could be replicated to test some of the hypotheses generated in this study.

9.4 What approaches could improve future CVM studies?

The reliability and validity of the CVM could be improved with increased familiarity with the good or service under valuation. Also, it is imperative to present a good that the respondents would ordinarily have wanted to consume in the scenario, so as to make the exercise more realistic, thereby encouraging people to reveal their true preferences. The ITNs used in this study could have helped to increase the validity and reliability of WTP estimates although this influence was not investigated in the thesis.

One thus disagrees with Kealey et al (1990) who suggested that the type of good used might have little influence on the reliability and predictive validity of the CV method, but that aspects of the method itself (e.g. explicit payment obligation) may be more important for high intention-behaviour consistency. However, Krupnick and Cropper (1992) support the argument that the nature of the resource is significant for improved reliability and validity, by showing that respondents with first-hand knowledge of a chronic lung disease were more willing to pay for the control of the disease. This was likely because the intervention represented a service that the respondents would normally want to consume.

As has been argued in this thesis, the use of context-specific question formats will decrease the divergence between stated and actual WTP; although it was not conclusively proven in this thesis. The discovery of what methods are context-specific requires the researchers to collect preliminary information on whether the question format resembles the price-taking mechanism and the extent that the format could be used to elicit intention-consistent values in the study setting. This information could be gathered from the consumers and suppliers of similar goods and services. What may appear face-valid to the researcher may not be so to the potential respondents and actual practitioners.

Unlike the design of the structured haggling technique, the content validity of many novel question formats was not explicitly determined, but rather relied on face validity. These efforts at designing new question formats include Davis (1963) who pioneered the design of question formats by developing the bidding game. Mitchell and Carson (1981; 1984) designed the payment card approach because the bidding game was prone to starting-point bias. Alberini, (1995) and Dalmau-Matarrodona (2001) found that the dichotomous choice method improves if a follow-up question is added.

In this study, the design of both the binary-with follow-up method and the structured haggling technique were modified by both the traders who know the practicality of different question formats (albeit informally) and the consumers on whom the question formats would be used. The traders and potential consumers also helped to modify the scenario. Thus, the versions of the question formats used in the surveys were from the collective input of the researchers, traders and consumers. However, caution needs to be exercised in the use of the input by the traders and consumers in case they introduce some biases in the scenario. For instance, the traders may want the side-effects of a resource like ITNs to be watered-down, so that the consumers may not be discouraged from buying. Conversely, the consumers may want a payment method that favours them, but that would lead to defaults in payments if a test of criterion validity is conducted.

While many studies considered it necessary to ensure that the scenarios they used were content-valid, none of them prior to the present one had explicitly determined the content validity for question formats. Not even the NOAA (1993) panel ensured that applying how people voted in referendum would be content-valid for eliciting the non-use value of environmental resources. An earlier study before the NOAA panel argued that "the CVM question is different from the normal price-taking behaviour where consumers react to posted prices; and is also different from public goods decisions voters make in a referendum" (Hoehn and Randall, 1987).

Hence, there are still controversies, surrounding the NOAA's recommended dichotomous-choice approach. Some authors have posited that it is less statistically efficient than the traditional open-ended or payment card approaches (Hanemann and Kanninen, 1996). Ryan et al. (1999) found that the dichotomous choice approach generated higher WTP estimates than the payment card approach. Smith (2000) provided other problems with the dichotomous-choice method that include 'yea saying'. The general conclusion by those authors was that due to the fallibility of the dichotomous choice format, economists should not necessarily adopt the recommendations of the NOAA guidelines regarding the use of that question format. They should instead continue to research the optimal way to conduct CVM studies.

The implication for ensuring that the research tools are content valid on eliciting more valid and reliable WTP is that the subjective views of researchers should not be the dominant input used for the final design of elicitation methods and scenarios used in WTP surveys. For improved reliability and validity, such designs should come from the consumers and possibly actual users (such as health care providers or traders) through qualitative and quantitative studies, with the researchers only providing the basic framework and modifying the instrument according to the input from consumers and suppliers. Pre-testing of the scenario is needed if the results are to be credible. Where the contingent scenario lacks credibility, respondents may be induced to strategically over-bid the belief that they will never be required to make the stated donations (Hoehn and Randall, 1987). They could also under-bid so that they do not have to pay for something.

In addition to the importance of determining content validity, it is also necessary to ensure that models used for assuring construct validity of elicited WTP are robust and this could be achieved through diagnostic testing of the models. The importance of using diagnostic tests to choose the appropriate econometric models was amplified in this study where the Tobit regression using the data generated from the structured haggling technique was the best according to the criteria set[21]. However, the Ramsey RESET test of functional form showed that it was not the correct model. Thus, without regression diagnostic tests, the Tobit model would have been chosen and the structured haggling technique adjudged to have been most construct-valid for WTP for ITNs.

The conclusions of many studies that have determined construct validity, without the conduct of appropriate diagnostic tests, are therefore questionable. Most of those studies simply relied on goodness of fit measures to assess the robustness of their regression models. This therefore raises the question of whether the construct validity of the CVM has been sufficiently proved.

[21] This was judged using the goodness of fit measures and theoretical constructs of the variables.

In order for econometric models and their conclusions to be acceptable as evidence of the construct validity of the elicited estimates of WTP, a minimum set of diagnostic tests must be performed. The minimum level should include test of functional form using the Ramsey RESET procedure and scatter plots for heteroscedascticity. These two tests are fairly simple to implement. The presence of heteroscedascticity could be more easily dealt with robust standard errors, transformations of the variables or the use of weighted least squares. Evidence of the wrong functional form means that the researchers should search for alternative estimators.

Modelling estimated WTP using different regression models is an excellent search process for determining the best models to use in interpreting the parameters. The robustness of the different parameters across different models strengthens the interpretations and conclusions on the effect of the variables on WTP. It should be noted that no single model is perfect for all data sets. Nevertheless, the Heckman's selection model estimated by maximum likelihood was consistently the best performing model.

The ultimate criterion for selecting the best econometric models rests principally on their robustness to regression diagnostics, and not solely be based on its ease of estimation (for example, the Tobit was easier to estimate than the Heckman's selection model, but fit the data less well). However, regression diagnostics also need to be undertaken in cases where single models are used; as econometric modelling without diagnostic testing is like a doctor concluding the diagnosis on a patient based on mere clinical examination, without supportive laboratory and other investigations.

However, the problems with the Heckman's model was the lack of clear guidelines in the literature on what variables should be included in the first part, so as to determine the cases that are selected for the second equation. It was suggested that the identifying variables in the first stage should be those related to the probability of a positive WTP, but not the level of WTP if positive, though the use of such variables are not mandatory (STATA, 1985-99). The method used in this study depended on the a priori logistic and OLS regression result of the decision to state a positive WTP and levels of WTP served to be an effective method for selecting these variables.

Characterising the zero WTP values to know the nature of those less likely to pay, together with the factors that explain the decision to pay, need to be determined by WTP studies. Thus, even in situations where continuous WTP values are elicited, the data should be initially converted to a binary distribution and econometric and/or non-parametric methods used to determine the predominant nature of the respondents in each group. This could make data analysis and interpretations more effective in providing insights on the effect of different socio-economic and other characteristics of the respondents, if the elicited WTP is used for actual resource allocation decisions.

This thesis is the first to use the phi coefficient in determining criterion validity of CVM. The phi coefficient offers a good test of criterion validity, as it relies less on researcher judgement and matches the criterion to all the respondents. As such, it could be adopted by other studies for easy interpretation and comparability of criterion validity. As shown in the review of empirical literature, most studies have principally relied on comparisons of mean or median of stated and actual WTP and others in addition compared regression models from both groups.

Future studies would be improved if they evaluate criterion validity results in relation to change factors, before providing strong assertions on the criterion validity of the CVM or of different question formats. In contrast to the arguments provided by Diamond and Hausmann (1994), lower than expected criterion validity estimates do not mean that the CVM is invalid, if their genuine causes can be identified. It was observed that while the BWFU had the lowest criterion validity values, it also had the highest level of reasons (change factors), which might explain why stated and actual WTP diverged.

Improving CVM studies/tests of reliability requires that the methodology for data collection and analysis for tests of reliability be re-examined, and should be in line with conventional practice in measurement theory from where the concept was initially borrowed. To assess the reliability of the CV method, replications should be conducted on the same group of people (Johannesson et al., 1993b). Improving the credibility of WTP reliability coefficients in health care requires studies to use adequate sample sizes. However, comparisons of means between groups using non-parametric or parametric tests could be adjuncts to the coefficients estimated. In cases where test-retest reliability is determined, it is important to empirically determine the causes of divergences between WTP in first survey and at retest in order to gain a full understanding of what

the coefficients really mean in terms of reliability of the CVM. Also, it useful to provide controls for the CVM reliability coefficients using the reliability coefficients of socio-economic variables, in order to aid interpreting the CVM reliability.

Different question formats were shown to generate different distributions of WTP, because they affect the respondents differently. This accounts for the need for different econometric models to model them. The mechanisms leading to the generation of different WTP by the different question formats is not clear and requires further investigation. However, it was observed that the bidding and structured haggling formats whose iterations were more related had similar distributions, when compared the binary with follow-up group with a different iteration.

Presently, the variety in methods of data analysis for all types of reliability and validity make comparisons between studies difficult. For example, one could not compare the models of theoretical validity because they may not have conducted regression diagnostics. Many studies that claimed to have determined reliability used analytic methods (like using the general fit of regression models) which were not measures of reliability. Varied data analyses were also found in the tests of criterion validity. A standardisation of analytic methods and compulsory tests needs to be developed in CVM for proper assessment of reliability and validity of the technique, and for easy comparability of results from different settings.

9.5 What are the implications of the findings for ITNs markets?

The divergences between stated and actual WTP should be anticipated in ITNs programmes and factors that will cause a positive divergence of WTP (i.e. make people to be actually more willing to pay), should be encouraged and amplified for optimal coverage. Conversely, the factors that could cause a negative divergence should be mitigated. Managers of ITNs programmes need to be aware that divergences in WTP could be in either direction and these changes will be greater the longer the interval between the survey to determine WTP and the actual sale of nets.

The improvement of coverage of ITNs programmes requires both demand and supply side interventions by the government and donors. The demand-side interventions may include means to increase the financial access of the people, while supply side interventions could include subsidies to the producers of nets, so that they would produce at lower cost and therefore sell to the consumers at reduced prices. Also, the duty on importation on nets and insecticides could be lowered or abolished so that the finished products would be brought into the country at reduced cost. The decision on whether to intervene on the demand or supply or a mixture of strategies, would require future cost-effectiveness analysis. However, since, there are no established net or insecticide producers or reliable importers in Nigeria, the feasible intervention at present would be demand-side interventions.

Financial and physical access to ITNs programmes need to be enhanced because the major factors that could limit the acquisition of nets as seen from the comparison of those who did and did not buy was distance to the sale's point and decreased levels of ability to pay. Econometric modelling of determinants of actual WTP confirmed that distance was a major limiting factor to net acquisition. People that belonged to households with greater malaria attacks were inclined to buy more nets. Therefore, the government or donors will have to design methods to ensure that ITNs schemes are based on providing equitable access, thereby reaching the poor and those at greatest risk of malaria, specifically children, pregnant women and immuno-compromised people.

Since people that lived further away from sales points were less likely to actually buy the nets, the obvious solution is to bring the nets closer to the people, but the problem is how to effectively achieve this. Improving physical access requires the use of sales mechanisms that are easy to implement and inexpensive (so as to not increase the cost of the nets). An option is to sell the nets through the primary health care (PHC) centres, but the main problem is that many communities do not have these centres. In fact, the PHC system is itself also struggling with ensuring physical access to its facilities.

The factors that determine the actual purchase of ITNs also have equity implications for ITNs programmes that are implemented on a user fee basis. Improvement of financial access will require that the pricing policy will take into account the fact that most people cannot afford to or were not willing to pay the official price for the nets. Selling nets at the official unit price of 450 Naira would lead to a very low level of coverage, so the pricing for the ITNs has to be appropriate for the target population. The median stated WTP amount could be used as a guide for appropriate pricing for the nets.

One way to improve financial access is by empowering the people through micro-credit and other income-generating schemes, so that they will have the money to buy the nets, amongst other basic household necessities at the official rates. An alternative will be for the government to supply the nets at subsidised rates to the people, but with the nets still being slightly more expensive than untreated nets in the market, to reduce the likelihood that nets are diverted to the open market. The communities could also ensure intra-communal altruism by the richer households contributing money to buy the nets for the poorest households.

Financing mechanisms that will ensure that those at increased risk of malaria are assisted to buy the nets should be designed and implemented. The people that are at higher risk of getting malaria could be found in all malaria endemic levels, and may signal that identification and "smart" (excluding all ineligibles) targeting of these people for ITNs sales could be a good strategy to increase coverage of ITNs. This means that it would be easier to access these people at the service points where they get malaria treatment or the management of other health conditions, and introduce/sell the nets to them there.

Armstrong Schellenberg et al. (1999) introduced a voucher system for pregnant women and infants to improve net coverage in the high-risk groups. Also, alternative payment strategies that will assist those with seasonal fluctuation in income, such as the subsistence farmers who comprise the largest occupational group in Nigeria, should be designed and implemented. Cham et al. (1997), Snow et al. (1999) and Makemba et al. (1995), looked at the impact of user fees on ITNs programme based on single payment mechanism, but did not investigate the effect of using alternative payment approaches, on increasing net coverage.

No particular distribution or marketing system is currently preferred in ITNs programmes. A number of studies have reviewed ITNs and insecticide distribution systems (Lines, 1996; Armstrong-Schellenberg et al., 1999 and 2001; Fraser-Hunt and Lyimo, 1998; Fielden 1996). Lines argued that the approach chosen must depend primarily on local circumstances as existing experience gives no clear general reason to prefer one option over another, in particular settings.

ITNs have been distributed using community-based methods in Africa (Makemba et al., 1995), and in South America (Kroeger et al., 1997). In the latter case, the authors found that ITN sale through village health workers was a viable way of increasing bed-net coverage. Methods like social marketing (Armstong-Schellenberg et al., 1999) and marketing by the private sector have also been used. The relative costs and outcomes of all these methods are not known; but it be the case that although social marketing methods are effective (Armstong-Schellenberg et al., 2001), they may be too expensive to sustain. Therefore, although social marketing could be used in Nigeria, it is potentially an expensive and unsustainable option, which the public sector may not be keen to undertake.

Distribution through the private sector is usually considered to be able to increase physical access, but it is unclear the type of private sector, referred to, there are many types with different characteristics. The type of private providers that have access to remote villages are the informal ones, consisting principally of unlicensed and/or itinerant drug sellers. To use them would require training to improve their services, together with skills in financial management. The ITNs could also be sold as consumer goods in the market and small village shops that sell provisions. Another option that is potentially cost-effective is to integrate the sales of ITNs into other established disease control programmes like community-directed treatment with ivermectin.

The sales of ITNs through established community-directed treatment with ivermectin (CDTI) distribution systems, used for the control of River Blindness (onchocerciasis) is an attractive option. Nigeria has more than 40 million people (out of a total population of 120 million) at risk of infection, and 7 million people infected by onchocerciais (Edungbola, 1991). The control of the disease is through mass distribution of ivermectin, with the use of community-directed distributors (CDDs). Most of the onchocerciasis-

endemic communities in Nigeria have CDTI schemes and hence, the system could be implemented in more than 30% of the country.

CDDs that already exist in the onchocerciasis-endemic communities, lack anything to do after the yearly ivermectin distribution exercise. This is one reason for the rapid turn-over of CDDs. The CDDs could be trained to handle the extra responsibility for ITN sales. This will also have the effect of mutually strengthening the CDTI system since the CDDs will be busy until the next ivermectin dosing round. However, it is not clear how this system would operate in practice or the processes involved in integrating sales of ITNs with ivermectin distribution, together with the costs and outcomes of such an integrated strategy.

The sales of ITNs could also be integrated into other community-based disease control programmes such as the sale of filters for the control of guinea worm, the implementation of directly-observed treatment for the control of tuberculosis, expanded programme on immunisation and family-planning services. In addition, village-health worker system could be re-invigorated for the twin role of drug treatment of uncomplicated malaria and selling ITNs. They would have to be part of the official heath care system, be formally employed and paid regular salaries. However, the processes, costs and outcomes of using any of these systems for the sale of ITNs and for net-re-treatment would need to be determined before any system is adopted.

In general, an effective system for selling and financing ITNs has to be flexible to accommodate community-specific preferences, and also to follow the general market trends in pricing (Onwujekwe et al., 2001). This will ensure that the sales and payment systems used will satisfy the demands of the consumers. However, the success of any distribution system will be determined by many factors such as: ready availability of nets and its coverage; the acceptability of the inherent payment mechanism to the buyers and sellers; as well as good financial management; and adequate monitoring and supervision of the system.

9.6 What are the distributional implications of the stated and actual WTP for resource allocation?

The disparity in the socio-economic classes of people predominantly willing to pay and those unwilling to pay was evident, and has implications for the analysis of WTP data. This is specifically on issues regarding equity of ITNs programmes and using WTP values for resource allocation decisions. The comparison of the two groups of respondents across the three question formats' groups revealed that the poorer respondents were less willing to pay.

Hence, despite its usefulness, the consensus is that a blanket application of the CVM without consideration of equity and ability of households to pay may undermine its purpose and usefulness (Boadu, 1992). Thus, the results have implications for using stated WTP to determine societal values for goods and services, and for prioritising resource allocation. This is because, the valuations of the rich could superceed that of the poor. A minority of rich respondents could overvalue the WTP, which will skew the data and increase the mean, thereby influencing resource allocation to their benefit. This is a major concern when using the CVM for resource allocation decisions.

In addition, most of those who were not willing to pay for the nets also spent less money on preventing and treating malaria. They were also more likely to be farmers. The correlation of these factors with WTP attests to the internal validity of the WTP method, as they accorded with the expectations from demand theory. Donaldson et al. (1997a) also assessed the effect of ability to pay on WTP by determining whether there was a statistically significant difference in WTP by income and social class characteristics. They found that those with a lower ability to pay levels were less likely to be willing to pay.

Findings that richer households were more willing to pay; as this study found, was also the reason that the Pareto optimum criterion was replaced by the Potential Pareto-improvement criterion. The potential Pareto criterion is based on the idea that if WTP is greater than compensated demanded, the change should go ahead, whether or not the compensation is paid. This criterion is consistent with the economic principle of efficiency (Donaldson, 1993). However, since no compensation is paid in reality, ways to ensure that the resource allocation decisions from WTP surveys are equitable should be designed.

The equity implications of estimates of WTP elicited from CVM require that they be addressed for equitable resource-allocation decisions. The analyst could leave the gains and losses from WTP expressed in monetary terms, make the distributional consequences explicit and leave it to the decision maker to judge whether or not a proposed change should take place (Donaldson, 1999). The second is actually to adjust monetary values through the use of distributional weights (Donaldson, 1999) or equity weights (Sassi et al., 2001). Hence, after diagnosing that different income or socio-economic classes have different estimates of WTP for the good or service and with higher groups being more willing to pay, a means to solving the problem would be the use of distributional weights. Such weights would enable policy makers to appreciate the problem better and then allocate resources in an equitable way.

The distributional weights may be derived from two main sources namely; the views of a (representative) sample of the population and the views of decision-makers (Sassi et al, 2001). Donaldson (1999) showed how to convert WTP values through the use of the weights. One way is to look at past decisions to establish such weights, assuming that the previous decisions resulted in an equitable solution, which may not be the case (Merewitz and Sosnick, 1971). Another method is to use the marginal rates of income taxation, based on the premise that a progressive income tax structure aims to distribute the real burden of any increased taxation amongst different income groups. A third method is to establish weights which are inversely related to income, usually in the form of a logarithmic function (Haberger, 1978; Philips et al., 1989).

The drawbacks from all these sources and uses of distributional weights were discussed by Donaldson (1999) and Sassi et al. (2001). The summary point is that they are all arbitrary. Furthermore, Sassi et al. (2001) felt that the empirical opinion towards rationing healthcare shows that current knowledge of societal preferences in the UK is not sufficient to form the basis for a set of equity weights. However, as Donaldson argued, it is also arbitrary not to adjust raw WTP values. For example, if a poor person and a rich person are each willing to pay £10 more, these values are treated as though they have equal significance in terms strengths of preference of the rich and poor person (Donaldson, 1999). Meanwhile, they are not and the use of different sets of distributional weights through sensitivity analysis would enable the policy makers to make more informed resource-allocation decisions that would be equitable, using estimates of WTP as the basis for such decisions.

9.7 Was consumer surplus valued fully by the WTP estimates in this thesis and of what use is the result for policy making?

Since the theoretical underpinning of the CVM is the valuation of consumer surplus, a consideration of whether the estimates of WTP elicited in this study fully valued consumer surplus is imperative. This will illustrate what level of consumer surplus was valued, and the implications for assuming that they represent societal valuation of the resources. This is because if the estimates WTP do not fully value consumer surplus, the societal benefits could be undervalued with severe implications for resource-allocation, as an intervention that is worthwhile may erroneously not be implemented.

People are often sceptical about monetary valuations of health benefits and express doubt about the reliability and validity of such estimates. Hence, only by demonstrating that elicited estimates of WTP are reliable and valid would such results be accepted as useful and representing credible measures of consumer surplus, especially by non-economists. Then, the resultant CBA or price-setting outputs that used such WTP measures as inputs could confidently be applied for policy making by both economists and non-economists.

The estimates of WTP provide a comprehensive measure of benefits as it combines multiple outcomes (Phillips et al., 1997). The estimates of WTP in this thesis being reasonably valid representations of the respondents' valuation of benefits of the ITNs and net re-treatment could be used as the measure of benefit in CBA. In addition, they could be used to inform policy makers on the optimal pricing for the goods and services. However, while the mean WTP would be used in CBA, the median WTP would be used to inform pricing policies, since the median is the price that at least 50% of the people were willing to pay.

The WTP valued could be for use, non-use and altruistic benefits, and thus for resources that have all these attributes, the consumer surplus would be expected to include elements of the three measures. Therefore, the consumer surplus from public goods would have either one or all these measures depending on the specification of the scenario. However, for private goods like ITNs or their re-treatment in this study, especially where the scenario made it explicit that the respondents were valuing the resources for personal use, the elicited WTP would be taken as a measure of consumer surplus. Nevertheless, an argument could be how the WTP for other peoples' nets figured in getting the total value of the resources to the respondents. The response would be that the different WTP are not additive, since the respondents knew from the scenario that they were evaluating the value of their satisfaction for buying the nets for other people.

However, because the ITNs and re-treatment services are, and were presented as private goods to the respondents, it is argued that the estimates of WTP were full measures of consumers' surplus for the nets and the mean WTP could be seen as the societal value for the products. Taking the argument further, the WTP for other peoples' nets could be taken to represent the lower bounds of societal valuation of the benefits of the ITNs. This is because the respondents took into their account their budget constraint and the fact that they had stated the WTP for own net, thus leading to lower mean WTP for other peoples' nets. Therefore, holding real income constant, the consumer surplus decreased because of the earlier purchase.

It could be argued that since ITNs may have positive externalities, such as mass action (where because many people are using the nets, others may get protected), there should be elements of altruism or non-use value in the WTP, to represent societal valuation of the resources. This is because the society would incorporate all the benefits of the resources in arriving at a value. However, the positive externalities of ITNs has not yet been proven and thus, it would be incorrect to expect the valuation of unproven beneficial aspects of ITNs programmes to be in society's welfare. Including the incorrect benefits in the valuation of consumer surplus may lead to either an inflation, or a deflation of the benefits, thus leading to erroneous cost-benefit computations and interpretations on whether it is worthwhile to implement the intervention.

There could be questions about the implications of this for contingent valuation studies determining the value of public goods, both in the health and other sectors. However, this should not be a problem as far as the researcher explicitly determines what to value and ensure that the values are specified in the scenario. Also, the researcher should determine the likely aspects of societal valuation of the resource through pilot qualitative or quantitative research to determine whether use, non-use and altruism should be in the scenario. Unfortunately, this was not undertaken in this thesis because the insight was not clear during the design of the empirical part of the study, since the assumption was that ITNs programmes are entirely private goods, but would be done in future studies should a mass effect be proven. However, the processes for the selection of the question formats and improving the validity of the scenario could be used for the valuation of public goods and in other fields of economics.

The use of the CVM to generate estimates of consumer surplus in the study area is still experimental and the results should be seen in that light. However, non-economists (and even economists) may consider that some of the unexpected discrepancies, such as the finding that some variables that had the wrong signs in tests of construct validity and the lower test-retest reliability coefficients, undermine any value of WTP estimates. However, no measurement that requires the normative responses of respondents and views of researchers in designing the study instruments and analysing the results is infallible. Researcher bias in designing the study tools and analysing the data collected, interviewer bias in collecting the data and interviewee bias in responding to questions could all act to undermine the accuracy of results.

The results should not be seen as perfect, but should be seen as reasonable representations of reality. Also, even when the estimates of WTP change, the changes could be reflections of actual life experiences. It does suggest that sensitivity analysis would be useful when using estimates of WTP as measures of benefit in CBA. The range of changes in WTP in this study could form the basis for stipulating the confidence limits of sensitivity analysis in future studies in the study area, so that resource-allocation decisions would reflect people's preferences, given changing situations.

9.8 What are the limitations of the thesis and how can they be addressed in future studies?

To date, there has been very few empirical modelling of the causes of divergences between actual and stated WTP, and as such there was no precedent to follow in choosing which variables to include in that aspect of the study. The qualitative evidence from the households through focus group discussions on possible causes of divergence was useful in identifying a few variables, but not how to measure them. In future studies, the findings of this study could form the basis for the expectations on the effect of different variables on divergences in WTP. Also, in-depth qualitative interviews with respondents with divergences in WTP could be used to understand the causes and how they affect the interpretation of the validity and reliability of the CVM, together with that of the question formats used.

The existence of starting-point bias in the WTP elicitation using the three formats was not tested, as it would have made the study design unduly complicated, and also would have required a huge sample size. Nevertheless, the same starting-points were used for all question formats, and so the starting-point bias, if present, is likely to be evenly spread. Previous studies in south-east Nigeria using the bidding game did not detect the presence of starting-point bias (Whittington et al., 1992; Onwujekwe and Nwagbo, 2001).

The WTP for ITNs alone was determined first, and that for ITNs re-treatment were determined later in the questionnaire. The WTP for the other immediately competing health-care needs of the respondents' treatment for malaria or onchocerciasis was not asked. However, in a resource allocation context, it is necessary to determine the relative WTP for a variety of services (Phillips and Zeckhauser, 1989). So, it is not clear whether the presentation of households competing needs would have affected the elicited WTP. Simultaneously asking about many services results in a WTP lower than the WTP obtained from asking about the projects separately (Diamond and Hausman, 1994). One could argue that if the scenario has these other elements, thus enabling the respondents to better appreciate their budget constraints in making trade-offs, more accurate measurements of value of the good could be determined.

The use of market behaviour on actual trading of ITNs and re-treatment services would have been better for developing the structured haggling iterations and the scenario. This may have given the real picture on the nature of the good from the consumers' perspective, the price-taking mechanism and information of perceived attributes that would have been included in the scenario. However, as the nets were not yet sold in the market, untreated nets were used to observe trading behaviour and to aid the development of the structured haggling technique.

There were particular problems in determining inter-rater reliability accurately, as it was difficult encouraging respondents to participate in more than one interview on the same day or within a few days of the first survey. This problem meant that instead of conducting the two interviews within a maximum interval of 3 days, we used a maximum of five days. This may have caused more people to genuinely change their minds, thus potentially lowering the inter-rater reliability of the estimates. It was also problematic having all the interviews one month later to repeat the surveys during the actual sales of ITNs, for optimal determination of test-retest reliability. Three of the interviewers had examinations, and so were unavailable to conduct the retest and had to be replaced by the supervisors that also administered the questionnaires to the respondents during the first survey, for generating data for inter-rater reliability.

The number of respondents whose stated WTP reached the price at which the nets were sold was not apparently enough for testing the relative positive predictive validity across the question formats. A large enough sample for calculation of positive predictive validity that can distinguish between the different question formats require that the price of the ITNs should be the median stated WTP, so that at least 50% of the respondents would be in that sample. However, the state malaria control programme that supplied the net insisted that their price be used, which was at least 100% higher than the median WTP. Thus, the project subsidised the price, but by enough to lower the price to the level of the median.

The one-year span of the study meant that it was not possible to compare the stated and actual WTP for net re-treatment. This was because, the earliest recommended period by the State Malaria Control Unit for re-treating the nets, was at least one year the first use. The use of WTP for net-treatment to test for criterion validity could have been able to better indicate the statistical difference between the three question formats in both phi

coefficient and positive predictive validity. This was because most respondents stated WTP amounts that were enough to cover the price of re-treating the ITNs. The state malaria control programme will conduct the re-treatment services, and it will be interesting to find out the response.

The differences found in some socio-demographic and economic indicators across the three question formats could lead to people querying the validity of the comparative performance of the question formats. However, differences in the variables across the three question formats is not expected to affect tests of content and criterion validity and the tests of reliability, due to the reasons already explained in sections 5.2.2 and 5.6.3. Also the fact that most of these variables significantly explained both stated and actual WTP in only one or pairs of question formats, showed that the results of comparative performance noted are acceptable. Nevertheless, future studies may need to sub-divide the household list into groups with similar socio-economic and asset holdings and then randomly assign similar numbers of respondents with similar characteristics to different question formats in order to improve the randomisation of the respondents.

Since there were multiple statistical tests conducted in the thesis, Bonferroni adjustments would have allowed for corrections so that one would not be over-enthusiastic about the statistically significant results found, as some statistically significant results could have occurred by chance. Bonferroni adjustments are useful when the same test is repeated in many sub-samples, such as when stratified analyses (by age group, sex, income status, etc) are conducted without an a priori hypothesis that the primary association should differ between these sub-groups (Perneger, 1998). However even in such situations, simply describing what was done and why, and discussing the possible interpretations of each result, should enable the reader to reach a reasonable conclusion without the help of Bonferroni adjustments (Perneger, 1998).

These limitations could be surmounted in future studies that are better funded and last longer. The better appreciation of policy-makers for the information generated by research will also help in accessing disease control tools, and the independence of existing price structures to decide how tests of actual validity should be conducted. Also, improved guidelines for CVM studies that account for different contexts could provide information that could be used to surmount some of the limitations found in this thesis.

CHAPTER 10: CONCLUSION AND RECOMMENDATIONS

The first section of this chapter is devoted to the general findings of the thesis. The second section presents key areas in which the thesis has added to knowledge in contingent valuation method (CVM), also including some recommendations related to the use of CVM. The third section deals with policy implications of the thesis. Key areas for future research are presented in the last section.

10.1 Key findings from the thesis

From the empirical part of the work, it is concluded that the contingent valuation method is a reasonably reliable and valid tool for determining people's willingness to pay for ITNs and net re-treatment in Nigeria. Also, the respondents, who were mostly illiterate people, understood the CVM questioning and gave what could be said to be truthful answers, as judged from the tests of validity. It was seen that even in cases where there were low criterion validity scores, there were genuine reasons for the divergences between stated and actual WTP. Thus, like Whittington et al. (1990), it is concluded that "the results suggest that it is possible to conduct a contingent valuation survey among a very poor, illiterate population and obtain reasonable, consistent answers".

Considering the results of all tests for reliability and validity, none of the different question formats in the study was significantly better than others. The structured haggling technique was the most content-valid of them all. On the other hand, the bidding game generated the best data for construct validity of ITN, while the structured haggling technique's data was the best for construct validity of ITNs re-treatment. However, the criterion and test-retest reliability results were inconclusive as higher level of genuine causes for divergences in WTP were mostly found in the BWFU and bidding game that had lower performance scores than the structured haggling group. The three question formats were indistinguishable using tests of inter-rater reliability.

The literature did not provide all the tools needed for data collection, analysis and interpretation of the findings. This is because, crucial to understanding the estimates of stated and actual WTP is an understanding of the reasons that lie behind decision-making of the respondents. However, advances have been made in CVM literature to unravel such reasons that lie behind this in order to understand situations wherein WTP could be said to biased or not. Incorporation of psychometric principles in CVM would have been quite helpful.

The findings on content validity of question formats could be generalised to all parts of south-east Nigeria, and in other contexts and areas where bargaining for goods and services is the normal price-taking mechanism. However, studies need to ensure that their question formats are content-valid and, if not, attempts should be made to improve their content validity. Similarly, the potential respondents and possible providers could help to improve the scenario and ensure that it describes what actually should be valued. This will ensure that the consumer surplus for the resource is fully valued.

Content validity tests of the question formats should be an important part of WTP studies. There should not be any rigid structure of question formats, and they should be modified based on the input of the consumers. The use of question formats that are not proven to be content valid in the context that they will be used should be discouraged. However, this argument could be countered as the bidding game and BWFU method also performed reasonably in this study. The response would be that they all had some levels of content-validity, and the BWFU was actually improved from the feedback of traders to make it more content-valid.

Thus, the use of context-specific question formats is advocated, as such methods would catalyse the respondents' thought processes as they assign values to the good in question, interpret the questions well, and thus give more valid and reliable responses. The study only used methods that the researcher adjudged and confirmed by the traders and consumers, to have some level of content-validity in the study area. The crux was to see whether a novel system that resemble the usual price-taking behaviour of the people in a particular context and study area, would perform better. The end point being that all elicited WTP would be able not only to be used to resolve some of the methodological debates in CVM, but also generate policy relevant estimates.

From the econometric modelling of WTP estimates, different question formats generate data that require different models. Some question formats would yield better construct-valid parameters with some models than with other models. However, construct validity would depend on how well the data for the modelling were measured, coded and stored. Improving results for construct validity will include better ways of measuring the variables, administering the questionnaires, training and supervising field workers.

The Heckman's estimator was consistently better than the Tobit model, while the OLS was better than the truncated regression models in estimating only the positive WTP responses. These differences in the models were determined through diagnostic testing of the models. Therefore, asserting the construct validity of WTP requires the diagnostic testing of different econometric models. Such testing is advocated for studies determining theoretical validity with econometric models.

While the type of question format mattered for comparing stated and actual WTP, it was not so crucial for determining actual WTP. This is because while the phi coefficients revealed some differences in the performance of the question formats, the actual sales recorded in each group were almost equal. Also, there were genuine causes of divergences between actual and stated WTP and levels varied systematically between question formats' groups. Hence, basing long-term resource-allocation decisions on only stated WTP might be problematic as genuine changes could lead to changes in consumer surplus over time.

Many cases of divergences between stated and actual WTP were not due to hypothetical bias. Also, stated WTP did not exceed actual WTP in contrast to previous studies that had compared both WTP estimates. The modelling of the causes of these divergences showed that in some situations, these were in line with expectations and therefore should not lead to a condemnation of the CVM, or question formats.

Many loopholes were identified in existing studies that have determined the reliability and validity of the CVM health care. Better study designs and reporting, together with further studies in all aspects of reliability and validity of the method are needed to build evidence either in favour, or against the CVM and the different question formats. The contributions to knowledge and specific conclusions on methodological and policy aspects of the thesis are presented below.

10.2 Methodological Implications

The thesis generated knowledge about the optimal way to conduct CVM studies. A starting point was the design and use of context-specific question formats. The need to ensure content-validity of CVM, together with the steps taken to ensure that the question formats and scenario were content valid before being administered to the respondents, are important contributions to knowledge, since this had never been done in previous contingent valuation studies.

Linked to optimal ways for conducting contingent valuation questioning was the development of a novel question format that mimicked more closely the price-taking behaviour of the study population. The explicit steps taken to design the novel technique, which ranged from literature review to the final pre-tests were unexplored frontiers prior to this thesis. It is hoped that future studies interested in the development and use of context-specific question formats could use some of the methods used in this thesis, to ensure that their question formats are content-valid.

The study also contributed to establishing the inter-rater and test-retest reliability of the CVM, especially that of different question formats. The reliability of the CVM has been grossly under-researched and the 18 reliability coefficients generated in this thesis will enrich the literature. The explicit determination of the causes of lower test-retest reliability and the knowledge that sometimes the lower test-retest reliability attest to the sensitive nature of the CVM is important knowledge. Hence, the reliability of the CVM is made clearer and more acceptable, especially since the conventional definitions, analytic tools and interpretations of the reliability estimates entrenched in the measurement theory were utilised in this thesis.

The methods for establishing the validity of the question formats were comprehensive and tests of criterion validity were developed from principles of measurement theory. Most of the previous comparisons of stated and actual WTP had been flawed because they generally estimated levels of WTP at different times from different respondents. Others elicited levels of WTP from two sub-groups; with one group stating only hypothetical WTP and the other offered a unit of the good being valued (for all the respondents). The establishment of criterion validity using a big sample, with all the respondents being offered the good to purchase, is a contribution to knowledge.

Both methodological and analytic weaknesses existed in the CVM literature in the area of testing criterion validity. The use of the phi coefficient as the ultimate measure of validity is new in the WTP literature and should provide a standard method of comparing results between different areas and contexts in the future. The comparison of criterion validity across the three question formats and determining the factors that determine actual WTP in comparison to those that determine stated WTP, is a new methodological approach in CVM.

Specific tests of validity like model specification approaches for the construct validity comparisons of the different question formats will be another addition to CVM methodology. This was because diagnostic testing was made an intrinsic part of model specification to guard against the use of spurious models to assert construct validity. It was seen that some models whose outputs were consistent with theoretical constructs of many parameters were wrongly specified. Thus, the thesis showed conclusively that models should be accompanied by careful diagnostic testing.

The structured approach for determination of the causes of divergences in WTP is additional knowledge that should ideally be part of any study that seeks to compare stated and actual WTP, or seeks to determine test-retest reliability. This would enable researchers to assert with confidence whether detected divergences are be attributable solely to hypothetical bias or unreliability of the estimates. The determinations would also show whether there were genuine reasons for the divergences that actually point to the validity or reliability of the CVM and to measure the direction of change appropriately.

10.3 Policy Implications

The CVM was both reliable and valid and hence can be used to credibly measure the monetary benefit of health programmes and as a tool for market research. Furthermore, an idea of the price that most people were willing to pay for ITNs in south-east Nigeria is now known, together with the potential demand for the good. This information will be useful in determining whether subsidies should be introduced and how much they should be.

The specific factors that determine whether people will or will not pay for ITNs, will be vital information for policy makers in their efforts to increase acquisition and usage of the nets. Finally, policy makers will find the factors that affect actual WTP of assistance in planning and implementing more efficacious ITNs sales programmes. Such information is part of this thesis.

The estimates from the WTP valuation will aid resource allocation and optimal pricing of disease control tools. However, comparisons of people willing and not willing to pay showed that using WTP estimates for resource allocation or other policy actions could lead to inequity. This is due to the fact that the preferences of the poorer members of society will not be protected, since they stated less WTP for the goods.

Nevertheless, most people were willing to pay for ITNs, and it is likely that this result could be generalised to other communities in Nigeria, especially south-east Nigeria, since almost all the communities there have very similar socio-economic and malaria profiles. This assertion is supported from the results of an earlier study in five communities of south-east Nigeria (Onwujekwe et al., 2001), where there were similar mean and median WTP for ITNs as those found in this thesis.

The thesis shows that to potentially increase coverage with ITNs, the factors that may discourage people from purchasing nets need to be mitigated. Distribution systems that will guarantee enhanced physical access should be put in place. Payment systems that will help the financially disadvantaged and high risk groups (e.g. children and pregnant women) also need to be designed and implemented to increase coverage with ITNs.

There should be a national policy to improve access to ITNs, and to encourage competition in pricing. The insistence of the Enugu State Malaria Control Unit on a rigid national price structure when it was evident that the people were not willing to buy the nets at the national price will not augur well for increasing net coverage.

The determinants WTP (stated and actual) showed that the values were the results of clear thinking by the people. These determinants should inform policy formulation of an ITNs distribution programme, as no such programme exists in Nigeria. The biggest challenge that will involve a multi-ministerial approach is to empower the people to be able to buy the nets through schemes that will improve their income earnings.

Accurate information on the possible causes of divergence between stated and actual WTP is of great importance in planning and forecasting trends in sales of nets. Causes of divergences should also inform policy makers and programme managers as to the best periods to intensify net sales. It can also help to determine promotional and advocacy approaches to increasing net sales and usage.

10.4 Future Studies

The thesis has also shown that there are still many areas for future work in CVM. It is crucial that more work on the CVM needs is conducted in Africa, where there is currently dire paucity of literature on CVM. Critical areas would include the development and testing of different question formats in different areas. Another critical area would be to develop better ways to present CVM scenarios in local contexts. Additionally, WTP for various types of health goods and services should be researched and the reliability and validity of estimates determined.

The question formats used in this thesis, especially the structured haggling technique could be used in future studies using the CVM in areas where haggling for goods and services is the dominant price-taking technique in the markets. However, the steps used in improving the content validity of the BWFU and the structured haggling technique could be used to improve a particular question format to be used in a study if none of the existing formats resemble price-taking there. Nevertheless, attempts could always be made in future studies to improve on the content-validity of their question formats, even if such mimic the price-taking behaviour in a particular area.

Future studies should examine the reasons behind divergences between stated and actual WTP and between inter-rater and test-retest reliability, so that conclusion on the existence of bias will be credible. More hypotheses need to be generated with more variables in order to inform and broaden the discussion on the causes and direction of the divergences between stated and actual WTP. Quantitative modelling of these divergences should be done in addition to qualitative interviews with the respondents. Such research should include those that diverged, as well as those that did not, so that reasons for consistency of responses and vice versa is known and accounted for.

It is also important that work should continue on determining the best question formats that could be used in different contexts and in different areas, using both validity and reliability tests (especially criterion validity). The studies that have compared criterion validity of different question formats had been constrained by either small sample sizes or other methodological problems and better study designs would be needed in future to overcome these problems and increase the power of the tests to detect differences in criterion validity. The comparisons may involve the use of formats that people judge to be non-content valid in contrast with those having acceptable levels of content-validity. Such a comparison will show the extent to which content-validity of the question format improves reliability and validity.

In the area of WTP for ITNs, future and broader studies could examine whether there are differences in WTP for ITNs in areas with different levels of malaria endemicity. This is because this thesis found a significant difference in WTP between those households that had malaria prior to the survey and those who did not. It would be interesting to ascertain whether as the incidence of malaria is reduced (for example through the wider use of ITNs), the WTP would change.

Finally, more comprehensive guidelines are needed for the conduct, analysis and reporting of tests of reliability and validity (especially criterion validity). In this vein, CVM may need to borrow from psychometric theory where such measures are more developed. Also, more work is needed on diagnostic tests to be presented with econometric models used to determine construct validity. All these will ensure that CVM findings are credible, easily assessed by external people, and amenable for decision-making purposes.

The methods used in this thesis are generalisable to other settings in different ways. The tests of validity and reliability used can be applied in any context and area. The only difference is that the tests have been used in a more organised way in this thesis, because one led into another in a continuum process. The content validity of the question formats were first determined and improved upon. At the same time the content validity of the scenario was improved. Then, stated WTP was elicited for tests of construct validity and inter-rater reliability, before finally actual WTP and retest were conducted for criterion validity and test-retest reliability. Thus, it is advocated that

future studies should also organise the various tests, as has been done in this research, in order to improve the reliability and validity of results.

All in all, the CVM could be used to elicit valid and reliable measures of consumer surplus for private health-care goods, amongst the Igbo people of Nigeria. However, though found to be reliable and valid, there was no dominant question format of the three utilised, in tests of reliability and validity, where one of the question formats was developed as part of this thesis to better mimic the price-taking behaviour of the Igbo people. Furthermore, differences between stated and actual WTP or a retest WTP after a period of time does not always mean that bias exists, but may actually show that the CVM was a valid and reliable approach, if the causes of the divergences are genuine. The factors that explain stated and actual WTP and divergences between the two point to policy-making areas and opportunities for increasing coverage and ensuring that equity is maintained in ITNs programmes. Further studies, using larger sample sizes and improved methods will better inform people on whether the type of question format used in different settings is a necessary pre-requisite for improved reliability and validity of CVM.

REFERENCES

Abel-Smith B and Rawal P (1992) Can the poor afford free health services: a case study of Tanzania. *Health Policy and Planning* 7(4), 329-341.

Acton JP (1975) Non-monetary factors in the demand for medical services: Some empirical evidence. *Journal of Political Economy*, 595-614.

AMREF (1997) *An assessment of the availability of bednets in Uganda.* African Medical and Research Foundation, Kampala.

Agbonifoh BA and Bakwenye N.M. (1985) Correlates of haggling behaviour. *Benin Journal of Social Science,*1:27-35. (quoted in Balogun, S. K. (1991). Bargaining strategy and sex as determinants of bargaining outcomes in open markets. *Nigerian Journal of Economic and Social Studies* 33(1), 31-38.

Akin J S, Guilkey DK and Denton HE (1995) Quality of services and demand for health care in Nigeria: a multinomial probit estimation; *Social Science and Medicine* 40(11), 1527-1537.

Akin JS, Griffin CC, Guilkey DK and Popkin BM (1985) *The demand for primary health services in the Third World.* Rowman and Allanheld; New Jersey.

Alberini A (1995) Testing willingness to pay models of discrete choice contingent valuation survey data. *Land economic* 71(1), 83-95.

Albernini A and Krupnick A (1998) Air quality and episodes of acute respiratory illness in Taiwanese cities: evidence from survey data. *Journal of Urban Economics* 44, 68-92.

Amemiya T (1984) Tobit models: A survey. *Journal of Econometrics* 24, 3-61.

Appel L, Steinberg E, Powe N, Anderson, G, Dwyer S and Faden R (1990) Risk reduction from low osmality contrast media: what do patients think it's worth? *Medical Care* 28, 324-337.

Arhin AD (1995). *Rural Health Insurance: A Viable Alternative to User Fees? A Feasibility Study in Ghana.* London, London School of Hygiene and Tropical Medicine: 74-123.

Armstrong Schellenberg JRM, Abdulla S, Nathan R, Mukasa O, Marchant TJ, Kikimbih N, Mushi AK, Mponda H, Minja H, Mshinda H, Tanner M and Lengeler C (2001) Effect of large-scale social marketing of Insecticide-treated nets on child survival in rural Tanzania. *The Lancet* 357,1241-1247.

Armstrong Schellenberg, JRM, Abdulla S, Minja H, Nathan R, Mukasa O, Marchant TJ, Mponda H, Kikimbih N, Lyimo E, Manchester T, Tanner M and Lengeler C (1999) KINET: a social marketing programme of treated nets and net treatment for malaria control in Tanzania, with evaluation of child health and long-term survival. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 93,225-231.

Arrow KJ (1984) *Collected papers of Kenneth J.* Arrow. Volume 3: Individual choice under certainty and uncertainty. The Belknap press of Harvard University press, Cambridge, Massachusetts.

Asenso-Okyere WK, Osei-Akoto I, Anum A and Appiah EN (1997) Willingness to pay for health insurance in a developing economy. A pilot study of the informal sector of Ghana using contingent valuation. *Health Policy* 42, 223-237.

Asgary A and Willis KG (1997) Estimating the benefits of construction measures to mitigate earthquake risks in Iran. *Environmental and Planning B-Planning & Design* 24(4), 613-624.

Bala MV, Wood LL, Zarkin GA, Norton EC, Gafni A and O'Brien B (1998) Valuing outcomes in health care: A comparison of willingness to pay and quality-adjusted life-years. *Journal of Clinical Epidemiology* 51(8), 667-676.

Balogun SK (1991) Bargaining strategy and sex as determinants of bargaining outcomes in open markets. *Nigerian Journal of Economic and Social Studies* 33(1), 31-38.

Basden GT (1938) *Niger Ibos.* London; Seeley, service and Co. Ltd.334-338.

Bateman IJ and Turner K (1993) The Contingent valuation method, in Turner K (ed) *Sustainable Environmental Economics and Management*. London: Belhaven.

Berger M, Blomquist G, Kendel D and Tolley E (1987) Valuing changes in health risks: a comparison of alternative measures. *Southern Economic Journal* 53, 967-984.

Berwick DM and Weinstein MC (1985) What do patients value? Willingness to pay for ultrasound in normal pregnancy. *Medical Care 23*(7), 881-893.

Binka F and Adongo P (1997) Acceptability and use of insecticide impregnated bednets in northern Ghana. *Tropical Medicine and International Health* 2(5), 499-507.

Binka FN, Kubaje A, Adjuik M et al. (1996) Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana District, Ghana: a randomised controlled trial. *Tropical Medicine and International Health* 1, 147-154.

Birdsall N and Chuhan P (1983) *Willingness to pay for health and water in rural Mali: Do WTP questions work?* Country policy department discussion paper No. 1983-8. The World Bank. Washington D.C.

Bishop RC and Herbelein TA (1979) Measuring values of extra-market goods: Are indirect methods biased? *American Journal of Agricultural Economics* 61, 926-30.

Bishop R, Welsh M and Herbelein T (1992) Some experimental evidence on the validity of contingent valuation. Department of Agricultural Economics, University of Wisconsin. Quoted in 'Loomis J et al (1997). Evaluating the validity of the dichotomous choice question format in contingent valuation. *Environmental and Resource economics* 10, 109-123.

Blumenschein K and Johannesson M (1996) Economic evaluation in health care. A brief history and future directions. *Pharmacoeconomics* 10(2),114 –122.

Blumenschein K et al (1998) Experimental results on expressed certainty and hypothetical bias in contingent valuation. *Southern Economic Journal* 65(1), 169-77.

Blumenschein K, Johannesson M, Yokoyama KK and Freeman, PR (2001) Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. *Journal of health economics* 20(3), 441-458.

Boadu FO (1992) Contingent valuation for household water in rural Ghana. *Journal of Agricultural Economics* 43(3), 458-465.

Boadway RW and Bruce N (1984) *Welfare economics*. Oxford: Blackwell.

Bowling A (1997) *Measuring Health: A review of quality of life measurement scales.* (2nd edition). Buckingham; Open University Press:10-11.

Brieger WR, Ramakrishna J and Adeniyi JD (1989-90) Community response to social marketing: filters for guinea worm control. *International Quarterly of Community Health Education* 19(1), 3-17.

Brookshire SD, Randall A and Stoll JR (1980) Valuing increments and decrements in natural resource service flows. *American Journal of Agricultural Economics,* 476-488.

Brown R and Green C (1981) Threats to health or safety: perceived risk and willingness to pay. *Social Science and Medicine* 15, 67-75.

Brown T, Champ P, Bishop R and McCollum D (1996) Which response format reveals the truth about donations to a public good. *Land Economics* 72(2), 152-166.

Brown TC, Barro SC, Manfredo MJ and Peterson GL (1995) Does better information about the good avoid the embedding effect? *Journal of Environmental Management* 44, 1-10.

Cameron TA and Quiggin J (1994) Estimation using contingent valuation data from a 'Dichotomous choice with follow-up' questionnaire. *Journal of Environmental Economics and Management* 27, 218-34.

Carmines EG and Zeller RA (1979) *Reliability and validity assessment.* London; Sage Publications.

Carson RT, Hanemann M, Kopp RJ, Krosnick JA, Mitchell RC, Presser S, Ruud PA and Smith VK (1997) Temporal reliability of estimates from contingent valuation. *Land Economics* 73,151-63.

Carson RT, Wright JL, Carson NJ, Alberini A and Flores NE (1994) *A bibliography of contingent valuation studies and papers.* Natural Resources and Damage Assessment Inc., La Jolla, California.

Carson RT (1991) *Constructed markets:* In J.B. Braden and C.D. Kolstad (Eds), Measuring the Demand for Environmental Quality, Elsevier/North-Holland, Amsterdam.

Cham MK, Olaleye B, D'Alessandro U, Aikins M, Cham B, Maine N, Williams LA, Mills A and Greenwood BM (1997) The impact of charging for insecticide on the Gambian National Impregnated bednet programme. *Health Policy and Plannning* 12(3), 240-247.

Champ PA, Bishop RC, Brown TC and McCollum DW (1997). Using donation mechanisms to value non-use benefits from public goods. *Journal of Environmental Economics and Management* 33, 151-162.

Clarke PM (2000) Valuing the benefits of mobile mammographic screening units using the contingent valuation method. *Applied Economics* 32(13), 1647-1656.

Clarke PM (2001) Testing the convergent validity of the contingent valuation and travel cost methods in valuing the benfits of health care. *Health Economics* (In press).

Cohen J (1987) *Statistical power analysis for the behavioural sciences (rev ed.).* New Jersey: Lawrence Erlbaum Associates.

Cooper J and Loomis J (1992) Optimal bid selection for dichotomous choice contingent valuation surveys. *Land Economics* 68, 211-222.

Cronbach LJ and Meehl PE (1955) Construct validity in psychological tests. *Psychological Bulletin* 52:281-302. Quoted in Carmines EG and Zeller RA (1979). *Reliability and validity assessment.* London; Sage Publications.

Culbertson V, Arthur T, Rhodes P and Rhodes R (1988) Consumer preferences for verbal and written medication information. *Drug Intelligence and Clinical Pharmacy* 22:390-396.

Cullis JG and West PA (1991) *The economics of health: An introduction.* Gregg Revivals: Hampshire, England.

Cummings RG, Harrison GW and Rustrom EE (1995) Home-grown values and hypothetical surveys: Is the dichotomous choice approach incentive-compatible? *American Economic Review* 85(1), 260-266.

Cummings RG, Brookshire DS and Schulze WD (1986) *Valuing environmental goods.* Rowman and Allanheld, New York.

Culyer AJ (1991) *The Economics of Health,* Volume II, Edward Edgar Publishing Limited, Aldershot.

Culyer AJ (1985) *Economics.* Oxford: Basil Blacwell Limited.

Davis D and Holt C (1993) *Experimental Economics.* Princeton University Press: Princeton.

D'Alessandro U, Olaleye BO, McGuire W, Langerock P, Bennet S, Aikins MK et al. (1995a) Reduction in mortality and morbidity from malaria in Gambian children following the introduction of a National insecticide impregnated bednet programme. *Lancet* 345, 479-483.

D'Alessandro U, Olaleye BO, McGuire W et al. (1995b) A comparison of the efficacy of insecticide-treated and untreated bednets in preventing malaria in Gambian children. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 89, 596-598.

Dalmau-Matarrodona E (2001) Alternative approaches to obtain optimal bid values in contingent valuation studies and to model protest zeros. Estimating the determinants of individual's willingness to pay for home care services in day care surgery. *Health Economics* 10, 101-118.

Davey P et al. (1998) Economic evaluation of insulin lispro versus neutral (regular) insulin therapy using a WTP approach. *Pharmacoeconomics* 13(3), 347-358.

Davis RK (1963). Recreation planning as an economic problem. *Natural Resources Journal* 3, 239-249.

Desvousges WH, Johnson FR, Dunford RW, Boyle KJ, Hudson SP and Wilson KN (1993) *Measuring natural resources damages with contingent valuation: tests of validity and reliability*. In Hausman JA (Ed.) *Contingent valuation: a critical assessment*. New York: North-Holland.

Dgedge M, Gomes A, Cossa H, Kumaranayake L and Lines J (1999) Willingness to pay for insecticide treated nets before and after implementation of ITN in a semi-rural district of south Mozambique. *Book of Abstracts. MIM African malaria Conference*, Durban, South Africa; p.B-29.

Diamond P and Hausmann J (1994) Contingent valuation: Is some number better than no number? *Journal of Economic Perspectives* 8(4): 45-64.

Dickie M, Fisher A and Gerkins S (1987) Market transactions and hypothetical demand data: A comparative study. *Journal of the American Statistical Association* 82(397), 69-75.

Diener A, O'Brien B, and Gafni A (1998) Health Care Contingent Valuation Studies: A Review and Classification of the Literature. *Health Economics* 7, 313-326.

Diez L (1998) Assessing the willingness of parents to pay to pay for reducing post operative emesis in children. *Pharmacoeconomics* 13: 589-595.

Dinwiddy C and Teal F (1996) *Principles of Cost-Benefit Analysis for Developing countries* Cambridge University Press; Cambridge.

Donaldson C (1990) Willingness to pay for publicly-provided goods: a possible measure of benefit. *Journal of Health Economics* 9, 103-118.

Donaldson C (1993) *Theory and practice of willingness to pay for health care.* Health Economics Research Unit Discussion paper 01/93. Departments of Public Health and Economics, University of Aberdeen.

Donaldson C, Mapp T, Ryan M and Curtin K (1996) Estimating the economic benefits of avoiding food-borne risk: is 'willingness to pay' feasible? *Epidemiology and Infection* 116, 285-294.

Donaldson C and Shackley P (1997) Does "process utility" exist? A case study of willingness to pay for laparoscopic cholecystectomy. *Social Science and Medicine,* 44(5), 699-707.

Donaldson C, Shackley P, Abdalla M and Miedzybrodzka Z (1995) Willingness to pay for antenatal carrier screening for cystic fibrosis. *Health Economics*, 4,439-452.

Donaldson C, Farrar S, Mapp T, Walker A and McPhee S (1997a) Assessing community values in health care: is the 'willingness to pay' method feasible? *Health Care Analysis* 5, 7-29.

Donaldson C, Thomas R and Torgerson DJ (1997b) Validity of open-ended and payment scale approaches for eliciting WTP. *Applied Economics* 29, 79-84.

Donaldson C, Shackley P and Abdalla M (1997c) Using willingness to pay to value close substitutes: carrier screening for cystic fibrosis revisited. *Health Economics* 6, 145-159.

Donaldson C, Jones AM, Mapp TJ and Olson JA (1998) Limited dependent variables in willingness to pay studies: applications in health care. *Applied Economics* 30, 667-677.

Donaldson C (1999) Valuing the benefits of publicly-provided health care: does 'ability to pay' preclude the use of 'willingness to pay'? *Social Science and Medicine* 49, 551-563.

Drummond ME, O'Brien B, Stoddart GL and Torrance G W (1997) *Methods for the Economic Evaluation of Health Care Programmes.* New York, Oxford University Press.

Duan N, Manning WG, Morris CN and Newhouse JP (1983) A comparison of alternative models for the demand for medical care. *Journal of Business and Economic Statistics* 1, 115-126.

Duborg WR, Jones-Lee MW and Loomes G (1994) Imprecise preferences and the WTP-WTA disparity. *Journal of Risk and Uncertainty* 9,115-133.

Dupuit J (1844) On the measurement of utility of public works. Translated by RH Barback in *International Economic Papers*, 2(1952):83-110. In Donaldson C (1993) *Theory and practice of willingness to pay for health care.* Health Economics Research Unit Discussion paper 01/93. Departments of Public Health and Economics, University of Aberdeen.

Eckerlund I, Johannesson M, Johansson PO, Tambour M and Zethraeus N (1995) Value for money? A contingent valuation study of the optimal size of the Swedish health care budget. *Health Policy* 34:135-143.

Edungbola LD (1991) Onchocerciasis control in Nigeria. *Parasitology Today* 7, 97-99.

Einarson T, Bootman J, McGhan W, Larson L, Gardner M and Donohue M (1988) Establishment and evaluation of a serum cholesterol monitoring service in a community pharmacy. *Drug Intelligence and Clinical Pharmacy* 22, 45-48.

The Enugu State Government (2001) *Enugu State of Nigeria - The Story of Untapped Potentials*. Enugu State Government, Enugu, Nigeria.

Evans RG (1984) *Strained Mercy: The Economics of Canadian Health Care.* University of British Columbia.

Feilden R et al. (1996). *Willingness and ability to pay for Hepatitis B vaccine.* Interim report. Ministry of Health. Bucharest, Romania.

Feilden RM (1996) *Chapter 3: Experiences of implementation.* In: Net Gain. A new method for preventing malaria deaths. Lengeler C., Cattani J. and de Savigny D. (editors). Geneva; World Health Organisation. Pp55-110.

Fink A (1993) *Evaluation Fundamentals: Guiding Health Programs, Research and Policy.* London: Sage Publications.

Fischoff B and Furby L (1988) Measuring Values: A conceptual framework for interpreting transactions with special reference to contingent valuation of visibility. *Journal of Risk and Uncertainty* 1, 147-184.

FMOH (1992) *The National health policy of Nigeria.* Federal Ministry of Health (FMOH). Lagos.

Flowers CR, Garber AM, Bergen MR and Lenert LA (1997) Willingness to pay utility assessment: feasibility of use in normative patient decision support systems. *Journal of the American Medical Informatic Association Proceedings/AMIA Annual Fall Symposium,* 223-227.

Folland S, Goodman AC and Stano M (1997) *The economics of health and health care (2nd ed.)*. New Jersey: Prentice-Hall.

Foster V, Bateman IJ and Harley D (1997) Real and hypothetical willingness to pay for environmental preservation: A non-experimental comparison. *Journal of Agricultural Economics* 48(2), 123-138.

Fraser-Hurt N and Lyimo E (1998) Insecticide-treated nets and treatment service: a trial using public and private sector channels in rural Tanzania. *Bulletin of the World Health Organisation* 76(6), 607-615.

Frykblom P (1997) Hypothetical question modes and real willingness to pay. *Journal of Environmental Economics and Management* 34(3), 275-87.

Fuchs VR (1968) The growing demand for medical care. *New England Journal of Medicine* 279(190), 190-5.

Gafni A (1998) Willingness to pay. What's in a Name? *Pharmacoecomics* 14(5), 465-470.

Gafni A (1991) Willingness to pay as a measure of benefits. Relevant questions in the context of puvlic decision-making about health care programmes. *Medical Care* 29, 1246-52.

Gertler P, Locay L and Sanderson W (1987) Are use fees regressive? The welfare implications of health financing proposals in Peru. *Journal of Econometrics* 36, 421-435.

Gertler P and van der Gaag J (1990) The willingness to pay for medical care: Evidence from two developing countries. The World Bank; Washington DC.

Golan E and Shechter M (1993) Contingent valuation of supplemental health care in Israel. *Medical Decision Making* 13,302-310.

Goldman F and Grossman M (1978) The demand for paediatric care: A Hedonic approach. *Journal of Polictical Economy* 86(2),259-280.

Gonzalez-Caban A and Loomis J (1997) Economic benefits of maintaining ecological integrity of Rio Mameyes, in Puerto Rico. *Ecological Economics* 21(1):63-75.

Goodman CA, Mnzava AE, Dlamini SS, Sharp BL, Mthembu DJ, Gumede JK (2001) Comparison of the cost and cost-effectiveness of insecticide-treated bednets and residual house-spraying in KwaZulu-Natal, South Africa. *Tropical Medicine and International Health* 6, 280-95.

Gore P and Madhavan S (1994) Consumers' preference and willingness to pay for pharmacist counselling for non-prescription medicines. *Journal of Clinical Pharmacy and Therapeutics* 9, 12-25.

Gravelle H and Rees R (1981) *Microeconomics.* New York: Longman Group Limited.

Green D, Jacowitz KE, Kahneman D and McFadden D (1998) Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resource and Energy Economics* 20(2), 85-116.

Gregory R (1986) Interpreting measures of economic loss: evidence from contingent valuation and experimental studies. *Journal of Environmental Economics and Management* 13,325-337.

Grossman M (1972) *The demand for Health: A theoretical and empirical investigation.* National Bureau of Economic Research; New York.

Guilford JP (1956) *Psychometric methods* (2nd edn). McGraw Hill; New York.

Gyldmark M and Morrison GC (2001) Demand for health care in Denmark: results of a national sample survey using contingent valuation. *Social Science & Medicine* 53(8), 1023-1036.

Haberger AC (1978) On the use of distributional weights in social cost-benefit analysis. *Journal of Political Economy* 86, S87-S120.

Hanemann WM (1984) Welfare evaluations in contingent valuation experiments with discrete responses: Reply. *American Journal of Agricultural Economics* 69, 332-341.

Hanemann WM (1991) Willingness to pay and willingness to accept: how much do they differ? *American economic Review* 81, 635-647.

Hanemann WM (1994) Valuing the environment through contingent valuation. *Journal of Economic Perspectives* 3, 1-23.

Hanemann M and Kanninen B (1996) *The statistical analysis of discrete-response CV data*. Working paper 794, Department of Agricultural and Resource Economics, University of California, Berkeley.

Harrington W and Portney PR (1987) Valuing the benefits of health and safety regulation. *Journal of Urban Economics* 22, 101-12.

Heckman JJ (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurements* 5/4, 475-492.

Heller PS (1982) A model of the demand for medical and health services in Penninsular Malaysia. *Social Science and Medicine* 16, 267-284.

Hellinger FJ (1989) Expected utility theory and risky choices with health outcomes. *Medical Care* 27(3), 273-9.

Herriges JA and Shogren JF (1996). Staring point bias in dichotomous choice valuation with follow-up questioning. *Journal of Environmental Economics and Management* 30, 112-131.

Hicks JR (1941) The rehabilitation of consumer surplus. *The Review of Economic Studies* 8, 108-116

Hicks JR (1943) The four consumer's surpluses. *The Review of Economic Studies* 11, 31-41

Hoddinott J (1992) Modelling remittance flows in Kenya. *Journal of African Economies* 1(2), 207-232.

Hoen JP and Randall A (1987) A satisfactory benefit cost indicator from contingent valuation. *Journal of Environmental Economics and Management* 14, 226-247.

Hongvivatana T and Manopimoke S (1991) *A baseline survey for preference for rural health insurance*. GTZ and Centre for Health Policy Studies, Mahidol University,

Thailand.

Jacobs P (1997) *The economics of health and medical care.* Aspen Publishers Inc: Gaithersburg, Maryland.

Johannesson M and Johansson PO (1996). To be or not to be, that is the question - an empirical study of the willingness to pay for an increased life expectancy at an advanced age. *Journal of Risk and Uncertainty* 13,163-174.

Johannesson M, Johansson PO and O'Conor RM (1996) The value of safety versus the value of public safety. *Journal of Risk and Uncertainty;*13:263-275.

Johannesson M and Johansson P (1997) The value of life extension and the marginal rate of time preference: a pilot study. *Applied Economics Letters* 4, 53-55.

Johannesson M, Liljas B and O'Conor RM (1997) Hypothetical versus real willingness to pay: Some experimental results. *Applied Economics Letters* 4(3), 149-151.

Johannesson M (1997) Some further experimental results on hypothetical versus real willingness to pay. *Applied Economic Letters* 4(8), 535-536.

Johannesson M, Liljas B, Johansson P (1998) An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions. *Applied Economics* 30, 643-647.

Johannesson M, Blomquist GC, Blumenschein K, Johansson P-O, Liljas B and O'Conor RM (1999) Calibrating hypothetical willingness to pay responses. *Journal of Risk and Uncertainty* 18, 21-32

Johannesson M (1996) *Theory and methods of economic evaluation of health care.* Dordrecht, Kluwer Academic Publishers.

Johannesson M (1993) The contingent-valuation method [comment]. *Medical Decision Making* 13(4), 311-2.

Johannesson M, Johansson P, Kristrom B, Borgquist L and Jonsson B (1993a) Willingness to pay for lipid lowering: a health production function approach. *Applied Economics* 25(8), 1023-1031.

Johannesson M, Johansson PO, Kristrom B and Gerdtham U-G (1993b) Willingness to pay for antihypertensive therapy--further results. *Journal of Health Economics* 12(1), 95-108.

Johannesson M and Jonsson B (1993). Economic evaluation of osteoporosis prevention. *Health Policy* 24(2), 103-24.

Johannesson M, Johansson PO and Jonsson B (1992) Economic evaluation of drug therapy: a review of the contingent valuation method. *Pharmacoeconomics* 1(5), 325-37.

Johannesson M (1992) Economics evaluation of lipid lowering: a feasibility test of the contingent valuation approach. *Health Policy* 20(3), 309-320.

Johannesson M, Jonsson B and Borgquist L (1991) Willingness to pay for antihypertensive therapy: results of a Swedish pilot study. *Journal of Health Economics* 10(4), 461-474.

Johnson FR, Desvouges WH, Ruby MC, Stieb D and De Civita P (1998) Eliciting health state preferences: an application to willingness to pay for longevity. *Medical Decision Making* 18 (suppl.), S57-S67.

Jones-Lee MW (1990) The value of transport safety. *Oxford Review of Economic Policy* 6:39-60.

Jones-Lee MW, Hammerton M and Philips PR (1985) The value of safety: Results of a National sample survey. *Economic Journal* 95, 49-72.

Kahneman D and Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47,263-91.

Kahneman D and Knetsch JL (1992) Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management* 22, 57-70.

Kaldor N (1939) Welfare propositions and interpersonal comparisons of utility. *Economic Journal* 49, 542-549.

Kanninen B (1995) Bias in discrete response contingent valuation. *Journal of Environmental Economics and Management* 28(1), 114-125.

Kartman B, StAlhammar N and Johannesson M (1997) Contingent valuation with an open-ended follow-up question: A test of scope effects. *Health Economics* 6, 637-639.

Kartman B, Andersson F and Johannesson M (1996) Willingness to pay for reductions in angina pectoris attacks. *Medical Decision Making* 16, 248-253.

Kealey MJ, Dovidio JF and Rockel ML (1988) Accuracy in Valuation is a matter of Degree. *Land Economics* 76, 158-71.

Kealey MJ, and Turner RW (1993) A test for the equality of close-ended and open-ended contingent valuations. *American Journal of Agricultural Economics* 75(2), 321-331.

Kealey MJ, Montgomery M and Dovidio JF (1990). Reliability and predictive validity of contingent valuation: Does the nature of the good matter? *Journal of Environmental Economics and Management* 19, 78-85

Kennedy P (1998) *A guide to econometrics.* 4th ed. Oxford: Blackwell publishers.

Kline P (1986) *A handbook of test construction.* Methuen; London.

Klose T (1999) The contingent valuation method in health care. *Health Policy* 47:97-123.

Kracmar JZ (1971) *Marketing research in the developing countries.* A Handbook. Praeger Publishers; New York.

Kristrom B (1990) A non-parametric approach to the estimation of welfare measures in discrete choice response valuation studies. *Land Economics* 66(2), 135-139.

Kroeger A, Meyer R, Mancheno M, Gonzalez M and Pesse K (1997) Operational aspects of bednet impregnation for community-based malaria control in Nicaragua, Ecuador, Peru and Colombia. *Tropical Medicine and International Health* 2(6), 589-602.

Krupnick AJ and Cropper ML (1992) The effect of information on health risk valuations. *Journal of Risk and Uncertainty* 5, 29-48.

Kuh JH, Stoll JR and Olson JK (1987) An economic assessment of the benefits of mosquito abatement in an organised mosquito control district. *Journal of the Americam Mosquito Control Association* 3(1), 8-13.

Lee S, Neumann P, Churchill W, Cannon M, Weinstein M and Johannesson M (1997) Patients' willingness to pay for autologous blood donation. *Health Policy* 40(1), 1-12.

Liljas B and Blumenschein K (2000). On hypothetical bias and calibration in cost-benefit studies. *Health Policy* 52, 53-70.

Lines J (1996) Mosquito nets and insecticides for net treatment: a discussion of existing and potential distribution systems in Africa. *Tropical Medicine and International Health* 1(5), 616-632.

Little IMD (1949) The foundations of welfare economics. *Oxford Economic Papers* 1:227-246. In Donaldson C (1993) *Theory and practice of willingness to pay for health care.* Health Economics Research Unit Discussion paper 01/93. Departments of Public Health and Economics, University of Aberdeen.

Little IMD (1957) *A critique of welfare economics*; Clarendon Press:London. In Donaldson C (1993) *Theory and practice of willingness to pay for health care.* Health Economics Research Unit Discussion paper 01/93. Departments of Public Health and Economics, University of Aberdeen.

Loehman ET, Park S and Boldt D (1994) Willingness to pay for gains and losses in visibility and health. *Land Economics* 70(4), 478-98.5

Loomis J, Brown T, Lucero B and Peterson G (1996) Improving validity experiments of contingent valuation methods: results of efforts to reduce the disparity of hypothetical and actual WTP. *Land Economics* 72, 450-61.

Loomis J, Brown T and Lucero B (1997) Evaluating the validity of the dichotomous choice question format in contingent valuation. *Environmental and Resource Economics* 10(2), 109-23.

Loomis JB (1989) Test-Retest reliability of the contingent valuation method: A comparison of general population and visitor responses. *American Journal of Agricultural Economics* 71, 76-84.

Lunander A (1998). Inducing incentives to understate and overstate willingness to pay within the open-ended and the dichotomous-choice elicitation formats: An experimental study. *Journal of Environmental Economics and Management* 35(1), 88-102.

Makemba AC, Winch PJ, Kamazima SR, Makame VR, Sengo F, Lubega PB, Minjas JN and Shiff CJ (1995) Community-based sale, distribution and insecticide impregnation of mosquito nets in Bagamoyo District, Tanzania. *Health Policy and Planning* 10(1), 50-59.

MacCormack CP and Snow RW (1986) Gambian cultural preferences in the use of insecticide-impregnanted bed nets. *Journal of Tropical Medicine and Hygiene* 89, 295-302.

Macmillan DC, Smart TS and Thorburn AP (1999) A field experiment involving cash and hypothetical charitable donations. *Environmental and Resource Economics* 14(3), 399-412.

Mathiyazhagan K (1998) Willingness to pay for rural health insurance through community participation in India. *Health Planning and Management* 13(1), 47-67.

Merewitz L and Sosnick SH (1971) *The budget's new clothes: a critique of planning-programme budgeting and benefit cost analysis*. Markham, Chicago.

McGuire A, Henderson J and Mooney G (1988) *The Economics of Health Care. An Introductory text.* London; Routledge and Keegan Paul.

Miedzybrodzka Z, Semper J, Shackley P, Abdalla M and Donaldson C (1995) Stepwise or couple antenatal carrier screening for cystic fibrosis?: women's preferences and willingness to pay. *Journal of Medical Genetics* 32, 282-283.

Miedzybrodzka Z, Shackley P, Donaldson C and Abdalla M (1994) Counting the benefits of screening: a pilot study of willingness to pay for antenatal carrier screening for cystic fibrosis. *Journal of Medical Screening* 1,82-83.

Mills A, Fox-Rushby J, Aikins M, D'Alessandro U, Cham K and Greenwood B (1994) Financing mechanisms for village activities in The Gambia and their implications for financing insecticide for bednet impregnation. *Journal of Tropical Medicine and Hygiene* 97(6), 325-332.

Mishan E J (1988) *Cost-Benefit Analysis.* (4th Edition) Unwin Hyman, London.

Mishan EJ (1976) The use of compensating and equivalent variations in cost-benefit analysis. *Economica* 43, 185-97.

Mishan EJ (1971) *Cost-benefit analysis:* An introduction. New York: Prager Publishers. Quoted in Brookshire et al (1980). Valuing increments and decrements in natural resource service flows. *American Journal of Agricultural Economics*, 478-488.

Mitchell RC and Carson RT (1981) *An experiment in determining willingness to pay national water quality improvements.* Draft report to the US Environmental Protection Agency, Washington DC.

Mitchell RC and Carson RT (1984) *A contingent valuation estimate of national freshwater benefits.* Technical report to the US Environmental Protection Agency, Washington DC.

Mitchell RC and Carson RT (1989) *Using Surveys to Value Public Goods: the Contingent Valuation Method.* Washington, D.C.: Resources for the Future.

Moatti JP, Auquier P, Le Coroller AG, Macquart-Moulin G (1995) QALYS or not QALYS: that is the question? *Rev Epidemiol Sante Publique* 43(6), 573-83.

Mooney G (1994) *Key issues in Health Economics.* Hertfordshire, Prentice Hall/Harvester Wheatsheaf.

Mooney G (1992) *Economics, Medicine and Health Care.* 2nd edition, Hertfordshire, Harvester Wheatsheaf.

Morrison GC (1998) Understanding the disparity between WTP and WTA: Endowment effect, Substitutability, or Imprecise preferences? *Economics Letters* 59(2), 189-194.

Morrison GC and Gyldmark M (1992) Appraising the use of contingent valuation. *Health Economics* 1, 233-243.

Mujinja PGM and Premji Z (1999) Willingness to pay for insecticide treated bed nets for malaria control: A case study of Bagamoyo bed-net project. *Book of Abstracts. MIM African malaria Conference*, Durban, South Africa; p.B-69.

Muller A and Reutzel T (1984) Willingness to pay for reduction in fatality risk: An exploratory survey. *American Journal of Public Health* 74, 808-812.

Munro BH (1997) *Statistical methods for health care research* (3rd edition). New York; Lippincott.

Musgrove RA (1959) *The theory of public finance-A study in public economy.* McGraw Hill:New York.

MVCU (2000) *Prevalence of malaria mobidity and mortality in Enugu state, 1995 to 1999.* Malaria and Vectors Control Unit (MVCU). Ministry of Health, Enugu, Nigeria.

Mwabu G and Mwangi W (1986). Health Care financing in Kenya: A simulation of welfare effects of user fees. *Social Science and Medicine* 22(7), 763-767.

Nahar S and Costello A (1998) The hidden cost of 'free' maternity care in Dhaka, Bangladesh. *Health Policy and Planning* 13(4), 417- 422.

National Oceanic and Atmospheric Administration (NOAA) (1993) Report of the NOAA panel on contingent valuation. *Federal Register* 58, 4602-4614.

National Population Commission (1999) *The 1998 National Census.* The National Population Commission, Abuja, Nigeria.

Neill H, Cummings RG, Ganderton PT, Harrison GW and McGuckin T (1994) Hypothetical surveys and real economic commitments. *Land Economics* 70, 145-54.

Neumann PJ and Johannesson M (1994) The willingness to pay for in vitro fertilisation: a pilot study using contingent valuation. *Medical Care* 32(7), 686-699.

Nevill CG, Some ES, Mung'ala VO, Mutemi W, New L, Marsh K, et al. (1996) Insecticide-treated bednets reduce mortality and severe morbidity from malaria among children on the Kenyan coast. *Tropical Medicine and International Health* 1, 139-148.

Norris T R (1952) *The theory of consumer's demand.* Yale University press: New Haven.

Nunnally JM and Bernstein IH (1994) *Psychometric Theory, 3rd edn.* New York: McGraw-Hill.

O'Brien B and Viramontes J (1994) Willingness to pay: a valid and reliable measure of health state preference? *Medical Decision Making* 14, 289-297.

O'Brien B, Novosel S, Torrance G and Streiner D (1995) Assessing the economic value of a new antidepressant: a willingness to pay approach. *Pharmacoeconomics* 8(1), 34-45.

O'Brien B and Gafni A (1996) When do the "dollars" make sense? Toward a conceptual framework for contingent valuation studies in health care. *Medical Decision Making* 16(3), 288-99.

O'Brien B, Goeree R, Gafni A, Torrance GW, Pauly MV, Erder H, Rusthoven J, Weeks J, Cahill M, and LaMont B (1998) Assessing the value of a new pharmaceutical: a feasibility study of contingent valuation in managed care. *Medical Care* 36(3), 370-384.

Ojediran B (1998) Odo-Oba: The village that feeds the multitude. *Sunday Sketch* 27, 8-9. (quoted in Balogun SK (1991) Bargaining strategy and sex as determinants of bargaining outcomes in open markets. *Nigerian Journal of Economic and Social Studies* 33(1), 31-38.

Okonkwo PO, Akpa A, Ihekwaba A, Nwagbo D, Umeh R, Adibua S, Ezike V and Ogbuokiri J (1991) Studies on onchocerciasis in forest-Savannah mosaic areas of Nigeria.1. Investigations in Gbaragu, Oji River. *Annals of Tropical Medicine and Parasitology* 85, 617-623.

Olsen JA and Smith RD (2001) Theory versus practice: A review of 'willingness to pay' in health and health care. *Health Economics 10, 39-52.*

Olsen J and Donaldson C (1998) Helicopters, hearts and hips: using willingness to pay to set priorities for public sector health care programmes. *Social Science and Medicine* 46, 1-12.

Olsen JA (1997) Aiding priority setting in health care: is there a role for the contingent valuation method. *Health Economics* 6, 603-612.

Onwujekwe OE, Chima R, Shu E, Nwagbo D and Okonkwo P (2001) Hypothetical and actual willingness to pay for Insecticide-treated nets: A study in five Nigerian communities. *Tropical Medicine and International Health* 6(7), 545-53.

Onwujekwe OE, Shu EN and Okonkwo PO (1999) Willingness to pay for the maintenance of equity in a local ivermectin distribution scheme in Toro, Northern Nigeria. *Public Health* 113, 193-194.

Onwujekwe OE, Shu EN, Nwagbo D, Akpala CO and Okonkwo PO (1998) Willingness to pay for community-based ivermectin distribution: A study of three onchocerciasis-endemic communities in Nigeria. *Tropical Medicine and International Health* 3(10), 802-808.

Onwujekwe OE (1999) Are insecticide-treated nets affordable? Relating costs of two sizes of nets and net re-treatment with basic household expenditures. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 93, 571-572.

Onwujekwe OE (2001) Searching for a better contingent valuation elicitation method in rural Nigeria. The Binary question with follow-up versus the bidding game technique. *Health Economics* 10, 147-158.

Onwujekwe O and Nwagbo D (2001) Investigating starting-point bias: A survey on willingness to pay for insecticide-treated nets. *Social Science and Medicine* (In press).

Pennie R, O'Conner A, Garvock M and Drake E (1991) Factors influencing the acceptance of Hepatitis B vaccine by students in health disciplines in Ottawa. *Canadian Journal of Public Health* 82, 12-15.

Perneger TV (1998). What's wrong with Bonferroni adjustments. *BMJ* 1998 316, 1236-1238.

Phelps CE and Newhouse J P (1974) Co-insurance, The price of time and demand for medical services. *The review of Economics and Statistics* 56, 334-342.

Phelps CE (1992) *Health Economics.* New York; HarperCollins Publishers Inc.

Philips PR, Russel IT and Jones-Lee MW (1989) The empirical estimation of individual valuation of safety: results of a national sample survey. In: Jones-Lee MW (Ed.). *The Economics of safety and physical risk.* Basil Blackwell, Oxford.

Phillips K, Homan R, Luft H, Hiatt P, Olson K, Keraney T and Heard S (1997) Willingness to pay for poison control centres. *Journal of Health Economics* 16, 343-357.

Phillips CV and Zeckhauser RJ (1989) Contingent valuation of damage to natural resources: How accurate? How appropriate? *Toxics Law Reporter* 4,520-529.

Premji Z, Lubega P, Hamisi Y, Mchopa E, Minjas J, Checkley W and Shiff C (1995) Changes in malaria-associated morbidity in children using insecticide treated mosquito nets in the Bagamoyo district of coastal Tanzania. *Tropical Medicine and Parasitology* 46, 147-153.

Ramsey SD, Sullivan SD, Psaty BM and Patrick D L (1997). Willingness to pay for antihypertensive care: evidence from a staff model HMO. *Social Science and Medicine* 44(12), 1911-1917.

Reiling SD, Boyle KJ, Philips ML and Anderson MW (1990) Temporal reliability of contingent values. *Land Economics* 66(2), 128-134.

Roemer M I (1961) Bed supply and hospital utilisation: A natural experiment. *Hospitals* 35,36-42.

Ross C, Steward C and Sinacore J (1995) A comparative study of seven measures of patient satisfaction. *Medical Care* 33(4), 392-406.

Rowe RD, D'Arge RC and Brokshire DS (1980) An experiment in the economic value of visibility. *Journal of Environmental Economics and Management,* 7:1-19.

Rowley J, Cham B and Pinder M (1999) Availability and affordability of insecticide for treating bednets in The Gambia. *Book of Abstracts. MIM African malaria Conference,* Durban, South Africa, p.B-85.

Russell S, Fox-Rushby J, and Arhin D (1995) Willingness and ability to pay for health care: A selection of methods and issues. *Health Policy and Planning* 10(1), 94-101.

Ryan M, Scott DA and Donaldson C (1999) *Valuing health care using willingness to pay: a comparison of the payment card and dichotomous choice methods.* Paper presented at the 2nd International Health Economics Association conference, Rotterdam.

Ryan M, Ratcliffe J and Tucker J (1997) Using willingness to pay to value alternative models of antenatal care. *Social Science and Medicine* 44(3), 371-380.

Ryan M (1997) Should government fund assisted reproductive techniques? A pilot study using willingness to pay. *Applied Economics* 29(7), 841-849.

Ryan M (1996) Using willingness to pay to assess the benefits of assisted reproductive techniques. *Health Economics* 5(6), 543-558.

Ryan M (1993) *Valuing the Benefits of Health Care: Conjoint Analysis or Contingent Valuation?* HERU Discussion Paper 02/93. Health Economics Research Unit, Department of Public Health and Economics, University of Aberdeen, 1993.

Sassi F, Archard L and Le Grand J (2001). Equity and the economic evaluation of healthcare. *Health Technology Assessment* 5 (3).

Scitvosky T (1941) A note on welfare propositions in economics. *Review of Economic Studies* 9, 77-88.

Seip K and Strand J (1992) Willingness to pay for environmental goods in Norway: A contingent valuation study with Real payment. *Environmental and Resource Economics* 2(1), 91-106.

Shain M and Roemer MI (1959) Hospital costs relate to the supply of beds. *Modern Hospital* 92, 71-73.

Smith RD (2000) The discrete-choice willingness to pay question format in Health economics: Should we adopt environmental guidelines? *Med Decis Making* 20, 194-206.

Smith RD (2001) The relative sensitivity of willingness to pay and time trade-off to changes in health status: An empirical investigation. *Health Economics* 10, 487-497.

Snow RW, McCabe E, Mbogo CNM, Molyneux CS, Some ES, Mung'ala VO and Nevill CG (1999) The effect of delivery mechanisms on the uptake of bed net re-impregnation in Kilifi District, Kenya. *Health Policy and Planning* 14(1), 18-25.

Sorum PC (1999) Measuring patient preferences by willingness to pay to avoid: The case of otitis media. *Medical Decision Making* 19(1), 27-37.

Stalhammar N (1996) An empirical note on willingness to pay and starting point bias. *Medical Decision Making* 16, 242-247.

STATA *manual version 6* (1985-1999) Texas. Stata Press.

Streiner DL and Norman GR (1995) *Health measurement scales. Practical Guide to their Development and Use.* Second edition. New York; Oxford University Press.

Tambour M and Zethraeus N (1998) Nonparametric willingness-to-pay measures and confidence statements. *Medical Decision Making* 18(3), 330-6.

Teisl MF, Boyle KJ, McCollum DW and Reiling SD (1995) Test-retest reliability of contingent valuation with independent sample pretest and posttest control groups. *American Journal of Agricultural Economics* 77, 613-619.

Thompson M, Read J and Laing M (1984) Feasibility of willingness to pay measurement in chronic arthritis. *Medical Decision Making* 4, 195-215.

Thompson M (1986) Willingness to pay and accept to cure chronic disease. *American Journal of Public Health* 76, 392-396.

Tobin J (1958) Estimation of relationships for limited dependent variables. *Econometrica* 21, 31-40.

Weaver M, Ndamobissi R, Kornfield R, Blewane C, Sathe A, Chapko M, Bendje N, Nguembi E and Senwara-Defiobona J (1996) Willingness to pay for child survival: results of a national survey in a central African Republic. *Social Science and Medicine* 43, 985-998.

Walraven G (1996) Willingness to Pay for District Hosptial Services in Rural Tanzania. *Health Policy and Planning* 11(4), 428-437.

Whitehead JC and Hoban TJ (1999) Temporal reliability in contingent valuation. *Land economics* 75(3), 453-465.

Whittington D, Pinheiro AC and Cropper M (1996) *The economic benefits of malaria control: A contingent valuation study in Marracuene, Mozambique.* Department of Environmental Sciences and Engineering and Citry Regional Planning, University of North Carolina at Chapel Hill.

Whittington D, Smith K, Okoroafor A, Liu JL and McPhail A (1992) Giving respondents time to think in contingent valuation studies: *Journal of environmental economics and management 22(3), 205-225.*

Whittington D, Luria DT, Wright AM, Choe K, Hughes JA, and Swarna V (1993) Household Demand for Improved Sanitation Services in Kumasi, Ghana. *Water Resources Research* 29(6), 1539-60.

Whittington D, Lauria DT and Mu X (1991) A study on water vending and willingness to pay for water in Onitsha, Nigeria. *World Development* 19(2-3), 179-198.

Whittington D, Briscoe J, Mu X and Barron W (1990) Estimating the Willingness to Pay for Water Services in Developing Countries: a Case Study of the Use of Contingent Valuation Surveys in Southern Haiti. *Economic Development and Cultural Change* 38(2), 293-312.

World Bank Water Demand Research Team (1993) The demand for water in rural areas: determinants and policy implications. *World Bank Research Observer* 8(1), 47-70.

Zethraeus N (1998) Willingness to pay for hormone replacement therapy. *Health Economics* 7, 31-38.

ANNEXES

Appendix 1: Studies that have examined construct validity

Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Abel-Smith and Rawal (1992)	User charges in public health facilities	Interviewer-administered Qs; close-ended	Not stated	None	Market research	nil	nil	Nil
AMREF (1997)	Bed-nets	Focus groups; open ended	Not stated	None	Market research	nil	nil	Nil
Alberini and Krupnick (1998)	Improved air quality	Questionnaires; dichotomous choice with follow-up bidding	Not stated	Log linear regression	Benefit	Duration, number of symptoms, episode is a cold, age, household income, education, respondent has had a serious respiratory illness, respondent has other non-respiratory chronic illness, respondent has health insurance and number of sick days available to respondent.	Duration, number of symptoms, episode is a cold, household income, education, respondent has had a serious respiratory illness, respondent has other non-respiratory chronic illness, respondent has health insurance and number of sick days available to respondent.	Nil
Arhin (1995)	Health insurance	Interviewer-administered Qs; open-ended	Not stated	None	Market research	nil	Nil	Nil
Armstrong Schellenberg et al (1999)	ITNs and the treatment	Not stated	Not stated	None	Market research	nil	nil	Nil
Asenso-Okyere et al (1997)	Health Insurance	FGDs and Interviwer-administered Qs; bidding game	Preferred by the respondents from FGDs and in-depth interviews	Ordered probit	Market research	Premium, other available premiums, income, last month's expenditure on health care, occupation, household dependency ratio, services offered, difficulty in paying for health care, education level, benefits available, current level of health	Sex, household dependency ratio and income, difficulty in paying for health care and last month's expenditure on health care.	Nil

						expenditures, sex, age, marital status and years of schooling		
Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Asgary and Willis (1997)	Construction measures to mitigate earthquake crisis	Self-administered Qs; open-ended	To provide a realistic elicitation procedure	Linear regression and similarity to another study in Iran	Benefit	Levels of earthquake hazard perception, levels of earthquake risk perception, fatalistic attitude, household monthly expenditure, family size, education, age, household decision to move, earthquake experience, knowledge about earthquakes	levels of earthquake risk perception, education and age.	Nil
Appel L J et al (1990)	Risk reduction from low osmolality contrast media	Interviewer-administered Qs; bidding	Patients had problem answering open ended questions in the pilot phase	Generalised linear interactive modelling	Benefit	Income, education, insurance, age, sex, race, mental status and prior experience.	None	Nil
Bala et al (1998)	Shignles control and to compare WTP with QALYs	Questionnaire:Double-bounded binary method	Its better than open-ended	Probit model	Benefit	Prior disease experience, present disease status, age, gender, household size, monthly income, WTP dollar amount	Gender (+) Prior disease experience (-)	Nil
Berger et al (1987)	Health risks	Self and mailed Qs; open ended	Not stated	Comparison with other measures	Benefit	nil	Nil	Nil
Berwick and Weinstein (1985)	Ultrasound in normal pregnancy	Interviewer-administered Qs; open-ended	Not stated	Linear regressions	Benefit	Age, household income and specific ultrasound & pregnancy variables	Did not state	Nil
Binka and Adongo (1997)	ITNs	Interviewer-administered Qs and FGD; No elicitation method	Nil	Nil	Benefit	nil	Nil	Nil

		shown						
Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Birdsall and Chuhan (1983)	Village health worker and water	Interviewer-administered Qs; open-ended	Not stated	OLS regression and convergent validity with indirect WTP measurement	Market research	Household income, number of literate adults, number of persons in school, age group of sick baby, number of sick persons, distance to nearest dispensary, distance to nearest drug outlet, quality of nearest dispensary, quality of nearest drug outlet, dispensary and drug outlet at common site, area of residence.	None	Nil
Boadu (1992)	Water	Interviewer-administered Qs; bidding game	To mimic local markets	OLS multiple regression analysis	Market research	Age, sex, education, household size, income, expenditures, distance from existing source, drought, illness and ownership.	Age, sex, education, household size, income, drought, illness and ownership.	Yes: Used Breusch and Pagan (1979) to check for heterosceda sticity and condition indexes & variance inflation to check for multicolline arity
Brieger et al (1996)	ITNs	Qusetionnaires: open-ended	Not stated	Not stated	Market research	Assets, status in household, education, previous net purchase	Sex (+), education (+) and previous net purchase (+)	Nil
Brown and Green (1981)	Increased safety	Questionnaire; open ended	Not stated	Not stated	Benefit	nil	Nil	Nil
Clarke (2000)	Improved access to mobile mammographic screening units	Telephone: Single & Double-bounded binary method	In keeping with many recent CV studies	Multivariate probit	Benefit	Income, bid amount, distance, intended use, KAP, marital status, education, age	Bid (-), distance (+), intended use (+)	Likelihood ratio index test for goodness of fit
Clarke (2001)	Mammographic screening	Telephone: Single bounded Binary method	Better than double bounded binary method	Probit and convergent validity	Benefit	Income, bid amt, distance, intended use, GP advice, screened before	Bid (-), access (+), GP advice (+)	Likelihood ratio index test for goodness of fit

Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Culbertson et al (1988)	Medication information	Self-administered Qs; open ended	Not stated	None	Benefit	nil	Nil	Nil
Dalmau-Matarrodona (2001)	Home care services in day care surgery	Questionnaire:Binary with follow-up	Recommended and mimics price taking	Tobit and Double-Hurdle methods	Benefit	Age, sex, employment, income, education, household size	Age, sex, employment, income, education, household size	Nil
Davey et al (1998)	Different insulin therapies	Interviewer-administered Qs; started with open ended to lead on to bidding	To increase realism to resemble current payment mechanism	None	Benefit	Type of diabetes, study centre, education, household income, current monthly spending on diabetes, satisfaction with current therapy, age and sex.	Age, type of diabetes, education and income.	Nil
Diez (1998)	Reduction of post-operative emesis	Interviewer-administered Qs; bidding	Respondents found the open ended difficult and a large sample size is needed for the binary method	Spearman correlation coefficients between WTP and explanatory variables	Benefit	Household income, degree of worry, education level, child had Presence of nausea and vomiting, number of children, number of operations and difficulty with WTP question.	Household income	Nil
Donaldson (1990)	Publicly-provided health care services	Interviewer-administered Qs; open-ended and bidding game	Bidding game to help those unable to answer the open-ended	None	Benefit	nil	nil	Nil
Donaldson et al (1996)	Avoiding food-borne risk	Mailed Qs; open ended	Not stated	Tobit analysis and asymptotic F test	Benefit	Sex, age, self perceived health, income previous illness, preference, education and income	Previous illness and preference	Nil
Donaldson and Shackley (1997)	Laparscopic cholecystectomy	Mailed Qs; payment card	Not stated	OLS multiple regression	Benefit	Age, marital status, social class, education level, self-perceived health, distance of residence and	Social class (though wrong sign) and income.	Yes: Residual plot and RESET test

						previous surgical knowledge		
Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Donaldson et al (1997)	Child health services	Interviewer-administered Qs; open-ended and bidding	Not stated	Tobit analysis	Benefit	Childage, sex, age, marital status, social class, previous work experience, length of interview, education, location, previous surgical experience, number of children and household income.	1. WTP for surgery Marital status, location B and social class A. 2. WTP for outpatient clinics number of children, locations A & B and social classes A & B.	Yes: mentioned homoscedasticity, normal distribution, correct functional form and independent predictor variables but did not present the result
Donaldson et al (1997)	Bone mineral screening	Self-administered Qs; open-ended and payment scale	To compare the two methods	Ordinary least squares regression analysis	Benefit	Age, car ownership, social class, education level, experienced falls, experienced flushes, height, experienced hysterectomy, ever smoked	Social class	Yes: RESET and test for heterosceda sticity, though tests not presented
Donaldson et al (1997)	Carrier screening for cystic fibrosis	Self-administered Qs; open-ended	Not stated	Compared results to a previous study & OLS regression .	Benefit	Age, marital status, education, home ownership, number of pregnancy weeks, number of children in the family, family history of cystic fibrosis and social class of main household income earner	Social class and single mothers	Nil
Donaldson et al (1998)	Maternity care	Mailed Qs; open-ended and payment scale	To compare the two methods and determine econometric models	OLS, Tobit, Heckman and Grouped regression	Benefit	Preferred care, age, marital status, had a baby before, information, who is main income earner, social classes, income classes, educational level, self-perceived health and	Age, sex	Yes: RESET, Normality and test for heterosceda sticity.

						reminder.		
Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Eckerlund et al (1995)	Health care	Telephone; binary method	Method close to market realism	Logistic regression	Benefit and market	Household income, age and sex	Household income	Nil
Einarson et al (1988)	Serum cholesterol monitoring service	Self-administered Qs; binary and payment card	Not stated	Comparison of WTP with attribute scores	Market research and benefit	nil	Nil	Nil
Feilden R et al (1996)	Hepatitis B vaccine	Interviewer-administered Qs; binary	Not stated	Comparison of WTP with ATP	Market research			Nil
Golan and Schecter (1993)	Supplemental health care	Interviewer-administered Qs; bidding game	Common in environmental goods studies	Pre-testing and checking for systematic biases and logit regression	Market research and benefit	Age, belief family size, income, socio-economic group, employment, neighbourhood, private medical expenses and education level.	Age, income, neighbourhood, private medical expenses and education level.	Nil
Gonzalez-Caban and Loomis (1997)	Ecological integrity	Interviewer-administered Qs; bidding technique to get close-ended answers	Recommended by NOAA	Multivariate logit	Benefit	Region, education and bid amount	Region, education and bid amount	Nil
Gore and Madhavan (1994)	Pharmacist counselling	Mailed Qs; WTP for ranges	Not stated	None	Benefit	nil	nil	Nil
Gyldmark and Morrison (2001)	Private insurance for four different health-care goods	Questionnaire: open-ended & payment card	It presents more information than binary questions	Heckman and OLS, but used log OLS	Benefit	Income groups, years in school, residence, age, marital status, occupation	Some income groups and school years	Normality and homoscedasticity
Hongvivatana and Manopimoke (1991)	Health Insurance	Interviewer-administered Qs; close ended	Not stated	Not stated	Market research	Not stated	Not stated	Nil
Johannesson et al (1991)	Anti hyper-tensive therapy	Mailed Qs; open-ended and discrete choice	To compare them	Logistic regression	Benefit	Income, age, sex, substitute and risk reduction	nil	% predictions
Johannesson (1992)	Lipid lowering	Mailed Qs; payment scale and binary	Not stated	None done	Benefit	nil	nil	Nil
Johannesson et al (1993)	Anti hyper-tensive therapy	Mailed Qs: Binary	To test a calibrated elicitation system	Logistic regression	Benefit	Income, education, age, sex and household size	Perceived changes in health status, price.	Yes: % correct predictions and Likelihood

								ratio index test
Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Johannesson et al (1993)	Lipid lowering	Self-adminQs; combination of open ended and binary systems	Not stated	Linear regression	Benefit	Perceived health gain, age, sex, household size, household income and education level	Perceived health gain, income, sex and education level.	Nil
Johannesson and Johansson (1996)	Increased life expectancy	Telephone; binary	NOAA recommendation	Logistic regression	Benefit	Age, bid, income, education and sex.	Bid, age, education and income	Yes: % correct predictions and Likelihood ratio index test
Johannesson and Johansson (1997)	Increased life expectancy at an advanced age	Telephone; binary	Not stated	Logistic regression = ordered and simple	Benefit	Age, bid, income, education, quality of life scale and sex.	Bid (-), income (+), quality of life scale (+)	Yes: % correct predictions and Likelihood ratio index test
Johnson et al (1998)	Longevity	Computer-based self-administered Qs; Binary	Not stated	Ordered probit	Benefit	Longevity SP pairs, age, education, sex, symptoms.	Education, some SP states.	Nil
Jones-Lee et al (1985)	Safety	Interviewer-administered Qs; guided open ended with bidding	It was better than payment card and bidding game	Wilcoxon matched pairs and multiple regression analysis	Benefit	Household income, age, social class, miles driven last year, car ownership, accident experience, missing mileage and missing income.	Age, household income and social class	Nil
Kartmann et al (1997)	Treatment of reflux oesophagitis	Questionnaire: binary with open-ended follow up	To test whether the follow-up is sensitive to scope effect	Probit and OLS regression analyses	Benefit	Age, gender, income, severe pain, bid, scope variable, second place and third place in sequence	Income, bid and scope variable	Nil
Kroeger et al (1997)	ITN	Questionnaire; open ended	Nil	Nil	Market research	nil	nil	Nil
Krupnick and Cropper (1992)	Cure for chronic lung disease	Self-administered Qs; ?open-ended	Not stated	linear regression	Benefit	Sex, marital status, household size, number of kids, education level, income, age, smoke, insurance, work, exercise etc.	Sex, number of kids and smoke.	Nil

Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Kuh et al (1987)	Mosquito control	Mailed Qs; payment card	Not stated	OLS multiple regression	Benefit	Household income, number of years residence, education level, household evaluation of effectiveness of the program, the size of current household tax contributions for mosquito control, concern for non-chemical control measures, expenditures for self control measures etc	Household income, education level, household evaluation of effectiveness of the program, the size of current household tax contributions for mosquito control, concern for non-chemical control measures, expenditures for self control measures	Nil
Lee et al (1997)	Autologous blood transfusion	Self-administered Qs; Binary method	Not stated	Logistic regression	Benefit	Age, sex, education, household size, household income, price listed, dread of blood transfusion, perceived risks.	Stated price, the dread variable and income.	Nil
Loehman et al (1994)	Visibility and Health	Interviewer-administered Qs; payment card	Not stated	Logistic regression	Benefit	Perceptions, income, urban, market, health status and information	Income, market, health status and information.	Nil
MacCormack and Snow (1986)	Insecticide treated nets	Interviewer-administered; open ended	Not stated	Nil	Market research	nil	Nil	Nil
Mathiyazhagan (1998)	Rural health insurance	Interviewer-administered Qs; open ended	Not stated	Logistic regression	Market research	Age, family size, caste, health condition, number of hospital episodes, number of working days lost due to ill-health, number of times doctor consulted, source of health care service	Family size, ST/SC caste, illness, three or more times of hospital episodes, number of working days lost due to ill-health, annual income, income flow characteristi	Nil

						utilised, annual income, income flow characteristics and occupational status.	cs and occupationa l class	
Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Miedzybrodzka et al (1995)	Carrier screening for cystic fibrosis	Questionnaire; open ended	Not stated	Linear regression	Benefit	Age, marital status, age at leaving full time education, home ownership, social class, gestation, number of children and family history of CF	Social class	Nil
Mujinja and Premji (1999)	Insecticide-treated nets	Not stated	Not stated	Logistic regression	Market research	Not stated	Household income and household size.	Nil
Muller and Reutzel (1984)	Reduction in fatality risk	Mailed Qs; open ended	Not stated	OLS multiple regression	Benefit	Income, car ownership, accident experience, exposure to potential accidents, risk aversion and mathematical ability.	Nil	Nil
Mills et al (1994)	Insecticides	Interviewer-administered Qs; open ended	Not stated	Kruskal-Wallis test	Market research	Not stated	Area	Nil
Nahar and Costello (1998)	Maternity services	Interviewer-administered Q; WTP yes or no rating	Not stated	Not stated	Market research	nil	nil	Nil
Neumann and Johannesson (1994)	In vitro fertilisation (IVF)	Self-administered Qs; close ended and bidding	Shown to work well with respect to response rate & acceptability	OLS with natural log of the variables	Benefit	Age, sex, marital status, number of children, current and expected household income, desire more children, inclination to use IVF, etc	Desire for more children, inclination to use IVF, expected income, education level and sex.	Nil
O'Brien and Viramontes (1994)	Cure for chronic lung disease	Interviewer-administered Qs; bidding game	Has more market realism than open-ended Qs & to investigate starting point bias	ANCOVA, Intra-clan correlation coefficient and multivariate analysis	Benefit	Age, sex, education level, income and various health state measures	nil	Nil

Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
O'Brien et al (1998)	Value of a new phramace-tical	Questionnaire:Double-Bidding game	Bidding games as search algorithms are better than single –question binary	ANCOVA and scope tests	Benefit	Age, sex, education, health status, annual pre-tax household income	nil	Nil No starting point bias WTA > WTP
Olsen and Donaldson (1998)	Different health care services	Interviewer-administered Qs; payment card (extra-earmarked taxation)	Resembled how respondents pay for health care services	Stated WTP for one service compared to what people are already paying	Benefit	Sex, education level, residence, own health, income, age, sex, agesq and time for interview	Sex and education level	Yes: Tested using probability and residuals vs fitted WTP plots
Onwujekwe et al (1998)	Ivermectin distribution	Interviewer-administered Q; bidding	Not stated	OLS regression	Market research	Expenditure on health care, expenditure on food, type of property owned, income class, type of saving scheme, occupation, education, marital status, age group, presence of illness, priority ranking of the disease, and perceived risk of contracting the disease.	Expenditure on health care, type of property owned, , type of saving scheme, education and presence of illness,	Nil
Onwujekwe et al (1999)	Altruism	Interviewer-administered Qs; open-ended	Not stated	OLS regression	Market research	Socio-economic variables	nil	Nil
Onwujekwe et al (2001)	ITNs re-treatment	Qusetionnaires: bidding game and BWFU	Mimicked price-taking	OLS	Market research	Residences, sex, priority ranking of malaria, WTP for own nets, exp on malaria, food expenditures, savings, celebrate.	Residences, sex (+), WTP for own nets (+), food expenditures (+), savings (+).	Nil
Onwujekwe et al (2001)	Altruism for ITNs	Qusetionnaires: open-ended	Mimicked how people give alms	Tobit	Market research	Residences, sex, no of residents, marital status, malaria incidence, WTP for own nets, expenditure on malaria,	Residence, sex (+), marital status (-), WTP for own nets (+), expenditure on malaria (+), savings	Nil

						expenditures, savings.	(+).	
Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Persson et al (1994)	Reduction of non-fatal injuries	Questionnaire; ?open-ended	Not stated	Linear regress-ion	Benefit	Income, age, distance, exposure an risk.	Risk	Nil
Pennie et al (1991)	Hepatitis B vaccine	Self-administered Qs; Binary	Not stated	Multiple regression	Market research	Did not state	Discipline of study, perceived risk if not immunised, and willingness to pay (the dependent variable was intentions).	Nil
Phillips et al (1997)	Poison control centers	Telephone: Bidding game	Thought it was recommended by NOAA	OLS	Benefit	Use, education, income, sex, age, household size, perception,	Females (+), younger people (+), income (+)	Nil
Ramsey et al (1997)	Anti hyper-tensive care	Mailed Qs; payment scale with embedded binary system	To simulate a bidding game	OLS multiple regression	Benefit	Quality of life rating, household income, education level, age, sex, marital status and current perceived health status rating	**WTP1:** Income & CRS **WTP2:**CRS, Income and age **WTP3:**CRS, Income and gender	Nil
Ryan et al (1997)	Alternative models of antenatal care	Mailed Qs; binary elicitation method	It reflects the way individuals buy things everyday & recommended by NOAA (1993)	Logistic regression	Benefit	Bid level, arm of the trial, type of care, age, number of antenatal appointments, satisfaction, perceived risk of health problems, perceived risk of pregnancy and income.	Bid level, age, satisfaction and income.	% predictions
Ryan et al (1999)	Three health programs	Questionnaire:payment card & binary	To compare them	Logit and ordered logit	Benefit	Age, education, gender, household size, income, perceived health status.	Did not show	Nil
Ross et al (1995)	Patient satisfaction	Self-administered Qs; open ended	Not stated	Not done	Benefit	nil	nil	Nil
Smith (2001)	Various levels of health care	Questionniare: open-ended	Most resembled the comparative TTO design	Discriminant validity	Benefit / utility	Nil	Nil	Nil

Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Sorum (1999)	Prevention of otitis media	Self-administered Qs; open-ended	Better than binary and bidding game in the study context	None	Benefit	nil	Nil	Nil
Tambour and Zethraeus (1998)	Hormone replacement therapy	Interviewer-administered Qs; binary	Used widely and recommended by NOAA (1993)	Convergent validity with rating scale and time-trade off technique	Benefit	nil	nil	Nil
Thompson (1986)	Cure chronic diseases	Interviewer-administered Qs; Open ended in %	Not stated	OLS regression	Benefit	31 independent variables: costs of drugs taken, personal investment income, total household income etc.	Age, investment income etc	Nil
Weaver et al (1996)	Child survival	Questionnaire; binary	Market realism	Logit model	Market research	Price, income	Price (-), income (+)	Nil
Zethraeus (1998)	Hormone replacement therapy	Interviewer-administered Qs; binary	To be able to estimate a demand function	Logistic regression	Benefit	Household income, the bid, education level, age and household size	The bid	Yes: % correct predictions and Likelihood ratio index test
Walvaren (1996)	Hospital services	Interviewer-administered Qs; open ended	Not stated	None	Market research	nil	nil	Nil
Weaver et al (1996)	Improved health facilities	Interviewer-administered Qs; close-ended	Not stated	Logistic analysis	Market research			Nil
Whittington et al (1990)	Water	Interviewer-administered Qs; bidding	Not stated	Ordered probit	Market research	Household wealth index, household with foreign income, occupation index, household educational level, distance from existing source, quality index of existing source, sex.	Household wealth index, household educational level, distance from existing source, quality index of existing source and sex.	Nil
Whittington et al (1991)	Water	Interviewer-administered Qs;	Not stated	OLS multiple regression	Market research	Slum, toilet, water storage, small trader,	Slum, toilet, small trader, retail vendor	Nil

		bidding		analysis and ordered probit		number of years of schooling, retail vendor and household income.	and houschold income.	
Authors	WTP for what?	Survey design; question format;	Rationale for the question format	Validity measures	Benefit or market research	Variables considered in validity tests	Statistically significant variables	Diagnostic tests
Whittington et al (1992)	Water	Interviewer-administered Qs; bidding	Not stated	OLS multiple regression analysis and ordered probit	Market research	Time-to-think, starting value of iterative bidding, head, sex, attitude towards public provision, talk to people about bids, asset measure, education, family size, fraction adult women, # of children, farming, expenditures on water, storage capacity, housing type and residence.	Not stated	Nil
Whittington et al (1992)	Sanitation services	Interviewer-administered Qs; bidding	Not stated	OLS multiple regression, Stewart's max. likelihood and ordered probit	Market research	Starting value of iterative bidding, time to think, sex, age, household head, owner of house, religion, knowledge, household income, wealth, years of education, trader, office worker.	Starting value of iterative bidding, owner of house, household income, years of education, trader.	Nil
Whittington et al (1996)	Hypothet-ical malaria vaccine	Interviewer-administered Qs; close ended	Not stated	Probit analysis	Benefit	Price, wealth, private connection, price x wealth, had malaria before, source, rooms, male, education, occupation and age.	Wealth, price x wealth and private connection	Nil
World Bank Water Demand Research Team (1993)	Water	Not stated	Not stated	Not stated	Market research	nil	nil	Nil

Appendix 2. Articles that compared stated and actual willingness to pay

Authors	WTP for what?	Compare what	Survey design; question format	Methodology of the validity measure	Divergence between hypothetical and real WTP
Blumenschein et al (1998)	Pair of sunglasses	Dichotomous choice contingent valuation method with real purchase decisions	Self-administered questionnaires; dichotomous questions	Used two different groups & tested differences in between-groups and in within-groups; calibration	No difference in the in-between groups using the calibration
Blumenschein et al (2001)	Pharmacist provided programme	Calibrated/Certainty of hypothetical WTP with real purchase decisions	Self/interviewer administered questionnaires; dichotomous questions	Used two different groups & tested differences in between-groups; calibration	Divergences without calibration/certainty; no divergences with calibration/certainty
Cummings et al (1995)	Juice-maker, chocolate and calculator	Hypothetical and actual dichotomous WTP	Self-administered Qs; dichotomous choice	Used two different samples; within and between groups).	Divergence between hypothetical and actual WTP – no ordering effect
Dgedge et al (1999)	Insecticide-treated nets	Hypothetical and actual WTP	Interviewer-administered Qs; bidding game	Re-interviewed the respondents two years after programme implementation	Moderate to severe divergence
Dickie et al (1987)	Strawberries	Demand relations between hypothetical and actual WTP	Interviewer-administered Qs; binary	Used two different groups; one group stated WTP only, while the other both stated and purchased the strawberries	No divergence noted between hypothetical and actual WTP
Foster V et al (1997)	Protection of birds	Real and hypothetical WTP	Mailed Qs; open ended bidding method	Compared real and hypothetical from different groups	Divergence noted; hypothetical 4X the real WTP
Frykblom (1997)	An environ-mental atlas	Hypothetical and real validity of close-ended and open-ended WTP elicitation methods	Classroom sessions: self-administered Qs; open ended and close ended	Used three different samples for the three different question modes	Close ended and open ended were similar, but both diverged from real WTP
Johannesson et al (1997)	Box of Belgian chocolates	Compare results with that from an earlier similar study	Room session; open-ended	Compared real and hypothetical WTP from two different groups	N0 divergence noted
Johannesson (1998)	Box of Belgian chocolates	Dichotomous choice contingent valuation method with real purchase decisions and test the effect of calibration on WTP	Self-administered questionnaires; dichotomous questions	Used two different groups & tested differences in between-groups and in within-groups; immediate purchases	Hypothetical yes overestimates real yes, and absolutely yes underestimates real yes; thus both of them provided a range for real WTP
Johannesson (1997)	Box of Belgian chocolates	Compare results with that from an earlier similar study	Room session; open-ended	Compared real and hypothetical WTP from two different groups	Divergence noted
Johannesson et al (1999)	Box of chocolates and pair of sunglasses	Calibrated hypothetical WTP with real purchase decisions	Self-administered questionnaires; dichotomous questions	Used two different groups & tested differences in between-groups and in within-groups; calibration	No divergence between the two measures
Loomis J et al (1996)	An art print	Real and hypothetical WTP	Classroom sessions: self-administered Qs; open ended	Real WTP in form of sealed bid auction	Divergence noted between Hypothetical . real WTP. However, reminding respondents to be honest improved the correlation
Authors	WTP for what?	Compare what	Survey design; question format	Methodology of the validity measure	Divergence between hypothetical and real WTP

Loomis et al (1997)	An art print	Compare dichotomous choice contingent valuation with real purchase decisions and with hypothetical & real open ended technique	Room sessions, dichotomous choice and open ended using four different samples for each elicitation	Used different statistical tests to compare hypothetical and actual WTP, and the two different elicitation methods	Some divergence noted, but none between dichotomous choice and open ended question format
Neill et al (1994)	An art print and a map	Real and hypothetical WTP	Classroom sessions: self-administered Qs; open ended	Real WTP in form of sealed bid auction	Divergence noted between Hypothetical . real WTP.
Onwujekwe et al (1999)	Insecticide-treated nets	Hypothetical and actual WTP	Interviewer-administered Qs; bidding game	Sold the nets to a sub-sample one month after eliciting hypothetical WTP	Mild divergence
Onwujekwe (2001)	ITNs	Determine better elicitation method	Interviewer-administered questionnaires; Bidding game and binary with follow up	Positive predictive validity and overall predictions from a sub-sample of the original respondents after 1 month of determining stated WTP	No difference between the two elicitation method, and no significant divergence between stated and actual WTP.
Rowley et al (1999)	Insecticide impregnation	Hypothetical and actual WTP	Interviewer-administered Qs	Re-interviewed a sub-sample some months after programme implementation	Mild to moderate divergence

Appendix 3: Reliability:

Reference	What good	Survey design; question format	Reliability approach	Correlation and other statistics	Conclusion
Desvousges et al (1983)	Natural resource damage	Questionnaires: Binary	Intra-analyst reliability: Regression analyses	Nil	WTP on non-use values were not valid
Carson et al (1997)	Protection of a sea	Questionnaire: Double-bounded binary	Surveyed diff. people 2 years apart: Chi-square and regression	Chi-square ($p>0.05$)	WTP was reliable
Golan and Shecter (1993)	Supplemental health care	Questionnaire: Bidding game	Checked general fit of the regression model	Nil	WTP was reliable
Kealey et al (1988)	Chocolate	Questionnaire: Binary and Open-ended (B & OE)	Test-retest 2 weeks apart: Paired t-tests and regression	t-tests ($p>0.05$)	WTP and B?OE were reliable
Loomis (1989)	Protection of Mono Lake	Questionnaire: open-ended	Test-retest 9 months apart: paired t-tests, correlation coefficients and regression	$r = 0.55, 0.42$ $r = 0.55, 0.78$ t-tests ($p>0.05$)	WTP was reliable
O'Brien and Viramontes (1994)	Health state preference	Questionnaire: Bidding game	Test retest reliability: Intra-class correlation using 20 people	$r = 0.66$	WTP was reliable
Reiling et al (1990)	Controlling late-season black flies	Questionnaire: open-ended	Surveyed diff. People 3 months apart; Paired t-tests	t-tests ($p>0.05$)	WTP was reliable
Stevens et al (1994)	Bald eagles and wild turkeys	Questionnaire: Binary	Test-retest 3 years apart: Paired t-tests, Correlation coefficient, regression	$r = 0.30$ t-tests ($p>0.05$)	WTP was reliable
Sorum (1999)	Treatment for acute otitis media	Qusetionnaire: Open-ended	Test-retest 6 months apart with 39 people: Spearman's rank correlation and non-parametric tests	$r = 0.21 - 0.69$	Reliable for groups but not for individuals
Teisl et al (1995)	Moose hunting	Questionnaire: Binary	Test-retest 5 months apart: Cramer's V correlation, Chi-square	$r = 0.57$ Chi-square ($p>0.05$)	WTP was reliable
Reference	What good	Elicitation method	Reliability approach	Correlation and other statistics	Conclusion
Thomson et al (1984)	Chronic arthritis	Questionnaire: Open-ended	Test retest reliability: Correlation coefficient	$r = 0.25$	Was not reliable
Whitehead and Hoban (1999)	Improved air and water quality	Telephone survey: Binary	Test-retest 5 years apart: Regression analysis and t-statistics	t-statistics ($p>0.05$)	WTP was conditionally reliable

APPENDIX 4: Interview guide for the baseline focus group discsussions.

1. What are the most common diseases of the people in your village?
2. What do you understand to be malaria or 'Iba'?
3. What are the causes of malaria?
4. Do mosquitoes constitute a problem in the village?
5. How would know that someone has malaria?
6. What do you do when someone has malaria?
7. How do you prevent people in your household from getting malaria?
8. How do you prevent mosquitoes from biting people?

[At this point, the respondents were given a brief talk on the scientific causes, prevention and treatment of malaria. There was also a brief introduction of ITNs].

9. How do you think that are the solutions to malaria problem?

[At this point, there was full presentation on ITNs by the moderator and the need for people to pay].

10. How do you suggest that the amount that people would be willing to pay for the ITNs would be elicited so that only truthful answers that will indicate the full value of the nets to the people would be elicited?
11. What factors could lead to some people changing the amounts they have stated they would pay when the nets are actually brought and sold in your village?
12. How should the sales of the ITNs be organised in your village, and what payment mechanisms should be used?

APPENDIX 5: Qualitative determination of malaria and mosquito nets perceptions and experiences

The FGDs were used to provide information about the perceptions and practices of the people concerning control of mosquitoes and malaria. The information generated was used to develop variables for the questionnaires. It also provided elements that were included in the WTP scenario, especially the choice of payment mechanism and sales strategies. The findings were also triangulated with those from the questionnaire surveys on similar topics in order to have clear pictures of malaria perceptions, practices and WTP for its control in the villages.

Most common diseases of the people

According to the people, the most common diseases, were malaria "Iba" in Igbo language), eye problems, arthritis and typhoid fever. The women from Ahani, together with the men from Amaetiti in addition mentioned skin nodules & rashes (symptoms of onchocerciasis that is hyper endemic in these communities) and high blood pressure. They all however agreed that Iba was the most common of them all.

The presence of fever was the major underpinning condition for determining that a condition was Iba, but not in all cases. This factor was brought to the fore in the next set of discussions that dealt with the classification of malaria or Iba. Iba was used the term used by the moderator in the discussion but the discussants used it interchangeably with malaria. Hence, malaria and Iba will be used interchangeably here.

Classification of malaria

Four broad groups of malaria were identified. The classification was based on some key symptoms associated with the different disease states. The different classifications were: malaria that comes with yellow eyes (Iba nchananya or Iba odo); Iba that comes with severe headache and joint pains (Iba okponku or Iba ngbajiukwu/aka or Iba ogbawaisi); Iba that is mild and may be associated with only fever and chills (Iba nkiti or Iba ntuoyi) and this was commonly called malaria or fever itself; Iba oji ("acute malaria" as a discussant put it); and a fifth group mentioned by an Amatetiti man was Iba ukwala (fever with cough) that comes during the rainy season and is cured by eating hot food.

Causes of malaria

All the FGD groups identified a battery of causes of malaria. However, most of the participants in all the groups identified mosquitoes as one of the causes of malaria. Others included overexposure to the sun, too much hard work ("oke doga doga"), drinking too much local gin and other alcoholic beverages, eating unhygienic food, poor diet, drinking contaminated water, houseflies, cockroach, blackfly and tse-tse fly. The rest were eating a lot of groundnuts, change in diet from traditional food and fruits to refined food. In the case of alcohol, one of the Ahani men said that raffia palm wine (ngwo) does not cause malaria, but rather it is oil palm wine (nkwu) that causes malaria. The groups did not believe that malaria could come from magic spells by witch doctors.

Mosquito nuisance

The participants all agreed that mosquitoes were a big problem in their communities. The Enugu-Akwu men said that in their case, this stemmed from the effects of the nearby Oji River to their village. The mosquito nuisance increases during the rainy season, when they start breeding in water collections in empty cans, stagnant water on the surfaces and in the bushes.

Symptoms of Malaria

When asked to discuss the symptoms of Iba, all the FGD groups correctly identified them. However, symptoms not commonly attributable to malaria (except in severe cases), but to other causes of fever like yellow eyes/urine were also mentioned by the groups. The Enugu-akwu women further classified the different types of Iba depending on the symptoms. They said that Iba that comes with fever, chills, headache and anorexia and lasts for about 4 days was Iba nkiti, while those that come with chest pain, dryness of the throat, weakness and lasts for more than 4 days with treatment was Iba typhoid.

Some of the participants noted that malaria could cause convulsions in children because of the associated high fever and if not treated on time. The discussant that added Iba oji to the classification also mentioned mental derangement as one of the symptoms of malaria. They did not believe that malaria could cause convulsions. However, they stated that if a child has a "high" fever, it might result to convulsions.

Health seeking

However, most of the discussants from all FGD groups recognised the use of orthodox medicine to treat malaria. Most of them patronise patent medicine dealers first, and if symptoms persist, they go to hospitals. Some people will take herbal medicine if the hospitals also fail to cure the illness. Nevertheless, many of the respondents would indulge in home treatment with paracetamol first, and if the fever does not subside, they would visit patent medicine dealers.

However, some individuals had some distinct health seeking peculiarities not generally shared by members of their groups. In this vein, two men from Enugu-Akwu would usually consult herbalists first and take herbal mixtures when they have malaria. Then, if this does not work, they will then go to the hospital. A man from Amaetiti would first take chloroquine and then follow it up with herbal medicine after 3 days (the other discussants in the group were surprised by this treatment regimen).

According to an Amaetiti woman: *"when I have Iba, I first take traditional (herbal) purge medicine, then buy and take orthodox medicine from patent medicine dealers"*. Yet another Amaetiti woman stated *"I self-prescribe drugs which I buy from patent medicine dealers, since I can not go to clinics/hospitals all the time. Sometimes, I take fruits, blood tonics and maloxine"*.

Some people said they concoct a home brew, boiled for between 20 to 30 minutes to cure malaria. This is by mixing leaves from paw-paw trees, lemon grass, lime and guava leaves. They drink and also bathe with the concoction. Others boil and drink the local anti-malaria herb called 'dogonyaro'.

The health seeking for treatment of child-hood malaria was not significantly different from that of the adults. However, the respondents were in agreement that they never give children herbal medicine since they are not strong enough to withstand the potency of the drugs and also will not like the bitter taste. Most of the discussants would first treat the children in the home with paracetamol and then go to patent medicine dealers who would "mix" drugs for the children. They sometimes take them to health centres and hospitals if the drugs from the patent medicine dealers do not work. None of the participants stores anti-malarials in his or her homes.

However, there was an individual peculiarity. According to an Amaetiti woman *"I first undertake home treatment with paracetamol first, then mix many herbs and epson salt, and then bathe them in it. Then if this doesn't work, I will buy them drugs from patent medicine dealers"*.

Malaria prevention with drugs

Some discussants said that preventive drugs like the popular Sunday-Sunday medicine was used by people especially the pregnant women. Only one woman actually gives paludrine to her children. According to one of the women *" I do not give preventive medicine when there is no sign of malaria"*.

Mosquito control

The discussants gave many methods on how mosquito nuisance could be controlled. These included closing doors and windows in the night to prevent mosquitoes from entering the house, applying mosquito repellent cream on the exposed body at night and clearing weed around the houses, wearing of long dresses, use of mosquito nets and physically killing mosquitoes. More mosquito control techniques were; destroying/emptying stagnant water collections, covering all water containers and insecticide sprays (commonly called Shelltox® after the first brand in the market) especially during the rainy season.

Some women from Ahani said that applying Kerosene on the body prevents mosquitoes from biting people. Some respondents maintained that sprinkling kerosene in the corners of the house drives away mosquitoes. Others were, putting of local leaves (osigbu) in the houses (these leaves which are of two varieties were claimed to be good mosquito repellents because of their smell). These herbs could also be burnt in and around the homes. Also mentioned were using ceiling fan and mosquito coils. One of the Enugu-akwu men said that putting rat-poison inside a pot of water attracts and kills mosquitoes

The evidence pointed to the low mosquito net ownership and use in these three villages. A few of the Amaetiti youths said that they had used bed-nets in secondary school boarding houses. One of the participant's families' uses bed-nets, while another's uses window-nets in their home in the urban city. None of the Enugu-Akwu women said that

they used mosquito nets due to the high cost, although one said that nursing mothers use nets to cover their babies.

However, one Amaetiti man had window-nets in his house, while another had bed-nets, and they said that these nets were efficient in keeping mosquitoes from biting people at night. Only one man from Enugu-Akwu had heard about ITNs through the radio. Many questioned their safety, since they feel that sleeping under nets treated with insecticides may harm the person. A woman from Amaetiti said, *"I will never use mosquito bed-nets even given free of charge. They make me feel as if I am being suffocated "*

When asked about the use of mosquito coils, the Enugu-Akwu men responded that *"most people do not use them because they are dangerous"*. This was because as most of the discussants who did not particularly like mosquito coils stated, they "cause cough". Most of the discussants do not use insecticide sprays because of their high cost and bad smell.

One of the Enugu-Akwu men said that *"anytime I have no money for* Shelltox®, *there is nothing I can do"*. According to another man *"God scares away the mosquitoes when I have no money to buy insecticide-spray"*.

[At this point, the respondents were given a brief talk on the scientific causes, prevention and treatment of malaria. There was also a brief introduction to ITNs. They were then asked to proffer solutions to the malaria problem].

Solution to malaria

The solutions proffered included the avoidance of risk factors and chemo-prophylaxis. Others were that promotion of good treatment practices and effective preventive measures. It was also suggested that people should always go to either patent medicine dealers or hospitals whenever they are ill. Mentioned also were preventing people from getting bitten by mosquitoes and discovery of drugs that can kill mosquitoes. They also suggested that mosquito nets could be used. The other solutions proffered were spraying of insecticides on stagnant water, malaria immunisation and spraying of insecticides in the air.

[At this point, there was full presentation on ITNs by the moderator and the need for people to pay. They were also asked their preferred sales and financing mechanisms for the ITNs].

ITNs sales

Most of the discussants agreed that the nets should be centrally sold, but the town crier should announce the availability of the nets and sales location. Some people also said that house to house sales could be used to sell to people unable to go to the central location to buy the nets. However, the Ahani youths were against door to door sales. As one of them explained *"if the ITNs are taken from house to house, people will value them less, because they will feel that the government was unable to sell them and is begging them to buy the nets"*.

ITNs financing

Most of the men favoured payment by instalment, but the youths favoured single full payment because of the difficulty in recovering debts. However, while some people preferred instalment payment after receiving the nets, some others favoured collection of nets after full instalment payment. A suggestion was made for the nets to be collected after full payment and that there should not be more than three instalments which could be on a weekly basis.

Another suggestion was that it would be better to accept partial payment after which the buyer collects the nets, and then completes payment later by instalment. All in all, single full payment was said to be the best payment system by all if people could afford to pay the full price at once. However, in the event that the price of the nets is high, then instalment payments should be allowed.

Pricing

Some people suggested that the nets should have a fixed but subsidised price. However, they were of the opinion that the community should not fix the price. Other participants felt that the price should be discussed at general community assemblies with the ITNs providers, so that the price at the nets would be sold would be arrived at. Other methods for setting the price of the nets were also suggested. A man from Enugu-akwu said that few God fearing people should be selected to set the prices. However, he did not give the criteria for selecting God fearing people.

One of the discussants said that community leaders should set the prices. To do this he said, the providers should give a certain range of prices, and the community leaders will then decide on one price. Some people felt that they should bargain with the community leaders in order to set the price. Finally, a unique method suggested for pricing was that "reasonable" men from various socio-economic groups should be assembled to set the prices - but no criteria was given on how to pick the people. In the words of a man from Amaetiti: *"The price of the nets should be very low. They should be fixed prices, and the ward leaders will tell all households about the price"*. Another stated that *"The trained field workers should educate the people on the importance of the nets, and then seek their opinion on the prices for the net. Then, the average price should be used as what people should pay"*.

General conditions for eliciting truthful WTP

In order to elicit truthful WTP, some of the discussants said that the benefits of the ITNs must be fully explained to the people before asking how much they would pay. In describing the importance of the nets, some people said that we must add that it is essential that all households have ITNs.

An insight gained was that only knowledgeable respondents would make truthful or realistic WTP offers for the nets. In this vein, some of the discussants recommended that the household heads should be educated about malaria and ITN before asking their WTP. It was recommended that we should explain the production cost of the net and all the costs involved in getting them down to the community.

Other people said that framing of the question was vital to get truthful responses, thereby recommending that the people should be asked how much they could afford to pay. In addition, they also recommended that we must inform people that the nets unlike ivermectin are not free of charge and that there would be no subsidy.

This establishment of credibility of the programme and that of the interviewers were found to be essential to make people be honest during the actual WTP elicitation. Hence, it was suggested that the community leaders should tell the people to be honest in answering questions. Also, some discussants recommended that the field workers

should also explain that they were working for people from the University of Nigeria Teaching Hospital.

Question formats to use for eliciting truthful WTP

Most of the discussants felt that we should bargain with the respondents to determine WTP. In bargaining for the price of the nets, it was suggested that we should start with high bids. It was further suggested that the producers of the net should first of all set the price and then people can haggle with the sellers. However, as a second choice, some people felt that the nets should have a fixed price and buyers asked whether or not they would purchase at the fixed price. Then, the answer should be followed up to find out their maximum WTP.

Further different question formats were suggested by individual groups of respondents. In this connection, a man from Amaeteti suggested that the people should be told the cost of the net, and then asked the maximum amount they would pay and the payment mechanism they preferred. Then in order to get the truthful maximum WTP, the respondents should then be asked how much extra they would be able to pay in six months time.

The women from Amaetiti suggested that the respondents should be shown the nets accompanied by a full explanation of the benefits. Then, the respondents would be told the cost price of procuring the nets and the insecticides respectively. Thereafter, the total cost of the nets should be told to the respondents followed by asking their maximum WTP.

Someone said that open-ended questions should not be used, but that people should be guided using a starting-point. This was because as he noted, people will start to strategize when asked how much they will pay as an open-ended question.

Factors leading to divergence between stated and actual WTP

The participants said that the major reason for the divergences would be either lack or immediate unavailability of money. The other reasons given were if the nets are sold at different prices to different people, or that the nets are unaffordable. However, the participants suggested that payment mechanisms that would help those who lack cash at time of sales, to be able to buy the nets later would decrease the divergence between stated and actual WTP.

Pre-testing: Pre-testing using the versions of the scenario and elicitation methods used in the study ensured that the scenario and elicitation formats were understood. No new changes were suggested on the structure of the elicitation methods. Modifications were carried out on the coding of some of the variables.

Discussion

FGDs brought out a number of points that were included in the design of the scenario that was used for eliciting WTP. These points were in reference to the terminology for malaria, advantages of ITNs over ordinary nets, payment vehicle, the delivery system and strategies for achieving truth-telling. Many of the variables that were included in the questionnaires were also derived from the FGDs.

Iba was the term used to describe malaria in the scenario. However, there was a transitional explanation informing the respondents that there were other possible causes of Iba as they indicated, but that the most common cause in their villages was malaria, brought about by mosquitoes as many of them stated. This was because, though the people were knowledgeable about malaria and the beneficial effects of using nets to protect themselves from the disease, their Iba classifications meant that some Iba episodes they call malaria may not be so. However, since amore than 85% percent of fever cases in Nigeria, are usually confirmed to be malaria (FMOH, 1992), the occasional wrong classification may not be a problem for their appreciation of the benefits of ITNs as a malaria preventive tool.

The advantages of ITNs over untreated nets were explained in the scenario. This was because, it was found that though the villagers were aware that mosquito nets could protect against mosquito bites and ultimately malaria, they were unaware of the availability of ITNs and the benefits. In addition, the scenario presented the case for protection using nets. This was due to very minimal level of net usage in the villages, and thus the benefit of net usage, albeit ITNs were described explicitly in the scenario.

In order that the payment vehicle and delivery systems for the ITNs and net re-treatment are as realistic as possible, the preferred options determined through the FGDs were used in the scenario. The use of preconceived notions on what should be the payment vehicle or optimal delivery system may lead to elicitation of invalid responses.

Thus, the original plan to stipulate a door-to-door mechanism for net sales was jettisoned, since the people felt that it would make the communities view the programme with suspicion. Also, the pre-FGDs decision to use cash-and-carry system for net sales so as to limit defaulting in payment was replaced by the recommended instalment payment system. The villagers preference for a common was consistent with the investigator's prior beliefs about its appropriateness. These preferences of the villagers as determined in the FGDs and included in the scenario were used during the actual sales of the ITNs.

Because, the people had many classifications for malaria and its causes, the questionnaire was designed in a way that the respondents were given the opportunity to indicate factors they thought that caused malaria and the place of preventive tools. Thus, the options for the variables on malaria perception, household use of protective measures and expenditures on malaria control were derived from the FGDs. This helped to restrict the options for most of the perceptions and practices variables to the most important options.

Also, because some respondents may still not understand the link between malaria and mosquitoes, at the end of the perceptions and practices part of the questionnaire that preceded the scenario, the respondents were given a transitional explanation on scientific causes of malaria. This helped to focus their minds on malaria together with its link to mosquitoes during the presentation of the scenario and subsequent WTP elicitation.

Finally, having and understanding the quantitative answers to the issues in the questionnaires helped the understanding and interpretation of the findings of the quantitative surveys, having the village perspectives as the base. Thus, the FGDs and the questionnaire findings were triangulated to give the best understanding of the malaria picture of the villages. The results showed that they were similar especially with regards to knowledge and practices to malaria and prevention of the disease with nets.

APPENDIX 6: Guide for the personal interviews with sellers of mosquito nets and FGDs with consumers.

1. What thoughts and observations came to your mind as the scenario was read to you?
2. What do you think of the elicitation methods, especially in terms of resembling actual processes in the market for arriving at final prices?
3. How do the different elicitation method resemble or differ from how you negotiate prices with your customers?
4. To what extent will the elicitation methods capture the actual price the consumers will normally pay?
5. Do elicitation methods give enough room for eliciting the maximum WTP of the consumers?
6. What are the flaws with the scenario and the elicitation method and how can these be remedied in order to approximate normal bargaining procedures in the real market?
7. How many haggling bids should be used in the structured haggling method?
8. Should there be a difference in the number of haggles for males and females?
9. How can the three elicitation methods and the scenario be improved so as to resemble the normal market transactions used to arrive at final prices for goods such as mosquito nets?

Appendix 7: The Igbo version of the questionnaire for the first survey

Ezi onye n'ege nti,

Anyi bu ndi olu ime obodo azuruazu, nke ndi olu University of Nigeria Teaching Hospital (UNTH) zipuru. O bu mmasi anyi i mara uche gi maka oria Iba, nke n'enye obodo anyi nsogbu, ma mara kwa uche gi ka a gesi gbochie oria a. N'ihi nkea, anyi choro igba ndi obodo ajuju onu, ka anyi wee mara uzo kachasi nma ga-adiri anyi nfe isi gbochie Iba na obodo anyi site na amuma ohuu akporo Ulo anwunta a gbara ogwu (Insecticide-treated nets - ITNs). I ga-enweta abamuru ajuju onua n'odiniru. Anyi nwere olile anya I were ihe anyi nwetara nyere obodo anyi aka igbochi oria Iba. Ka osinadi, ihe obula izara ka a ga-eji kwado amuma a. Ihe nile I ga-aza ga-adi n'etiti mu na gi. Gi onwe gi isonye zaa ajuju ndi a, bu site na mkpebi gi. I nwere ike, I gahgi aza ajuju o bula I choghi iza.

Instruction: Please fill the next section at the beginning and at the end of the interview.

What is the name of the respondent' s community?

What is the respondent's home address_____

What is the name of the interviewer _____

When was the date of interview	**Dateint[ ]**
What time did the interview start	**Timestart[ ]**
What time did the interview end	**Timeend[ ]**

Instruction: Please, enter the appropriate number representing the answer given in the spaces provided in both the question and code columns.

QUESTION	**CODE**
SECTION A: PERSONAL DATA (of respondent) A chorom ibido jua gi ajuju maka gi n'ezi na-uno gi	
1. Kedu onodu gi n'ezi n'uloa? 1 = Onye isi ezi n'ulo. 0 = Onye nnochite anya ezi n'ulo **[Instruction: ONLY the household head or his/spouse should be interviewed]**	**Status[ ]**
2. Madu ole bi n'a ezi n'uloa?	**Livein[ ]**
3. Madu ole ndi ezi n'uloa gbara afo ndia? a.Ha erubeghi afo ise b.Site n'afo ise rue n'afo iri n'ise c.Karia afo iri n'ise	**<5years [ ]** **5-15years[ ]** **>15years[ ]**
4. Kedu aha gi? _____	
5. Ibu nwoke ka I bu nwanyi? 1 = nwoke 0 = nwanyi	**sex[ ]**
6. Afo ole ka I di?	**Age[ ]**
7. I guru akwukwo? 1 = eee 0 = mba	**School[ ]**
7a. O buru n'iguru akwukwo, afo ole ka o were gi iji gua akwukwo?	**Schoolyears [ ]**
8. Kedu olu bu isi olu iji enweta ego? [ ]___	**Occupation [ ]**
9. I nuru di (ma o bu nwunye)? 1 = eee 0 = mba	**Married[ ]**
9a. Kedu onodu alum di na nwunye gi ugbua? 1 = anyi bi. 2 = di m (mao bu nwunye m anwuala). 3 = anyi agbasa go.	**Maritsat[ ]**

SECTION B: Information on malaria and mosquito

Introduction:(Read out to the respondent)

Biko, o ga-amasi m iju gi ufodu ajuju banyere oria a na-akpo Iba. O bu nnukwu oria n'enye ndi madu nsogbu mgbe nile. Otutu madu n'aria oria a ihe kariri otu ugboro n'afo. Mgbe ufodu, o na-adi njo nke ukwu o ka chasi n'ebe umu ntakiri no.

Instruction: Please, enter the appropriate number representing the answer given in the spaces provided in both the question and code columns. Read out the answer options and let the respondent answer mba, eee, ma obu amaghi m (mba = 0; eee = 1; amaghi m = 2)

Question	**Code**
10. Gini ka I chere n'ebuta iba? **(zaa mba, eee, ma obu amaghi m site na ihe ndi m ga-aguputara gi)** 10a = ikuru ajo ikuku 10b = anwunta ita 10c = nri manu 10d = ino noke anwu 10e = Iňu ajo miri 10f = Iru oru siri ike 10g = o nwere ihe ozo? (biko kwuputa):	**Air [ ]** **Mosquito[ ]** **Oil[ ]** **Sun[ ]** **Water [ ]** **Work [ ]** **Others[ ]**
11. Kedu uzo enwere ike isi gbochie iba? **(zaa mba, eee, ma obu amaghi m site na ihe ndi m ga-aguputara gi)** 11a = Iňu ogwu bekee 11b = iji ulo anwunta 11c = I gbu anwunta 11d = iji ogwu n'egbu umu aruru anagba ma munye oku (coil/shelltox) 11e = I zere ino n'anwu 11f = iji ahihia na-achupu anwunta 11g = onwere ihe ozo? (biko kwuputa):	Drugs [ ] Nets[ ] KillM [ ] Insecide[ ] Avoidsun [ ] Herbs [ ] Others[ ]

Instruction: Please, enter the number representing the answer given in the spaces provided in both the question and code columns. (Read out the options for the 1st parts of questions 12,13 & 14, and let the respondent answer mba = 0; eee = 1)

12.Dika I si ahuta anwunta, o nenye ezi n'ulo gi nsogbu? Mba = 0, Eee = 1 12a. Ole otu ezi n'ulo gi si ahuta anwunta dika nsogbu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 1 = Nnukwu 2 = etiti 3 = obere	Mworry [ ] PR [ ]
13. I chere na onye obula n'ezi n'ulo gi nwere ohere inweta Iba? 0 = mba 1 = eee 13a. Ole otu isi ele ohere onye obula n'ezi n'ulo gi I nweta iba? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 1 = Nnukwu 2 = etiti 3 = obere	Mchance [ ] RC [ ]
14. Ugboro ole ka iba n'abia kari n'ezi n'ulo gi n'afo? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 1 = opekatampe kwa onwa 2 = otu onwa rue onwa isii 3 = ihe kariri onwa isii 4 = ihe pekariri otu ugboro n'afo 14a. Iba o na-abia n'ezi n'ulo gi ugboro ugboro? [ ] 0 = Mba 1 = eee	PC [ ] Moccur[ ]

Igwo n'igbochi Iba

Ugbua, achorom iju gi ufodu ajuju etu ezi n'ulo gi si mefuo ego maka igwo n'igbochi oria iba n'onwa gara-aga.

15. Madu ole n'ezi n'uloa yara iba n'onwa gara-aga? **HADMAL1 [ ]**

15a. Afo ole ka-ndia yara iba gbara?	**Less than 5 years [ ]** **5 to 15 years [ ]** **More than 15 years [ ]**

15b. Na mgbako gi, ego ole ka ezi n'ulo gi mefuru na onwa gara-aga iji gwoo iba onye nke mbu yara site n'ihe ndia m ga aguputara gi (edeputara n'okpuru)?

Ebe ejere	Ugboro ole ka ejere?	Ego ole ka akwuru maka igwo iba	Ego ole ka akwuru maka igba motor	Onye kwuru ugwo ogwu	Ebee ka anatara ogwugwo	**CODE Total**
15b1 Ebe a na-ere ogwu						**Ch**
15b2 Ulo ogwu						**Ho**
15b3 Obere ulo ogwu						**Cl**
15b4 Ebe a na-enye ogwu						**Di**
15b5 Ebe a na -choputa oria						**La**
15b6 Dibia igbo						**He**
15b7 Ebe a na ere ogwu na ihe ndi ozo						**Md**
15b8 Ebe a na-ere ogwu n'ulo						**Hd**
15b9 Ndi n'ahu maka aru ike n'ime obodo.						**Vw**
15b10 Ndi n'agba afa/ndi ekpere						**Sh**
15b11 Onwere ihe ozo (biko kwuputa)						**Ot**
15b12 = Total						**Trtmal 1**

15c. Afo ole ka-onye yara ahu di? **AGEMAL1[ ]**

15d. Na mgbako gi, ego ole ka ezi n'ulo gi mefuru na onwa gara-aga iji gwoo iba onye nke abua yara site n'ihe ndia m ga aguputara gi (edeputara n'okpuru)?

Ebe ejere	Ugboro ole ka ejere?	Ego ole ka akwuru maka igwo iba	Ego ole ka akwuru maka igba motor	Onye kwuru ugwo ogwu	Ebee ka anatara ogwugwo a	**CODE Total2**
15d1 Ebe a na-ere ogwu						**Ch**
15d2 Ulo ogwu						**Ho**
15d3 Obere ulo ogwu						**Cl**
15d4 Ebe a na-enye ogwu						**Di**
15d5 Ebe a n'achoputa oria						**La**
15d6 Dibia igbo						**He**
15d7 Ebe a na ere ogwu na ihe ndi ozo						**Md**
15d8 Ebe a na-ere ogwu n'ulo						**Hd**
15d9 Ndi n'ahu maka aru ike n'ime obodo.						**Vw**
15d10 Ndi n'agba afa/ ndi ekpere						**Sh**
15d11 Onwere ihe ozo (biko kwuputa)						**Ot**

15d12 = Total						**Trtmal 2**

15e. Afo ole ka-onye yara ahu di? **AGEMAL2[ ]**

16. Na mgbako gi, ego ole ka ezi n'ulo gi mefuru n'onwa gara-aga maka I gbochi iba na nsoogbu anwunta site n'ihe ndia m ga aguputara gi (edeputara n'okpuru)?

Ihe e gotara	Ole odi	Ego ole ka otu bu	Mgbako ta ego ole obu	Onye gotara	Ebee ka e gotara	**CODE Prvntmal**
16aOgwu Ogbugba						**IJ[**
16b Nkpuruogwu						**TA[**
16c Ogwu mmiri						**SY[**
16dUde na-achu anwunta						**BC[**
16eUloanwunta(net)						**MN[**
16fOgwu aruru a na-agba agba (shelltox)						**IS[**
16g Ogwu anwunta a na-amunye oku (coil)						**MC[**
16hMkporogwu na ahihia						**HE[**
16iOnwere ihe ndi ozo (kwuputa)						**OT[**
16jMgbakota						**PT[**

Obere Nkowa ***(Read out to the respondent)***

Onwere ike nwe ihe ndi ozo n'ebuta iba dika i nwere ike I kwubu. Kaosinadi, anwunta bu isi ihe n'ebuta ya. Anwunta ndia n'atakari n'abali mgbe madu n'araru ura. Mgbe anwunta tara onye nwere iba, ma mesia taa onye n'enweghi ya, o ga-eburu nje o tatara onye n'aria iba bufee n'aru onye nke abuo, bu onye nwere ike malite riawa iba. N'uzo di otua ka iba si agbasa site n'a otu onye fee onye ozo. ***(Stop and ask the respondent: I nwere ajuju banyere otu iba si agbasa site n'otu onye fee onye ozo?).***

SCENARIO (Read out to the respondent)

Dika I nwere ike mara, ogwu anwunta a na-amunye oku (mosquito coils), nke a na-agba agba, mkporogwu na ahihia w.d.g, bu uzo ufodu esi egbochi anwunta ita. Uzo ozo esi egbochi anwunta ita madu n'araru ura ma o bu dina n'akwa bu site n'iji ulo anwunta nkiti(ordinary mosquito nets). Udidi ulo anwunta a bu ulo anwunta akwa (bed nets). Nsogbu ulo anwunta nkita a bu na-anwunta ka ga nebe na ya ma taa onye n'araru ura n'ime ya. O dighi egbu ma o bu chupu anwunta. Kaosinadi, o nwere usoro ohuru eji agba ulo anwunta ogwu opuru iche. Net ndia a gbara ogwu ka a na-akpo ulo anwunta agbara ogwu (Insecticide-treated nets = ITNs). Ulo anwunta a gbara ogwu na-achupu ma gbukwa anwunta. O na-ahu na anwunta ebeghi na ya wee taa onye n'araru ura n'ime ya. Nkea bu n'ihi ogwu puru iche a na-agba ulo anwunta ndia. Nchoputa e mere n'uwa nile egosiwo na netia na ebelata ugboro ole madu nwere ike iria iba. Net ndia a gbara ogwu n'egbu kwa umu aruru ndi ozo dika chinchi, igwu na ihe ndi ozo. ***(Stop and ask the respondent: Ihe di iche na net nkiti ya na ndi agbara ogwu na uru di na ya odoro gi anya?).***

Ogwu a na-agba net ndi a enweghi mmeru aru o bula nye ndi okenye ma umu ntakiri. Ka osinadi, obere metuta ogwua nwere bu so n'izuka mbu ebidoro jiwe netia. Obere mmetuta a nwere ike mee onye ahu ka o dika onye nwere azuzu. Nkea n'eme mgbe madu na net ahu agbara ogwu jekoro. Mgbe izu-uka mbu gasiri, mmetuta ahu akwusi. Mmetuta ahu adighi eme onye obula ji netia wee dina, kama n'eme madu ole na ole, ihe dika otu onye n'ime madu iri ise. Ka osinadi, ogwu dika paracetamol n'agwo obere mmetuta a.***(Stop and ask the respondent: I nwere ajuju?).***

Ekwesikwara iji ogwu opuruiche gbaghari netia ogwu opekatampe otu ugboro kwa afo ka o wee ga n'iru n'arusi oru ike. Ndi ga-agbaghari netia ogwu bu otu ndi aga aroputa n'ime obodo anyi na ndi na ahu maka igbochi iba nime Enugu state. Ezi n'ulo obula nwekwara ike iji aka ha gbaghari net ha ogwu ma enye sia ha obere ozuzu. I ga-akwu ugwo maka igbaghari net gi ogwu. ***(Stop and ask the respondent: I nwere ajuju?).***

Obu ezi okwu na otutu ezi n'ulo na-emefu nnukwu ego maka igwo na igbochi iba. Ndi n'aria iba na ndi n'eleta ha n'etufu oge ha ji aru oru na nke ha ji agu akwukwo. Ya mere, netia a agbara ogwu ga ewetuta ugboro ole madu nwere ike iria iba na imefu ego oge obula maka igwo iba n'ezi n'ulo gi.

Anyi nwere atumatu iweta netia agbara ogwu na obodo a, ma nwee olile anya na I ga-egoro netia, maka gi onwe gi na ndi ezi n'ulo gi nile. Nke a bu n'ihi na anyi agaghi enye net ndia na mmaka. Ndi a ga-eresi netia bu nani ndi kwuru ego ole zuru ego ejiri mebe net ndia. Ndi ga-ere netia bu ndi ime obodo

anyi a ga-enye ozuzu. Ha ga-ere netia ofu ebe ga adi onye obula nso, nya na ulo n'ulo. A ga-amalite ire netia mgbe o gasiri out onwa ajuju onua kwusiri. A ga-agwa ndi madu ngwa ngwa ebutara netia. I nwere ike n'akwubinye ugwo kwa izu uka, ma I ga akwuzu ugwo n'ime otu onwa. Ka osinadi, ikwuzu ugwo otu mgbe kachasi nma. I ga eburu net ma i kwuzuo ugwo. Aziza gi nye ajuju ndi a n'eso ka a ga-eji nwe mkpebi banyere amuma a.

Question	Code
18. Net ndia n'abia n'udi di iche iche. Kama nke anyi n'enye maka orire bu nke nnukwu (family size nets for 4 by 6 beds) dika I huru. **(Show the net to the respondents).** O n'abia kwa n'agwa di iche iche kama nke anyi ga ere na acha ocha. I tule ego gi onwe gi n'ezi n'ulo gi na-enweta n'onwa, aku na uba unu, ego ezi n'ulo gi na emefu, ohere unu inweta iba na nsogbu anwunta, I ga-enwe mmasi I gotara onwe gi netia? (biko,chee maka onwe gi ugbua maka emesia, agam aju gi mmasi idi ikwu ugwo maka ndi ezi n'ulo gi ndi ozo) [ ] 1 = Eee **(go to Q18a)** 0 = Mba **(go to Q18b and then Q21).**	**WTPYN1** [ ]

18a. Buru n'uche na I ga-akwu ego obula I ga-ekwu ngwa ngwa ebutara neti n'otu onwa n'abia, anyi choro ichoputa ego ole kachasi I nwere mmasi ikwu maka neti nke gi.	
18a1. Ego ole a n'ere otu neti bu 550 Naira; I nwere mmasi ikwu ego ole a? **1 = Eee(18a2) 0 = Mba (Q18a3)**	**Bid1** [ ]
18a2. O buru n'onwee otu onodu si gbanwe, onu ahia netia ria, wee buru zia 650 Naira, I nwere mmasi ikwu ego ole a maka otu net?**(Q18a4).**	**Bid2** [ ]
18a3. O buru n'onu ahia neti a buru 450 Naira, I nwere mmasi ikwu ego ole a?**(Q18a4)**	**Bid 3**[ ]
18a4. Kedu ego ole kachasi I nwere mmasi ikwu maka out net?	**WTP1**[ ]

18b. Gini mere na I choghi igo net maka gi onwe gi? **(Instruction: After the answers are given, go to number 21):** 18b1 = Ego adighi 18b2 = Achoghim ya 18b3 = Enweghim bed (akwa) 18b4 = Onwere ihe ndi ozo (Kwupute)	**Nomoney** [ ] **Nointerest**[ ] **No bed** [ ] **Others**[ ]

[Instruction: After question 18b, go to question number 21]

19. O buru n'obi siri gi ike maka I goro netia, agagbanye gi card n'akwughi ugwo. A na ele zi anya na I ga-akwu ugwo goro netia oge ebutara ya. Idenye aha n'akwu kwo putara na ituwo ahia neti ahu. Mgbakota onu ogugu madu ole turu ahia neti ahu ka a ge-eji bute ole a choro. Ya mere, ego ole bu nke ikpeazu I nwere mmasi ikwu maka otu neti nke gi?	**BidWTP1** [ ]

20. Ole mkpebisi ike gi ikwu ego ole ahu I kwuru? A gamu aguputara gi ufodu asisa, mana I ga-ewere otu. 1 = mkpebisi ike 2 = mkpebi 3 = ekpebighi 4 = ekpebisighi ike 5 = I maghaghi	**CalibYN1** [ ]

[Instruction: After question 20, go to question number 25]

21. N'agbanyere n'ichoghi igoro net maka gi onwe gi, anyi choro ichoputa ego ole kachasi I nwere mmasi ikwu maka neti maka ndi ezi n'ulo gi. Buru n'uche na I ga-akwu ego obula I ga-ekwu ngwa ngwa ebutara neti n'otu onwa n'abia	
21a. Ego ole a n'ere otu neti bu 550 Naira; I nwere mmasi ikwu ego ole a? **1 = Eee(21b) 0 = Mba (Q21c)**	**Bid1**
21b. O buru n'onwee otu onodu si gbanwe, onu ahia netia ria, wee buru zia 650 Naira, I nwere mmasi ikwu ego ole a maka otu net?**(Q21d).**	**Bid2**
21c. O buru n'onu ahia neti a buru 450 Naira, I nwere mmasi ikwu ego ole a?**(Q21d)**	**Bid 3**
21d. Kedu ego ole kachasi I nwere mmasi ikwu maka otu net?	**WTP2**

[Instruction: If also not willing to pay for other household members, ask question 22 and then go to question number 31].

22. Gini mere na I choghi igo net maka ndi ezi n'ulo gi? **(Instruction: After the answers are given, go to number 21):**22a = Ego adighi 22b = Achoghim ya 22c = Enweghim bed (akwa) 22d = Onwere ihe ndi ozo (Kwupute)	**Nomoney** **Nointerest** **No bed** **Others**

23. O buru n'obi siri gi ike maka I goro netia maka ndi ezi n'ulo gi, agagbanye gi card n'akwughi ugwo. A na ele zi anya na I ga-akwu ugwo goro neti ndia oge ebutara ya. Idenye aha n'akwu kwo putara na ituwo ahia neti ahu maka ndi ezi n'ulo gi. Mgbakota onu ogugu madu ole turu ahia neti ahu ka a ge-eji bute ole a choro. Ya mere, ego ole bu nke ikpeazu I nwere mmasi ikwu maka neti ndi ezi n'ulo gi buru kwa n'obi neti ole I choro I gotara ha?	**BidWTP2** **[ ]**

24. Ole mkpebisi ike gi ikwu ego ole ahu I kwuru? A gamu aguputara gi ufodu asisa, mana I ga-ewere otu. 1 = mkpebisi ike 2 = mkpebi 3 = ekpebighi 4 = ekpebisighi ike 5 = I maghi	**CalibYN1** **[ ]**

25. I buru n'uche na odimkpa ka igoputara ndi ozo n'ezi n'ulo gi netia ka e wee gbochie ofuma anwunta ita ha, na ibuta iba, neti ole ka I ga-egoro n'eego ole ahu [ ] I kwuru na asisa I nwere na ajuju 19 ***(Instruction: remind the respondent the amount s/he quoted in question 19)***?	**#nets1** **[ ]**

Instruction: If not willing to buy at that price but is willing to pay, go to question 25a.
If not willing to pay at all, go to question 25b.
If willing to buy net(s) at that price, go to question 25c.

25a. Ebe I choghi igotara netia maka ndi ezi n'ulo gi n'onu ego I kwuru I ga-ego nke gi, anyi choro ichoputa ego ole kachasi I ga-ego otu neti maka ndi ezi n'ulo gi ndi ozo. Buru n'uche na I ga-akwu ego obula I ga-ekwu ngwa ngwa ebutara neti n'otu onwa n'abia	
25a1. Ego ole a n'ere otu neti bu 550 Naira; I nwere mmasi ikwu ego ole a? **1 = Eee(25a2) 0 = Mba (Q25a3)**	**Bid7**
25a2. O buru n'onwee otu onodu si gbanwe, onu ahia netia ria, wee buru zia 650 Naira, I nwere mmasi ikwu ego ole a maka otu net?**(Q25a4).**	**Bid8**
25a3. O buru n'onu ahia neti a buru 450 Naira, I nwere mmasi ikwu ego ole a?**(Q25a4)**	**Bid 9**
25a4. Kedu ego ole kachasi I nwere mmasi ikwu maka otu net?	**WTP3**

(Instruction: if the respondent does not wish to buy a net for other household members, ask question 25b and then go to question number 31)

25b. Gini mere na I choghi igo net maka ndi ezi n'ulo gi? **(Instruction: After the answers are given, go to number 21):**25b1 = Ego adighi 25b2 = Achoghim ya 25b3 = Enweghim bed (akwa) 25b4 = Onwere ihe ndi ozo (Kwupute)	**Nomoney** **Nointerest** **No bed** **Others**

25c. Na mkpokota, neti ole ka I ga-egoro maka ndi ezi n'ulo gi?	**#nets2**

26. O buru n'obi siri gi ike maka I goro netia maka ndi ezi n'ulo gi, agagbanye gi card n'akwughi ugwo. A na ele zi anya na I ga-akwu ugwo goro neti ndia oge ebutara ya. Idenye aha n'akwu kwo putara na ituwo ahia neti ahu maka ndi ezi n'ulo gi. Mgbakota onu ogugu madu ole turu ahia neti ahu ka a ge-eji bute ole a choro. Ya mere, ego ole bu nke ikpeazu I nwere mmasi ikwu maka neti ndi ezi n'ulo gi buru kwa n'obi neti ole I choro I gotara ha?	**BidWTP3**

27. Ole mkpebisi ike gi ikwu ego ole ahu I kwuru? A gamu aguputara gi ufodu asisa, mana I ga-ewere otu. 1 = mkpebisi ike 2 = mkpebi 3 = ekpebighi 4 = ekpebisighi ike 5 = I maghaghi	**CalibBYN2**

28.Neti ole a, o ga-ezuru ndi ezi n'ulo gi nile? 1 = Eee 0 = Mba (Q28a)	**Enets[ ]**
28a. Kedu ndi ogbo ga-ano n'ejighi neti alaru ula? 28a1 = Less than 5 years 28a2 = 5 - 15 years 28a3 = More than 15 years 28a4 = Others (Please specify)	**<5yrs[ ]** **5 - 15yrs[ ]** **>15yrs[ ]** **Others[ ]**
28b. Gini mere ha agaghi eji neti laru ula?	

ITNs re-treatment scenario

Dika esi kwubu, o di mkpa igbaghari netia ogwu opekatampe otu ugboro kwa afo ka o wee ga n'iru n'irusi oru ike. Nke a bu maka na ike netia n'ebelata mgbe oge n'aga. Okwesighi ka asa netia asa rue mgbe a ga-agbaghari ya ogwu n'ihi na isa ya ugboro ugboro n'ewetulata irusi oru ike ya. Ndi ga-agbaghari netia ogwu ga-abu ndi aroputara na obodo anyi, bu ndi ga-ano n'ebe ga adi onye obula nso wee gbaghari netia ogwu. Ezi n'ulo obula nwe kwa ike igbaghari neti ha ogwu maka osighi ike ime. Ndi n'ahu maka igbochi iba na state a (Enugu state Malaria Control Unit) ekwenyewo ibuta ogwu mgbe obula ndi obodo choro ya.

29. Anyi choro ichoputa ego ole kachasi I nwere mmasi ikwu maka igbaghari neti di n'ezi n'ulo gi ogwu.	
29a. Ego ole a ga-eji gbaghari otu neti ogwu bu 100 Naira; I nwere mmasi ikwu ego ole a? **1 = Eee(29b) 0 = Mba (Q29c)**	**BidR1 [ ]**
29b. O buru n'onwee otu onodu si gbanwe, onu ahia eji agbanye neti ogwu ria, wee buru zia 150 Naira, I nwere mmasi ikwu ego ole?**(Q29d).**	**BidR2 [ ]**
29c. O buru n'onu ahia eji agbanye neti ogwu buru 50 Naira, I nwere mmasi ikwu ego ole a?**(Q29d)**	**BidR3[ ]**
29d. Kedu ego ole kachasi I nwere mmasi ikwu maka igbagahri otu net ogwu?	**WTPBR[ ]**

30. Ole mkpebisi ike gi ikwu ego ole ahu I kwuru? A gamu aguputara gi ufodu asisa, mana I ga-ewere otu. 1 = mkpebisi ike 2 = mkpebi 3 = ekpebighi 4 = ekpebisighi ike 5 = I maghaghi	**CalibBYN3 [ ]**

Altruism scenario

Dika onwere ike buru na ima, ufodu ezi n'ulo mmadu dara ogbenye nke bu na ha agaghi egotanwu neti ndia agbara ogwu. Anyi nwere olile anya n'ndi mmadu ga enwe ike tunye ego ka a were gotara ndi ogbenye a neti agbara ogwu. Obu ndi committee a ga-aroputa n'ime obodo ga-ahu ka aga esi were ego a ga-atuta were gotara ndi ogbenye neti.

31. Ego ole kachasi ka I ga-atunnye kwa afo ka iwere nye aka ka ndi ogbenye a wee new netia?	**AltruiWTP [ ]**

32. Ole mkpebisi ike gi ikwu ego ole ahu I kwuru? A gamu aguputara gi ufodu asisa, mana I ga-ewere otu. 1 = mkpebisi ike 2 = mkpebi 3 = ekpebighi 4 = ekpebisighi ike 5 = I maghaghi	**CalibBYN3 [ ]**

SECTION C: Other socio-economic factors

Kita, achorom iju gi ufodu ajuju etu ezi n'ulo gi si rie ego maka igwo oria ndi-ozo (n'abughi iba) n'onwa gara-aga.

33. Madu ole n'ezi uloa yara oria ndi ozo n'onwa gara-aga? **HADMIIL1 [ ]**

33a. Afo ole ka-ndia yara iba gbara?	**Less than 5 years [ ]** **5 to 15 years [ ]** **More than 15 years [ ]**

33b. Kedu ihe onye nke izizi a yara?

33c. Na mgbako gi, ego ole ka ezi n'ulo gi mefuru na onwa gara-aga iji gwoo oria ndi-ozo (n'abughi iba) onye nke izi yara site n'ihe ndia m ga aguputara gi (edeputara n'okpuru)?

Ebe ejere	Ugbolo ole ka ejere?	Ego ole ka akwuru maka igwo iba	Ego ole ka akwuru maka igba motor	Onye kwuru ugwo ogwu	Ebee ka anatara ogwugwo	**CODE Totalill1**
33c1 Ebe a na-ere ogwu						**Ch**
33c2 Ulo ogwu						**Ho**
33c3 Obere ulo ogwu						**Cl**
33c4 Ebe a na-enye ogwu						**Di**
33c5 Ebe a na -choputa oria						**La**
33c6 Dibia igbo						**He**
33c7 Ebe a na ele ogwu na ihe ndi ozo						**Md**
33c8 Ebe a na-ere ogwu n'ulo						**Hd**
33c9 Ndi n'ahu maka aru ike n'ime obodo.						**Vw**
33c10 Ndi n'agba afa/ ndi ekpere						**Sh**
33c11 Onwere ihe ozo (biko kwuputa)						**Ot**
33c12 = Total						**Trtill1**

33d. Afo ole ka-onye yara ahu di? **AGEILL1[ ]**

33e. Kedu oria onye nke abua yara?

33f. Na mgbako gi, ego ole ka ezi n'ulo gi mefuru na onwa gara-aga iji gwoo oria ndi-ozo (n'abughi iba) onye nke abua yara site n'ihe ndia m ga aguputara gi (edeputara n'okpuru)?

Ebe ejere	Ugbolo ole ka ejere?	Ego ole ka akwuru maka igwo iba	Ego ole ka akwuru maka igba motor	Onye kwuru ugwo ogwu	Ebee ka anatara ogwugwo	**CODE Totalill2**
33f1 Ebe a na-ere ogwu						**Ch**
33f2 Ulo ogwu						**Ho**
33f3 Obere ulo ogwu						**Cl**
33f4 Ebe a na-enye ogwu						**Di**
33f5 Ebe a na -choputa oria						**La**
33f6 Dibia igbo						**He**
33f7 Ebe a na ele ogwu na ihe ndi ozo						**Md**
33f8 Ebe a na-ere ogwu n'ulo						**Hd**
33f9 Ndi n'ahu maka aru ike n'ime obodo.						**Vw**
33f10 Ndi n'agba afa/ndi ekpere						**Sh**
33f11 Onwere ihe ozo (biko kwuputa)						**Ot**
33f12 = Total						**Toti2**

33g. Afo ole ka-onye yara ahu di? **AGEILL2[ ]**

34. Ego ole ka ezi n'ulo mefuru n'igota ihe oriri n'ahia n'ime out izu uka gara-aga site n'ihe oriri ndia m ga-aguputara gi?

Ihe egotara	Oha anuaa	Ego ole	Onye gotara	**CODE (buyfd)**
34a. Gari				**Garri**
34b. Agwa				**Beans**
34c. Abacha				**Cassava**
34d. Fio-Fio				**Fio-Fio**
34e. Ede				**Ede**
34f. Ji				**Yam**
34g. Osikapa				**Rice**
34h. Oka				**Corn**
34i. Azu				**Fish**
34j. Anu				**Meat**
34k. Akwukwo nri				**Veg**
34l. Onwere ihe ndi ozo (guputa)				**Others**
34m.Total				**Buyfd**

35. Oburu n'onwere nri ezi n'ulo gi meputara ma rikwaa n'izuka gara-aga, onu ego ole ka ha ga-aputa ma oburu n'ere ha n'ahia?

Ihe e meputara, e riri kwaa	Oha anuaa	Ego ole	CODE (Profd)
35a. Gari			**Garri**
35b. Agwa			**Beans**
35c. Abacha			**Cassava**
35d. Fio-Fio			**Fio-Fio**
35e. Ede			**Ede**
35f. Ji			**Yam**
35g. Osikapa			**Rice**
35h. Oka			**Corn**
35i. Azu			**Fish**
35j. Anu			**Meat**
35k. Akwukwo nri			**Veg**
35l. Onwere ihe ndi ozo (guputa)			**Others**
35m.Total			**Profd**

36. Gi ma o bu onye obula n'ezi n'ulo gi egotatuwo udidi neti nkiti? 1=Eee(Q36a) 0=Mba(Q37)	**Ppnets**
36a. Kedu udidi neti n'ego ole ka i/o kwuru I goro otu neti?	**Bednets1** **Windnet1** **Doornet1**
36b. Neti ole ka i/o goro?	**Bednets2** **Windnet2** **Doornet2**
36c. Ole afo i/o goro neti ikpeazu?	**Yearnet**

37. I ma ihe a n'ere neti nkiti nke akwa (ordinary bednet)? 1= eee(Q37a) 0 = Mba(Q37c)	**NetcstYN [ ]**
37a. Kedu etu oha?	**Size[ ]**
37b. Ego ole ka-ana ere ya?	**Netcost [ ]**
37c. I mara ebe I nwere ike I gota neti? 1 = eee(37d) 0 = mba(Q38)	**Buynet [ ]**
37d. Kedu ebe I nwere ike I gota neti? (Biko kwuputa)	

38. I nubu go maka neti a gbara ogwu tupu ajuju a, m na-agba gi? 1 = eee (38a) 0 = Mba(Q39) 38a. Ebee ka-inuru make neti agbara ogwu? (Bikoo kwuputa)	**HITNS[ ]**

39. Gi ma o bu onye obula n'ezi n'ulo gi egotatuwo udidi neti agbara ogwu? 1=Eee(Q39a) 0=Mba(Q40)	**Ppnets2 [ ]**
39a.Oburu na onwe go onye gotara netia agbara ogwu, ego ole ka i/o kwuru I goro otu neti?	**BedITN1[ ]** **WindITN1[ ]** **DoorITN1[ ]**
39b. Neti ole ka i/o goro?	**BedITN2[ ]**

39c. Ole afo i/o goro neti ikpeazu?	**WindITN2[ ]** **DoorITN2[ ]** **Yearnet2[ ]**

40. Gi ma o bu onye o bula n'ezi n'ulo gi agbatuwo ma o bu gbagharia neti ogwu? 1=eee(Q40a) 0=Mba (Q41)	**Treatnets [ ]**
40a. Ego ole ka i/o kwuru maka otu neti?	**Amtrtnet[ ]**

Aku n'uba ezi n'ulo

41. Site n'ime ihe ndi nga aguputara gi, ole ndi ezi n'ulo gi nwere? Biko, zaa eee ma obu mba. 1 = eee 0 = Mba	
41a. Igwe n'ekwu okwu (Radio)	**Radio[ ]**
41b. Anyinya igwe (Bicycle)	**Bicycle[ ]**
41c. Igwe aka eji akwo ihe (Manual grinding machine)	**Gmachine[ ]**
41d. Ogba tum tum (Motor cycle)	**Mcycle[ ]**
41e. Ugbo ala (Motor car)	**Mcar[ ]**
41f. Oche (Upholstery chairs/Cane chairs)	**Chairs[ ]**
41.g.	**IndexP[ ]**

42. Kedu udidi ulo ezi n'ulo a bi na ya? **[Instruction; interviewer to fill in after observing the house]**	
42a. Kedu ihe eji kpuchi elu uno a? 1 = Atanyi(Thatch) 2 = Zinc 3 = Cement 4 = Others (Please specify)	**Roof[**
42b. Kedu ihe ejiri we ruo ulo a? 1 = Mud 2 = plastered Mud 3 = block 4 = others (please specify)	**Wall**

43. [Instruction: Ask respondent whether there is any other information s/he would like to give to help us understand how to use the ITNs to fight Iba in our community]

44. Time interview ended:

Name and signature of the interviewer:

Appendix 8: The English version of the questionnaire for the first survey

Dear respondent,

We are trained local workers sent by staff of the University of Nigeria Teaching Hospital (UNTH) Enugu. We are interested in your opinion about the disease 'Iba' that troubles this community and ho to control it. To achieve this aim, we wish to conduct a questionnaire survey on the feasibility of starting a community-based malaria control project in this community with a new technology called Insecticide-treated nets (ITNs). You shall in the future get the results of this exercise, and we hope to use your responses to help your community to control Iba. Please be confident to answer what you really feel, since whatever you answer will affect the decision on whether or not the programme will be established. All information given will be private. Your participation is voluntary, and you do not have to answer questions you do not want to answer.

Instruction: Please fill the next section at the beginning and at the end of the interview.

What is the name of the respondent' s community?

What is the respondent's home address_____

What is the name of the interviewer _____

When was the date of interview	**Dateint[ ]**
What time did the interview start	**Timestart[ ]**
What time did the interview end	**Timeend[ ]**

Instruction: Please, enter the appropriate number representing the answer given in the spaces provided in both the question and code columns.

QUESTION	**CODE**
SECTION A: PERSONAL DATA (of respondent) I would like to begin by asking a few questions about you and your household.	
1. What is your status in this household? [ ] 1 = household head. 0 = representative of household **[Instruction: ONLY the household head or his/spouse should be interviewed]**	**Status[ ]**
2. How many people live in this household? [ ]	**Livein[ ]**
3. How many members of your household belong to the following age-groups? a. less than 5 years b. 5 to 15 years c. More than 15 years	**<5years [ ]** **5-15years[ ]** **>15years[ ]**
4. What is your name? _____	
5. Sex [ ] 1 = male 0 = female	**sex[ ]**
6. How old are you? [ ]	**Age[ ]**
7. Did you go to school? [ ] 1 = yes 0 = no 7a. If #6 is yes, what was the total number of years that you spent schooling?	**School [ ]** **Schoolyears [ ]**
8. What occupation is your major source of getting money? [ ]_____	**Occup [ ]**
9. Are you married [ ] 1 = yes 0 = no 9a. What is your current marital status? [ ] 1 = Single. 2 = married. 3 = widowed. 4 = divorced/separated.	**Married [ ]** **Maritalstat [ ]**

SECTION B: Information on malaria and mosquito

Introduction:(Read out to the respondent)

Please, I would like to ask you some questions about the disease called Malaria (iba). It is a serious disease affecting everybody all the time. Many people suffer from malaria attacks more than once in a year. Sometimes, it could be very serious, especially among children.

Instruction: Please, enter the appropriate number representing the answer given in the spaces provided in both the question and code columns. Read out the answer options and let the respondent answer yes, no or do not know (No = 0; yes = 1; do not know = 2)

Question	**Code**
10. What do you think or feel is/are the cause(s) of malaria? **(I will read out some options, just tell me yes, no or do not know)**	
10 a = breathing bad air [ ]	Air [ ]
10b = bite from mosquito [ ]	Mosquito[ ]
10c = oily food [ ]	Oil[ ]
10d = overexposure to the sun[ ]	Sun[ ]
10e = drinking bad water[ ]	Water [ ]
10f = doing strenuous work [ ]	Work [ ]
10g = others; please specify	Others[ ]
11. How can malaria be prevented? **(I will read out some options, just tell me yes, no or do not know)**	
11a = western medicine [ ]	Drugs [ ]
11b = mosquito nets [ ]	Nets[ ]
11c = kill mosquitoes	KillM [ ]
11d = use insecticides (coil/spray) [ ]	Insecide[ ]
11e = avoid the sun [ ]	Avoidsun [ ]
11f = using local leaves as insecticides [ ]	Herbs [ ]
11g = others; please specify	Others[ ]

Instruction: Please, enter the number representing the answer given in the spaces provided in both the question and code columns. (Read out the options for the 1st parts of questions 12,13 & 14, and let the respondent answer no = 0; yes = 1)

12. In your own view, do mosquitoes worry your household? [ ] 0 = No 1 = Yes	Mworry [ ]
12a. If #12 is yes, to what extent does your household see mosquito nuisance as a problem? **(I will read out three options for you to select one)** [ ] 1 = big 2 = medium 3 = small	PR [ ]
13. Do you think that anyone in your household stands a chance of getting malaria? [ ] 0 = No 1 = Yes 2 = Do not know	Mchance [ ]
13a. If #13 is yes, how do you rate the chance of anyone in your household contracting malaria? **(I will read out three options for you to select one)** [ ] 1 = big 2 = medium 3 = small	RC [ ]
14. How often does malaria occur in your household per year? **(I will read out three options for you to select one)** 1 = at least monthly 2 = one to six months 3 = six months to one year 4 = less than once yearly	PC [ ]
14a. Does malaria occur often in your household? [ ] 0 = No 1 = Yes 2 = Do not know	Moccur [ ]

Treatment and prevention of malaria:
Now, I would like to ask you some questions on how your household spent money to prevent and treat malaria within the past month.

15. How many people within your household had malaria within the past month? **Hadmal [ ]**

15a. What was/were the age-groups of the people that had malaria?	**Less than 5 years [ ]** **5 to 15 years [ ]** ***More than 15 years [ ]***

15b. In your own estimate, how much did your household spend within the last month to treat malaria suffered by household member 1 from the sources listed below?

Item	Number of visits	Amount spent on treatment	Amount spent on transportation	Who paid	Where was it bought	**CODE Treatmal1**
15b1.Chemist 15b2. Hospital 15b3. Clinic 15b4. Dispensary 15b5. Laboratory 15b6. Herbal home 15b7.Mixed dispensers 15b8.Home dispensers 15b9.Village health workers 15b10.Spiritual healers 15b11.Others (specify) 15b12. Total						**[ ]**

15c. What was the age of the ill person? **AGEMAL1[ ]**

15d. In your own estimate, how much did your household spend within the last month to treat malaria suffered by household member 1 from the sources listed below?

Item	Number of visits	Amount spent on treatment	Amount spent of transportation	Who paid	Where was it bought	**CODE Total2**
15d1.Chemist 15d2. Hospital 15d3. Clinic 15d4. Dispensary 15d5. Laboratory 15d6. Herbal home 15d7.Mixed dispensers 15d8.Home dispensers 15d9.Village health workers 15d10.Spiritual healers 15d11.Others (specify) 15d12. Total						**[ ]**

15e. What was the age of the ill person? **AGEMAL2[ ]**

16. In your own estimate, how much did your household spend within the last month to prevent malaria/mosquito nuisance on the items listed below?

Item	Quantity	Price of one	Total amount spent	Who purchased	Where purchased	CODE Prvntmal
16a.Injection						IJ
16b.Tablets						TA
16c.Syrup						SY
16d.Body cream repellent						BC
16e.Mosquito nets						MN
16f.Insecticide spray						IS
16g.Mosquito coils						
16h.Herbs						MC
16i.Others (specify)						HE
Total						OT
						PT

Transitional explanation (Read out to the respondent)

There may be other causes of Iba as you may have already said. However, mosquito bites is the major cause. These mosquitoes commonly bite in the night especially when people are sleeping. When a mosquito bites someone that has the malaria illness and later bites someone without the illness, it will transfer the germ collected from the ill person to the second person, who may now develop the illness. In this way, malaria is spread from person to person. ***(Stop and ask the respondent: Do you have any question about the way malaria is transferred from person to person?).***

SCENARIO (Read out to the respondent)

As you may be aware, mosquito coils, insecticide sprays, herbs etc are some of the ways used to prevent mosquito bites. Another method towards preventing mosquitoes from biting people is through the use of ordinary mosquito nets. A form of these nets is mosquito bed-nets. A problem with these ordinary nets is that mosquitoes still perch on them to bite people sleeping inside. Also, these ordinary nets neither kill nor chase away mosquitoes. However, there is a new system of treating mosquito nets with special chemicals. These treated nets are called Insecticide-treated nets (ITNs). They chase away and kill mosquitoes. They also ensure that the mosquitoes do not perch on the nets to bite somebody sleeping inside. These are as a result of the chemicals used to treat the nets. Studies around the world have shown that these treated nets can reduce the number of times one gets malaria. These nets also get rid of other insects like bed bugs, lice etc. ***(Stop and ask the respondent: Is the difference between simple and ITNs clear to you, together with the advantages of ITNs?).***

The chemicals used to treat the nets are harmless to both adults and children. However, mild side effects of the insecticides occur mostly in the first week of use of the Insecticide-treated nets. These side effects may make the person to feel that he/she has catarrh and presents with catarrh-like symptoms. These occur especially on direct contact with the nets, but thereafter there would be no symptoms after the first week. These side effects do not occur amongst everybody using the nets but only in a few people; about one out of fifty people. However, common drugs like paracetamol cure these symptoms. ***(Stop and ask the respondent: Do you have any question?).***

The nets however need at least yearly re-treatment with the chemicals in order to maintain their effectiveness. This net re-treatment will be organised by a committee to be selected from your community in conjunction with the Enugu State Malaria Control Unit. Individual households can also buy the chemicals and re-treat their nets after receiving a small training. You will have to pay for the net re-treatment. ***(Stop and ask the respondent: Do you have any question?).***

It is a fact that most households spend a lot of money either treating or preventing malaria. The sufferers and people looking after them also lose valuable work-hours and school-hours. Thus, the treated nets will ultimately reduce the time lost due to malaria and constant money spent to treat it in your household.

We have a programme to bring the treated nets to your community and hope that you can purchase them for both yourself and other members of your household. This is because the nets will not be given free of charge. The bed-nets will be sold only to those who have offered to pay enough money to cover the price of the nets. Trained community members will sell the treated nets at either central locations in your community and or door-to-door. The sales will commence within one month from the end of this interview. You can make weekly instalment payments, but total payments should be made in a maximum period of four weeks. You shall receive the nets after you have paid fully. However, cash-and-carry system is encouraged. Your answers to the following questions will be used to make decisions about design of the programme.

Question	Code
18. The nets come in several sizes, but the one we are offering for sale are *family size* nets (for 4 by 6 beds)) as you can see **(Show the net to the respondents)**. The nets also come in several colours. Taking into consideration your average monthly household income, wealth, expenditures, exposure to malaria and mosquito nuisance, would you be willing to buy a net for your personal use? (please think of yourself for now, as we will ask you about your willingness to pay for other household members later) [ ] 1= Yes **(go to Q18a)** 0 = No **(go to Q18b and then Q21)**	**WTPYN1** [ ]

Bidding Game

18a. Bearing in mind that you will be required to pay the amount that you will state once the nets are available within the next one month, what is the maximum amount that you are willing to pay for a net for your personal use?	
18a1. Are you willing to pay 550 for a net? 1 = Yes (Q18a2) 0 = No (Q18a3)	**Bid1** [ ]
18a2. What if the price is 650, will you be willing to pay? (go to Q18a4)	**Bid2** [ ]
18a3. What if the price is 450, will you be willing to pay? (go to Q18a4)	**Bid 3**[ ]
18a4. What is the maximum amount you are willing to pay for a net?	**WTP1**[ ]

18b. What are all the reason(s) you would not buy? **(Instruction: after the answers are given, go to number 21)** **18b1** = Lack of money 18b2 = Lack of interest 18b3 = Lack of bed 18b4 = Others (please specify)	No money No interest No bed Others

[Instruction: After question 18b, go to question number 21]

19. If you are sure of buying a net for your personal use, a free registration card will be issued to you and you are expected to purchase the net when it is available. Your being registered implies that you have ordered the net. The total number of those that have ordered the ITNs will be used to procure a certain quantity. Therefore, what is the final amount you are willing to pay for your personal net?	**BidWTP1** [ ]

20. With what certainty are you of actually paying the above amount? I will read out some options and you shall chose only one. 1 = very surely yes 2 = surely yes 3 = surely no 4 = very surely no 5 = do not know	**CalibBid1** [ ]

[Instruction: After question 20, go to question number 25]

21. Despite the fact that you are not willing to pay for a net for yourself, we would like to know the maximum amount that you are willing to pay per net for your other household members. Please, bear in mind that you will be required to pay the amount that you will state once the nets are available within the next one month.	
21a. Are you willing to pay 550 for a net? 1 = Yes (Q21b) 0 = No (Q21c)	
21b. What if the price is 650, will you be willing to pay? (Q21d)	**Bid4** [ ]
21c. What if the price is 450, will you be willing to pay? (Q21d)	**Bid5** [ ]
21d. What is the maximum amount you are willing to pay for a net?	**Bid 6**[ ] **WTP2**[ ]

[Instruction: If also not willing to pay for other household members, ask question 22 and then go to question number 31].

22. What are all the reason(s) you would not buy? **(Instruction: after the answers are given, go to number 31)** 22a = Lack of money 22b = Lack of interest 22c = Lack of bed 22d = Others (please specify)	No money [ ] No interest [ ] No bed [ ] Others [ ]

23. If you are sure of buying the nets for other household members, a free registration card will be issued to you and you are expected to purchase the nets when they are available. Being registered implies that you have ordered the nets for your household members and the total number of those that have ordered the ITN will be used to procure a certain quantity. Therefore what is the final amount you are willing to pay per net for other household members bearing in mind the number of nets you wish to purchase for them?	**BidWTP2** [ ]

24. With what certainty are you of actually paying the above amount? I will read out some options and you shall chose only one. 1 = very surely yes 2 = surely yes 3 = surely no 4 = very surely no 5 = do not know	**CalibBid2** [ ]

25. Bearing in mind that it is important that you buy the nets for other members of your household in order to assure maximum household protection from mosquito bites and ultimately malaria, how many nets would you purchase at this price of [ ] that you gave in question 19 to protect other household members from mosquitoes and malaria?. ***(Instruction: remind the respondent the amount s/he quoted in question 19)***	**#netsbid1** [ ]

Instruction: If not willing to buy at that price but is willing to pay, go to question 25a.
If not willing to pay at all, go to question 25b.
If willing to buy net(s) at that price, go to question 25c.

25a. Since, you are not willing to buy the nets for the household members at the price that you aid that you will buy yours, what is the maximum amount that you will pay for each net for other household members?	
25a1. Are you willing to pay 550 for a net? 1 = Yes (Q25a2) 0 = No (Q25a3)	**Bid7** [ ]
25a2. What if the price is 650, will you be willing to pay? (go to Q21a4)	**Bid8** [ ]
25a3. What if the price is 450, will you be willing to pay? (go to Q21a4)	**Bid 9**[ ]
25a4. What is the maximum amount you are willing to pay for a net?	**WTP3**[ ]

(Instruction: if the respondent does not wish to buy a net for other household members, ask question 25b and then go to question number 31)

25b. What are all the reason(s) you would not buy? **(Instruction: after the answers are given, go to number 31)** 25b1 = Lack of money 25b2 = Lack of interest 25b3 = Lack of bed 25b4 = Others (please specify)	No money [ ] No interest [ ] No bed [ ] Others [ ]

25c. How many nets in total will you purchase for other members of your household?	**#nets2** [ ]

26. If you are sure of buying the nets for other household members, a free registration card will be issued to you and you are expected to purchase the nets when they are available. Being registered implies that you have ordered the nets for your household members and the total number of those that have ordered the ITN will be used to procure a certain quantity. Therefore what is the final amount you are willing to pay per net for other household members bearing in mind the number of nets you wish to purchase for them?	**BidWTP3** **[ ]**

27. With what certainty are you of actually paying the above amount? I will read out some options and you shall chose only one. 1 = very surely yes 2 = surely yes 3 = surely no 4 = very surely no 5 = do not know	**CalibBid2** **[ ]**

28. Will the number of nets be enough for everyone in your household?1 = yes 2 = No (Q28a) 28a. What age-groups will not be sleeping under the nets? 28a1 = Less than 5 years 28a2 = 5 - 15 years 28a3 = More than 15 years 28a4 = Others (Please specify) 28b. Why would they not sleep under the nets?	**Enets[ ]** **<5yrs[ ]** **5 - 15yrs[ ]** **>15yrs[ ]** **Others[ ]**

ITNs re-treatment scenario

As already mentioned, the mosquito nets need at least once yearly re-treatment with the special chemicals for them to function effectively. This is because the effect of the insecticide wears out with time. The nets should not be washed until the time of re-treatment as washing reduces their effectiveness. The re-treatment could be easily done by selected community members at a central place in your community. It could also be done by individual households. The Enugu state Malaria Control Unit has agreed to provide the insecticides whenever needed by the community members.

29. What is the maximum amount that you are willing to pay (per net) to have the net(s) in your household re-treated once yearly? 29a. Are you willing to pay 100 to re-treat a net? 1 = Yes (Q29b) 0 = No (Q29c) 29b. What if the price is 150, will you be willing to pay? (go to Q25d) 29c. What if the price is 50, will you be willing to pay? (go to Q25d) 29d. What is the maximum amount you are willing to pay to re-treat a net?	 **BidR1 [ ]** **BidR2 [ ]** **BidR3[ ]** **WTPBR[ ]**

30. With what certainty are you of actually paying the above amount? I will read out some options and you shall chose only one. 1 = very surely yes 2 = surely yes 3 = surely no 4 = very surely no 5 = do not know	**CalibBYN3** **[ ]**

Altruism scenario

As you may be aware, some households or individuals in the community may be too poor to purchase the insecticide-treated bed nets. We hope that people would be able to contribute some money so that the ITNs could be purchased for the poor people. A committee that will be selected from within your village will manage the contributions for the poor people. They will also buy the nets and give them to the poor people.

31. What is the maximum amount that you are willing to contribute this year in order to help these poor people to acquire the nets?	**AltruiWTP** [ ]

32. With what certainty are you of actually paying the above amount? I will read out some options and you shall chose only one. 1 = very surely yes 2 = surely yes 3 = surely no 4 = very surely no 5 = do not know	**CalibBYN3** [ ]

SECTION C: Other socio-economic factors

Now, I would like to ask you some questions on how your household spent money to treat other illnesses (not including malaria).

33. How many other people in your household suffered from other illnesses within the past month? **HADILL1** [ ]

33a. What was/were the age-groups of the people that had malaria?	**Less than 5 years** [ ] **5 to 15 years** [ ] **More than 15 years** [ ]

33b. What was wrong with household member 1?

Item	Number of visits	Amount spent on treatment	Amount spent of transportation	Who paid	Where was it bought	**CODE Treatill1**
33c1.Chemist 33c2. Hospital 33c3. Clinic 33c4. Dispensary 33c5. Laboratory 33c6. Herbal home 33c7.Mixed dispensers 33c8.Home dispensers 33c9.Village health workers 33c10.Spiritual healers 33c11.Others (specify) 33c12. Total						[ ]

33d. What was the age of the ill person? **AGEILL1**[ ]

33b. What was wrong with household member 2?

33f. In your own estimate, how much did your household spend within the last month to treat malaria suffered by household member 1 from the sources listed below?

Item	Number of visits	Amount spent on treatment	Amount spent of transportation	Who paid	Where was it bought	**CODE Treatill2**
33f1.Chemist 33f2. Hospital 33f3. Clinic 33f4. Dispensary 33f5. Laboratory 33f6. Herbal home 33f7.Mixed dispensers 33f8.Home dispensers 33f9.Village health workers 33f10.Spiritual healers 33f11.Others (specify) 33f12. Total						[ ]

33g What was the age of the ill person? **AGEILL2[ ]**

33f. What was the age of the ill person? **AGEILL2[ ]**

34. How much did your household spend to purchase of food from the market in the past one week on the various items that I will read out?

Item	Quantity	Amount	Who purchased	**Buyfood**
34a.Gari 34b.Beans 34c.Cassava (akpu) 34d.Fio-Fio 34e.Cocoyam (ede) 34f.Yam 34g.Rice 34h.Corn 34i.Fish 34j.Meat 34k.Vegetables 34l.Others (specify) 34m.Total				[ ]

35. If the food items that your household produced that it also consumed in the past one week were bought from the market, how much will they cost?

Item	Quantity	Amount	**Profood**
35a.Gari 35b.Beans 35c.Cassava (akpu) 35d.Fio-Fio 35e.Cocoyam (ede) 35f.Yam 35g.Rice 35h.Corn 35i.Fish 35j.Meat 35k.Vegetables 35l.Others (specify) 35m.Total			[ ]

36. Have you or anyone in your household ever bought any form of ordinary mosquito nets?1=yes (Q36a) 2=No (Q37)	**Ppnets [ ]**
36a. What type of net and how much did you pay per net?	**Bednets [ ]** **Windnet[ ]** **Doornet [ ]**
36b. How many nets were purchased?	**Bednets [ ]** **Windnet[ ]** **Doornet [ ]**
36b. What was the year you made the last purchase?	**Yearnet [ ]**

37. Do you know what an ordinary mosquito net costs in the market? 1= Yes(Q37a) 0 = No(Q37c)	**NetcstYN [ ]**
37a. What is the size of the net?	**Size[ ]**
37b. How much is one unit of it sold?	**Netcost [ ]**
37c. Do you know where you can purchase the net? 1 = Yes(37d) 0 = No(Q38)	**Buynet [ ]**
37d. Where can you purchase the net?	

38. Have you ever heard of insecticide-treated nets? 1 = Yes (38a) 0 = No(Q39)	**HITNS[ ]**
38a. Where did you learn about the ITNs? (Please explain)	

39. Have you or anyone in your household ever bought any form of insecticide-treated nets? 1=yes 2=No	**Ppnets2[ ]**
39a. How much did you pay per treated net?	**BedITN1 [ ]** **WindITN1[ ]** **DoorITN1[ ]**
39a. How many treated nets were purchased?	**BedITN2 [ ]** **WindITN2[ ]** **DoorITN2[ ]**
39c. What was the year you made the last purchase?	**Yearnet2[ ]**

40. Have you or anyone in your household ever treated/re-treated your mosquito nets?1=yes (Q40a) 2=No (Q41)	**Treatnets [ ]**
40a. How much did you pay per net?	**Amtrtnet [ ]**

Assets of the household

41. From the assets that I am going to read out to you, could please tell me the ones that your household has? Please answer yes or no. 1 = yes 0 = No	
41a. Radio	**Radio**
41b. Bicycle	**Bicycle**
41c. Manual grinding machine	**Gmachine**
41d. Motor cycle	**Mcycle**
41e. Motor car	**Mcar**
41f. Upholstery chairs/Cane chairs	**Chairs**

42. What type of housing does the household live in? **(interviewer to describe)**	
42a. What type of roof? [ ] 1 = Thatch 2 = Zinc 3 = Cement 4 = others (please specify)	**Roof [ ]**
42b. What type of wall are the rooms made of [ ] 1= mud 2 = plastered mud 3 = block 4 = others (please specify)	**Wall [ ]**

43. [Instruction: Ask respondent whether there is any other information s/he would like to give to help us understand how to use the ITNs to fight Iba in our community]

44. Time interview ended:

Name and signature of the interviewer:

Appendix 9: The Igbo version of the questionnaire for the last survey

Ezi onye n'ege nti,

Dika I matara, anyi bu ndi olu ime obodo azuruazu, nke ndi olu University of Nigeria Teaching Hospital (UNTH) zipuru. Cheta na nime onwa gara-aga, anyi juru gi juputara ego ole I inwere mmasi ikwu maka neti a gbara ogwu, were mata kwuo etu aga esi kpokoba mkpochapu iba n'ime obodo nke anyi. Anyi amalite go ire netia agbara ogwu. Ka osinadi, o bu mmasi anyi iju gi ufodu ajuju yiri nke anyi juburu gi n'mbu. Ufodu ajuju ga-adi iche n'nke agbara gi n'mbu. Aziza gi ga-enyere anyi aka n'ima ga anyi ka esi mee ka amuma a ga esi were gaa n'iru ma gbanye kwa nkporogwu n'obodo anyi. Ihe nile I ga-aza ga-adi n'etiti mu na gi. Gi onwe gi isonye zaa ajuju ndi a, bu site na mkpebi gi. I nwere ike, I gaghi aza ajuju o bula I choghi iza.

Instruction: Please fill the next section at the beginning and at the end of the interview.

What is the name of the respondent' s community?

What is the respondent's home address_____

What is the name of the interviewer?_____

When was the date of interview?	**Dateint[ ]**
What time did the interview start?	**Timestart[ ]**
What time did the interview end?	**Timeend[ ]**

Instruction: Fill in questions 26, 36 and 37 from your register before starting the interview.

Instruction: Please, enter the appropriate number representing the answer given in the spaces provided in both the question and code columns.

QUESTION	**CODE**
SECTION A: PERSONAL DATA (of respondent) [*Instruction: It must be the same respondents that was interviewed during the first interview*] A chorom ibido jua gi ajuju maka gi n'ezi n'ulo gi	
1. Kedu onodu gi n'ezi n'uloa? 1 = Onye isi ezi n'ulo. 0 = Onye nnochite anya ezi n'ulo **[Instruction: ONLY the same responents that were interviewed during the first survey should be interviewed]**	**Status[ ]**
2. Madu ole bi n'a ezi n'uloa?	**Livein[ ]**
3. Madu ole ndi ezi n'uloa gbara afo ndia? a.Ha erubeghi afo ise b.Site n'afo ise rue n'afo iri n'ise c.Karia afo iri n'ise	**<5years [ ]** **5-15years[ ]** **>15years[ ]**
4. Madu ole bi na-ezi n'uloa obuwa nyere ibu ka-opekanata mpe, si n'onwa gara aga agbara gi ajuju, wee rue ugbua? 0 = Opekanata mpe, 1 = Obuwa nyere ibu 2 = Odi entu odi bu	**Livein2[ ]**
5. Kedu aha gi? _____	**sex[ ]**
6. Ibu nwoke ka I bu nwanyi? 1 = nwoke 0 = nwanyi	**Age[ ]**
7. Afo ole ka I di?	**Sch[ ]**
8. I guru akwukwo? 1 = eee 0 = mba	
9. Kedu olu bu isi olu iji enweta ego? [ ]_____	**Occupation [ ]**
10. I nuru di (ma o bu nwunye)? 1 = eee 0 = mba	**Marital[ ]**
11. Onwere nwa ohuru a muru n'ezi n'uloa site n'oge ikpeazu a gabara gi ajuju were rue ugbua? 1 = Eee 0 =Mba	**Birth[ ]**
12. Onwere onye obula n'ezi n'uloa di ime? 0 = Eee 0 = Mba	**Preg [ ]**

SECTION B: Information on malaria and mosquito

Introduction:(Read out to the respondent)

Biko, o ga-amasi m iju gi ufodu ajuju banyere oria Iba.

Instruction: Please, enter the number representing the answer given in the spaces provided in both the question and code columns.

13. Dika I si ahuta anwunta, o nenye ezi n'ulo gi nsogbu? 0 = Mba, 1 = Eee	Mworry [ ]
13a. Kedu ka ihe mgbanwe n'etu isi ahuta nsogbu anwunta nenye ezi n'ulo gi ha site n'oge mbu agbara gi ajuju wee rue ugbua? 0 = Opelatara mpe 1 = Obuwanyere ibu 2 = O nweghi ihe di iche.	Mworry2[ ]
14. Ole otu ezi n'ulo gi si ahuta anwunta dika nsogbu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 1 = Nnukwu 2 = etiti 3 = obere	PR [ ]
15. I chere na onye obula n'ezi n'ulo gi nwere ohere inweta Iba? 0 = mba 1 = eee	Mchance [ ]
16. Kedu ka ihe mgbanwe n'otu isi ahuta ohere onye obula n'ezi n'ulo gi inweta iba ha. Mgbanwe a bu site site n'oge mbu agbara gi ajuju wee rue ugbua? **(Achorom ka iwere otu aziza site n'ihe ato m ga-aguputara gi)** 1 = Nnukwu 2 = etiti 3 = obere	Mchance 2[ ]
17. Kedu ka mgbanwe n'ugboro ole iba n'aya gi ma obu ndi ezi n'ulo gi ha, site n'oge mbu agbara gi ajuju wee rue ugbua? **(Achorom ka iwere otu aziza site n'ihe ato m ga-aguputara gi)** 0 = Opelatara mpe 1 = Obuwanyere ibu 2 = O nweghi ihe di iche.	Moccur[ ]

Igwo n'igbochi Iba

Ugbua, achorom iju gi ufodu ajuju etu ezi n'ulo gi si mefuo ego maka igwo n'igbochi oria iba n'onwa gara-aga.

18. Madu ole n'ezi n'uloa yara iba site n'oge mbu agbara gi ajuju wee rue ugbua?	**HADMAL[ ]**
19. Afo ole ka-ndia yara iba gbara?	
Less than 5 years	**Less than 5 years[ ]**
5 to 15 years	**5 to 15 years [ ]**
More than 15 years	**More than 15 years [ ]**

20. Na mgbako gi, ego ole ka ezi n'ulo gi mefuru na onwa gara-aga iji gwoo iba onye nke mbu yara site n'ihe ndia m ga aguputara gi (edeputara n'okpuru)?

Ebe ejere	Ugboro ole ka ejere?	Ego ole ka akwuru maka igwo iba	Ego ole ka akwuru maka igba motor	Onye kwuru ugwo ogwu	Ebee ka anatara ogwugwo	**CODE Total**
20a1 Ebe a na-ere ogwu						**Ch**
20a2 Ulo ogwu						**Ho**
20a3 Obere ulo ogwu						**Cl**
20a4 Ebe a na-enye ogwu						**Di**
20a5 Ebe a na -choputa oria						**La**
20a6 Dibia igbo						**He**
20a7 Ebe a na ere ogwu na ihe ndi ozo						**Md**
20a8 Ebe a na-ere ogwu n'ulo						**Hd**
20a9 Ndi n'ahu maka aru ike n'ime obodo.						**Vw**
20a10 Ndi n'agba afa/ndi ekpere						**Sh**
20a11 Onwere ihe ozo (biko kwuputa)						**Ot**

20a12 = Total						**Trtmal1**

20b. Afo ole ka-onye yara ahu di? **AGEMAL1[ ]**

21. Na mgbako gi, ego ole ka ezi n'ulo gi mefuru na onwa gara-aga iji gwoo iba onye nke abua yara site n'ihe ndia m ga aguputara gi (edeputara n'okpuru)?

Ebe ejere	Ugboro ole ka ejere?	Ego ole ka akwuru maka igwo iba	Ego ole ka akwuru maka igba motor	Onye kwuru ugwo ogwu	Ebee ka anatara ogwugwoa	**CODE Total2**
21a. Ebe a na-ere ogwu						**Ch**
21b. Ulo ogwu						**Ho**
21c. Obere ulo ogwu						**Cl**
21d. Ebe a na-enye ogwu						**Di**
21e. Ebe a n'achoputa oria						**La**
21f. Dibia igbo						**He**
21g. Ebe a na ere ogwu na ihe ndi ozo						**Md**
21h. Ebe a na-ere ogwu n'ulo						**Hd**
21i. Ndi n'ahu maka aru ike n'ime obodo.						**Vw**
21j. Ndi n'agba afa/ ndi ekpere						**Sh**
21k. Onwere ihe ozo (biko kwuputa)						**Ot**
21l. = Total						**Trtmal2**

21m. Afo ole ka-onye yara ahu di? **AGEMAL2[ ]**

22. Na mgbako gi, ego ole ka ezi n'ulo gi mefuru n'onwa gara-aga maka I gbochi iba na nsoogbu anwunta site n'ihe ndia m ga aguputara gi (edeputara n'okpuru)?

Ihe e gotara	Ole odi	Ego ole ka otu bu	Mgbako ta ego ole obu	Onye gotara	Ebee ka e gotara	**CODE Prvntmal**
22aOgwu Ogbugba						**IJ[**
22b Nkpuruogwu						**TA[**
22c Ogwu mmiri						**SY[**
22dUde na-achu anwunta						**BC[**
22eUloanwunta(net)						**MN[**
22fOgwu aruru a na-agba agba (shelltox)						**IS[**
22g Ogwu anwunta a na-amunye oku (coil)						**MC[**
22hMkporogwu na ahihia						**HE[**
22iOnwere ihe ndi ozo (kwuputa)						**OT[**
22jMgbakota						**PT[**

Obere Nkowa ***(Read out to the respondent)***

Onwere ike nwe ihe ndi ozo n'ebuta iba dika i nwere ike ichere ma-obu mara. Kaosinadi, anwunta bu isi ihe n'ebuta ya. Anwunta ndia n'atakari n'abali mgbe madu n'araru ura. Mgbe anwunta tara onye nwere iba, ma mesia taa onye n'enweghi ya, o ga-eburu nje o tatara onye n'aria iba bufee n'aru onye nke abuo, bu onye nwere ike malite riawa iba. N'uzo di otua ka iba si agbasa site n'a otu onye fee onye ozo. ***(Stop and ask the respondent: I nwere ajuju banyere otu iba si agbasa site n'otu onye fee onye ozo?).***

SCENARIO (Read out to the respondent)

Dika I nwere ike mara, ogwu anwunta a na-amunye oku (mosquito coils), nke a na-agba agba, mkporogwu na ahihia w.d.g, bu uzo ufodu esi egbochi anwunta ita. Uzo ozo esi egbochi anwunta ita madu n'araru ura ma o bu dina n'akwa bu site n'iji ulo anwunta nkiti(ordinary mosquito nets). Udidi ulo anwunta a bu ulo anwunta akwa (bed nets). Nsogbu ulo anwunta nkita a bu na-anwunta ka ga nebe na ya ma taa onye n'araru ura n'ime ya. O dighi egbu ma o bu chupu anwunta. Kaosinadi, o nwere usoro ohuru eji agba ulo anwunta ogwu opuru iche. Net ndia a gbara ogwu ka a na-akpo ulo anwunta agbara ogwu (Insecticide-treated nets = ITNs). Ulo anwunta a gbara ogwu na-achupu ma gbukwa anwunta. O na-ahu na anwunta ebeghi na ya wee taa onye n'araru ura n'ime ya. Nkea bu n'ihi ogwu puru iche a na-agba ulo anwunta ndia. Nchoputa e mere n'uwa nile egosiwo na netia na ebelata ugboro ole madu nwere ike iria iba. Net ndia a gbara ogwu n'egbu kwa umu aruru ndi ozo dika chinchi, igwu na ihe ndi ozo. ***(Stop and ask the respondent: Ihe di iche na net nkiti ya na ndi agbara ogwu na uru di na ya odoro gi anya?).***

Ogwu a na-agba net ndi a enweghi mmeru aru o bula nye ndi okenye ma umu ntakiri. Ka osinadi, obere metuta ogwua nwere bu so n'izuka mbu ebidoro jiwe netia. Obere mmetuta a nwere ike mee onye ahu ka o dika onye nwere azuzu. Nkea n'eme mgbe madu na net ahu agbara ogwu jekoro. Mgbe izu-uka mbu gasiri, mmetuta ahu akwusi. Mmetuta ahu adighi eme onye obula ji netia wee dina, kama n'eme madu ole na ole, ihe dika otu onye n'ime madu iri ise. Ka osinadi, ogwu dika paracetamol n'agwo obere mmetuta a.***(Stop and ask the respondent: I nwere ajuju?).***

Ekwesikwara iji ogwu opuruiche gbaghari netia ogwu opekatampe otu ugboro kwa afo ka o wee ga n'iru n'arusi oru ike. Ndi ga-agbaghari netia ogwu bu otu ndi aga aroputa n'ime obodo anyi na ndi na ahu maka igbochi iba nime Enugu state. Ezi n'ulo obula nwekwara ike iji aka ha gbaghari net ha ogwu ma enye sia ha obere ozuzu. I ga-akwu ugwo maka igbaghari net gi ogwu. ***(Stop and ask the respondent: I nwere ajuju?).***

Obu ezi okwu na otutu ezi n'ulo na-emefu nnukwu ego maka igwo na igbochi iba. Ndi n'aria iba na ndi n'eleta ha n'etufu oge ha ji aru oru na nke ha ji agu akwukwo. Ya mere, netia a agbara ogwu ga ewetuta ugboro ole madu nwere ike iria iba na imefu ego oge obula maka igwo iba n'ezi n'ulo gi.

Anyi ebutala netia agbara ogwu na obodo a, ma nwee olile anya na I ga-egoro netia, maka gi onwe gi na ndi ezi n'ulo gi nile. Nke a bu n'ihi na anyi agaghi enye net ndia na mmaka. Ndi a ga-eresi netia bu nani ndi kwuru ego ole zuru ego ejiri mebe net ndia. Ndi na-ere netia bu ndi ime obodo anyi a ga-enye ozuzu, bu ndi dika mu. Ebe a ga-ano ere netia bu n'ebe ga-adi onye obula nso ije. Kama, a ga-ebugara netia n'ulo ndi enwere ike bia gota netia n'ebe a n'ere ya. I nwere ike n'akwubinye ugwo kwa izu uka, ma I ga akwuzu ugwo n'ime izu ano. Ka osinadi, ikwuzu ugwo otu mgbe kachasi nma. I ga eburu neti a ma i kwuzuo ugwo. Aziza gi nye ajuju ndi a n'eso ka a ga-eji nwe mkpebi banyere amuma a.

23. Ndi WHO e butala netia. Onu ego ya bu 450 Naira maka ottu net. I ga-enwe mmasi I gotara onwe gi netia? (biko,chee maka onwe gi ugbua maka emesia, agam aju gi mmasi idi ikwu ugwo maka ndi ezi n'ulo gi ndi ozo) **Instruction: Show the net to the respondent] 1 = Eee (go to Q26) 0 = Mba (go to Q24).**	**WTPYN1 [ ]**

Instruction: If willing to purchase a net at the stated price, go to question 26, but if not willing to buy, go to question 24. If not willing to pay for self but for other household members go to Q30].

24. Gini mere na I choghi igo net maka gi onwe gi? **(Instruction: After the answers are given, go to number 26): Reasons for refusing to buy:** _____

Code for reasons the respondent refused to buy for him/herself. Codes: 1 = Eee 0 = Mba

24a = Ego adighi 24b = Achoghim ya 24c = Enweghim bed (akwa) 24d = Onwere ihe ndi ozo (Kwupute)	**Nomoney [ ]** **Nointerest[ ]** **No bed [ ]** **Others[ ]**

[Instruction: After question 24d, go to question 25 if respondent wishes to buy but not at the fixed price of the net, but if not willing to but at all, go to question 26]

25. N'agbanyere n'inweghi ike ikwu 450 Naira bu ihe a n'ere ya, mana ichoro igoro neti maka gi onwe gi. Ego ole kachasi ka I nwere mmasi ikwu?

25a. Oburu na ihe a ga-ere out netia bu 440 Naira; I nwere mmasi ikwu ego ole a? 1 = Eee(25b) 0 = Mba (Q25b)	**Bin1 [ ]**
25b. Kedu kwanu ego ole kachasi I nwere mmasi ikwu? ***[Instruction; no mater the answer go to 25c]***	**BWTP1[ ]**
25c. Ihe a ga-ere netia bu 450 Naira, I nwere mmasi ikwu? 1 = Eee 0 = Mba	**Bin2 [ ]**

[Instruction: Go to question 26].

[Instruction: Fill in question 26 from your register before starting the interview]

26. Fill in the amount the respondent stated that s/he was willing to pay for personal ITNs in the first survey.	**BWTPOrig [ ]**
26a. Code same amount or different. 0 = same 1 = different	**WTPDiff [ ]**

[Instruction: If first and present WTP amounts differ, go to question 27 and ask the respondent why s/he changed his/her mind after reminding the person his/her WTP amount from the first survey. However, if the same go to question 28].

Code for reasons the respondent changing his/her mind 1 = Eee 0 = Mba

27. Ego ole ikwuburu n'iga akwu maka ottu netia n'onwa gara aga agbara gi ajuju di iche n'ego ole ikwuru tata. Biko, kedu ihe mere igiri wee gbanwe uche gi banyere ego ole I nwere mmasi ikwu?	**Nomoney [ ]** **Nointerest[ ]** **No bed [ ]** **Others[ ]**

28. Kedu, otu osiri di gi mma ikwu ugwo neti maka gi onwe gi? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi).** 1 = Ikwubinye ugwo kwa izu uka, ma kwusia ugwo a n'ime izu ano. 2 = Ikwuzu ugwo otu mgbe ma buru neti ozigbo 3 = Onwere ottu ozo (Kwupute)	**PaymntM[ ]**

Instruction: Ask Q29 to only those willing to purchase the net at 450Naira. For other people, go to Q30.

29. Ego ole ka-ichoro ikwu ugbua maka neti nke gi?	**Dwnpay[ ]**

30. Neti ole ka I ga-egoro ndi ezi n'ulo gi n'onu ego ottu neti bu 450 Naira?	**#netsHH [ ]**

[Instruction: if the respondent does not want to buy the nets for other household members at 450 Naira per net ask why (Q31) and go to Q32, but if willing to buy, go to Q33].

31. Gini mere na I choghi igo net maka ndi ezi n'ulo gi? **(Instruction: After the answers are given, go to number 32): Reasons for refusing to buy:**
Code for reasons the respondent refused to buy. 1 = Eee 0 = Mba

31a = Ego adighi 31b = Achoghim ya 31c = Enweghim bed (akwa) 31d = Onwere ihe ndi ozo (Kwupute)	**Nomoney [ ]** **Nointerest[ ]** **No bed [ ]** **Others[ ]**

32. N'agbanyere n'inweghi ike ikwu 450 Naira bu ihe a n'ere ya, mana ichoro igoro neti maka gi onwe gi. Ego ole kachasi ka I nwere mmasi ikwu?

25a. Oburu na ihe a ga-ere out netia bu 440 Naira; I nwere mmasi ikwu ego ole a? 1 = Eee(25b) 0 = Mba (Q25b)	**Bin1 [ ]**
25b. Kedu kwanu ego ole kachasi I nwere mmasi ikwu? ***[Instruction; no mater the answer go to 25c]***	**BWTP1[ ]**
25c. Ihe a ga-ere netia bu 450 Naira, I nwere mmasi ikwu? 1 = Eee 0 = Mba	**Bin2 [ ]**

[Instruction: Go to question 33].

[Instruction: Fill in question 33 from your register before starting the interview]

33. Fill in the amount the respondent stated that s/he was willing to pay for ITNs for other household members in the first survey. 33a. Code same amount or different. 0 = same 1 = different	**BWTPOrig2** **[ ]** **WTPDiff2 [ ]**

[Instruction: If inital and present WTP amounts differ, go to question 34 and ask the respondent why s/he changed his/her mind. However, if the same go to question 35].

Code for reasons the respondent changing his/her mind. 1 = Eee 0 = Mba

34. Ego ole ikwuburu n'iga akwu maka ottu netia n'onwa gara aga agbara gi ajuju di iche n'ego ole ikwuru tata. Biko, kedu ihe mere igiri wee gbanwe uche gi banyere ego ole I nwere mmasi ikwu?	**CODES** **Nomoney [ ]** **Nointerest[ ]** **No bed [ ]** **Others[ ]**

35. Kedu, otu osiri di gi mma ikwu ugwo neti maka ndi ezi n'ulo gi? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 1 = Ikwubinye ugwo kwa izu uka, ma kwusia ugwo a n'ime izu ano. 2 = Ikwuzu ugwo otu mgbe ma buru neti ozigbo 3 = Onwere ottu ozo (Kwupute)	**PaymntM2[ ]**

Instruction: Ask Q36 to only those willing to purchase the net at 450Naira. For other people, go to Q37.

36. Ego ole ka-ichoro ikwu ugbua maka neti ndi ezi n'ulo gi?	**Dwnpay2[ ]**

[Instruction: Fill in question 37 from your register before starting the interview]

37. Oge mbu agbara gi ajuju, I kwere nkwa na-iga enwe ike tunye _____ Naira ka a were gotara ndi ogbenye neti agbara ogwu. Obu ndi committee ndi obodo roputara go ga-ahu ka aga esi were ego a ga-atuta were gotara ndi ogbenye neti. I nwere mmasi ikwu ego ole I kwere n'nkwa ugbua? 1 = yes (Q40) 0 = No (Q38)	**AaltWTP[ ]**

38. Gini mere na I choghi zi itunye ego a I kwere n'nkwa gi? **(Instruction: After the answers are given, go to number 40): Reasons for refusing to contribute:**

Code for reasons the respondent refused to buy. 1 = Eee 0 = Mba

38a = Ego adighi 38b = Achoghim ya 38c = Enweghim bed (akwa) 38d = Onwere ihe ndi ozo (Kwupute)	**Nomoney [ ]** **Nointerest[ ]** **No bed [ ]** **Others[ ]**

[Instruction: After question 38 go to question 39 if respondent wishes to contribute but not at the fixed price of the net, but if not willing to but at all, go to question 40]

39. O buru n'inwere mmasi itunye ego, mana obu onu ego ozo, gini bu ego ole inwere mmasi itunye ugbua?	**NaltWTP** **[ ]**

[Instruction: Fill in question 40 from your register before starting the interview]

40. Fill in the amount the respondent stated that s/he was willing to pay for altruistic WTP in the first survey. 40a. Code same amount or different. 0 = same 1 = different	**HWTPOrig** **[ ]** **ALTDiff [ ]**

[Instruction: If initial and present amounts differ, go to question 41 and ask the respondent why s/he changed his/her mind. However, if the same go to question 42].

Code for reasons the respondent changing his/her mind 1 = Eee 0 = Mba

41. Ego ole ikwuburu n'iga atunye n'onwa gara aga agbara gi ajuju di iche n'ego ole ikwuru tata. Biko, kedu ihe mere igiri wee gbanwe uche gi banyere ego ole I nwere mmasi itunye?	**Nomoney [ ]** **Nointerest[ ]** **No bed [ ]** **Others[ ]**

42. Kedu, etu osiri di gi mma isi were tunye ego ole I kwuru? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 1 = Ikwubinye ugwo kwa izu uka, ma kwusia ugwo a n'ime izu ano. 2 = Ikwuzu ugwo ottu mgbe 3 = Onwere ottu ozo (Kwupute)	**PaymntA[ ]**

43. Ego ole ka-ichoro ikwu ugbua maka neti nke gi?	**DwnpayA[ ]**

SECTION C: OTHER SOCIO-ECONOMIC FACTORS

44. Ego ole ka ezi n'ulo mefuru n'igota ihe oriri n'ahia n'ime out izu uka gara-aga site n'ihe oriri ndia m ga-aguputara gi?

Ihe egotara	Oha anuaa	Ego ole	Onye gotara	CODE (buyfd)
44a. Gari				**Garri**
44b. Agwa				**Beans**
44c. Abacha				**Cassava**
44d. Fio-Fio				**Fio-Fio**
44e. Ede				**Ede**
44f. Ji				**Yam**
44g. Osikapa				**Rice**
44h. Oka				**Corn**
44i. Azu				**Fish**
44j. Anu				**Meat**
44k. Akwukwo nri				**Veg**
44l. Onwere ihe ndi ozo (guputa)				**Others**
44m.Total				**Buyfd**

45. Oburu n'onwere nri ezi n'ulo gi meputara ma rikwaa n'izuka gara-aga, onu ego ole ka ha ga-aputa ma oburu n'ere ha n'ahia?

Ihe e meputara, e riri kwaa	Oha anuaa	Ego ole	CODE (Profd)
45a. Gari			**Garri**
45b. Agwa			**Beans**
45c. Abacha			**Cassava**
45d. Fio-Fio			**Fio-Fio**
45e. Ede			**Ede**
45f. Ji			**Yam**
45g. Osikapa			**Rice**
45h. Oka			**Corn**
45i. Azu			**Fish**
45j. Anu			**Meat**
45k. Akwukwo nri			**Veg**
45l. Onwere ihe ndi ozo (guputa)			**Others**
45m.Total			**Profd**

46. Onwere ka etu gi ma obu ndi ezi n'ulo gi siri wee mefue nnukwu ego n'onwa gara-aga? 1 = Eee 0 = Mba	**Mexpend [ ]**

47. Kedu etu ego gi ma obu ndi ezi n'ulo gi mefuru n'ihe ndia m ga-aguputara gi ha?

Item	Ego ole	Onye kwuru
47a. Inu di ma obu nwunye	**Marriage**	
47b. Ikwa ozu	**Funeral**	
47c. I gwo oria (n'abughi iba)	**Healthcare**	
47d. I mu nwa	**Birth**	
47e. Ugwo akwu kwo	**Schoolfees**	
47f. Igo oru	**Labourers**	
47g. I kuputa nwa	**NamingC**	
47h. Oriri mme mme	**Festivals**	
47i. Ezurum (anyi) Ohi	**Robbery**	
47j. Odachi / Ihe ntumadi (oghom)	**Disasters**	
47k. Ije n'ulo Ikpe/ nke ndi uwe oji (court/police cases)	**Cases**	
47l. Izo ani	**Landdis**	
47m. Onwere ihe ndi ozo (biko kwupute)	**Others**	
Total	**Majorexp**	

SECTION D: Special change factors

Oga amasim imata ihe mmetuta ihe ndia mga enyocha nwere n'ego ole I nwere mmasi ikwu ugbua maka netia agbara ogwu, ma itule ya n'onu ego ikwuburu mgbe agbara gi ajuju onu ikpeazu?

I. Income effect

Question	Code	Biko kowa aziza gi
48. Na mgbakota, ego ezi n'ulo gi n'enweta, otinyere etinye ka obelatara, ka oha etu oha site nime oge ikpeazu agbara gi ajuju onu rue ugbua? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = opelatara (Q49) 1 = otinyere (Q49) 2 = oha etu oha (Q50)	**IncomCh** [ ]	
49. Kedu mmetuta mgbanwe ego ole gi na ndi ezi n'ulo gi na enweta nwere n'ego ole inwere mmasi ikwu ugbua maka netia, ma itule ya n'oge ikpeazu agbara gi ajuju onu rue ugbua? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = opelatara 1 = otinyere 2 = onweghi mmetuta	**Incomef** [ ]	

II. Bed effect

Question	Code	Biko kowa aziza gi
50. I goro akwa ohuru site nime site nime oge ikpeazu agbara gi ajuju onu rue ugbua? 1 = Eee (Q51), 0 = Mba(52)	**Beds** [ ]	
51. Kedu mmetuta onwere n'ego ole inwere mmasi ugbua ikwu maka netia, ma itule ya n'oge ikpeazu agbra gi ajuju onu rue ugbua? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Opelatara 1 = Otinyere 2 = Onweghi mmetuta	**Bedeff** [ ]	

III. Effect of WTP technique, knowledge and attitude.

Factor	Effect	Biko kowaa aziza gi
52. Inwela ihe dika ottu onwa iji wee chee ego ole inwere mmasi ikwu maka netia agbara ogwu. Ya mere, kedu mmetuta ogea ijiri chee uche nwere n'ego ole inwere mmasi ikwu ugbua? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Opelatara; 1 = Otinyere; 2 = Onweghi mmetuta.	**Think** [ ]	
53. Ole ottu ajuju onu agbara gi ihe kari ofu ugboro ji	**Interv**	

metuta ego ole inwere mmasi ikwu ugbua maka netia agbara ogwu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Opelatara; 1 = Otinyere; 2 = Onweghi mmetuta.	**[ ]**	
54. N'echiche gi, kedu mmeso ndi ka n'onu ogu ndi ndu obodo nwere n'amuma netia agbara ogwu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Ajo mmeso (Q55) 1 = Ezi mmeso (Q55) 2 = A maghim(Q56)	**AttitdCL [ ]**	
55. Ole ottu mmeso ndi ndu obodo n'amuma netia agbara ogwu ji metuta ego ole inwere mmasi ikwu ugbua maka netia agbara ogwu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Opelatara; 1 = Otinyere; 2 = Onweghi mmetuta.	**A.ef [ ]**	
56. Kedu udi akuko/ihe I n'anu maka netia agbara ogwu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Ajo akuko(Q57) 1 = Ezi akuko (Q57) 2 = Onweghi (Q58)	**Einfo[ ]**	**(here Q56 ask; kedu ihe inuru?)**
57. Ole mmetuta udi ihe inuru gbasara metia agbara ogwu si emetu ego ole inwere mmasi ikwu ugbua maka netia agbara ogwu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Opelatara; 1 = Otinyere; 2 = Onweghi mmetuta.	**E.ef [ ]**	
58. Ndi agbata obi gi, ha egotawo ka ha aka ga-egota netia agbara ogwu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Mba (Q59) 1 = Eee (Q59) 2 = Amaghim (Q60)	**Neib[ ]**	
59. Ole mmetuta maka ndi agbata obi igo ma-bou egoghi netia si metuta ego ole inwere mmasi ikwu ugbua maka netia agbara ogwu? **(Achorom ka iwere otu aziza site na ihe ato m ga-aguputara gi)** 0 = Opelatara; 1 = Otinyere; 2 = Onweghi mmetuta.	**N.ef [ ]**	
60. I nwere ike gwam ma onwere ihe ndi ozo mere site n'mgbe ikpeazu agbara gi ajuju wee rue ugbua metutara ego ole inwere mmasi ikwu ugbua maka netia a gbara ogwu?	**Others [ ]**	

61. Ole oge o n'ewe gi na nkeji, iji rute n'ebe a n'ere netia agbara ogwu ma-ijiri ottu ijikariri eme njem?	**SALESD [ ]**

Aku n'uba ezi n'ulo

62. Site n'ime ihe ndi nga aguputara gi, ole ndi ezi n'ulo gi nwere? Biko, zaa eee ma obu mba. 1 = eee 0 = Mba	
62a. Igwe n'ekwu okwu (Radio)	**Radio[ ]**
62b. Anyinya igwe (Bicycle)	**Bicycle[ ]**
62c. Igwe aka eji akwo ihe (Manual grinding machine)	**Gmachine[ ]**
62d. Ogba tum tum (Motor cycle)	**Mcycle[ ]**
62e. Ugbo ala (Motor car)	**Mcar[ ]**
62f. Oche (Upholstery chairs/Cane chairs)	**Chairs[ ]**
62.g. Index movable assets	**IndexM[ ]**

63. Kedu udi ulo ezi n'ulo a bi na ya? **[Instruction; interviewer to fill in after observing the house]** 63a. Kedu ihe eji kpuchi elu uno a? 1 = Atanyi(Thatch) 2 = Zinc 3 = Cement 4 = Others (Please specify) 63b. Kedu ihe ejiri we ruo ulo a? 1 = Mud 2 = plastered Mud 3 = block 4 = others (please specify)	**Roof[ ]** **Wall[ ]**

65. (Instruction: ask only people not willing to buy the nets). Kedu ihe mere n'igotaghi netia agbara ogwu? **(A ga m aguputara gi ufodu aziza, ma chokwa ka-iza eee ma obu mba. 0 = Mba 1 = Ece**

Reasons	Codes
1 = O gwuru agwu 2 = E jighim ego 3 = O galara onu 4 = A maghim ebe a n'ere ya 5 = Onwere ihe ndi ozo (Biko kwupute)	**Finished [ ]** **No cash [ ]** **Expensive [ ]** **Lackof K[ ]** **Others [ ]**

66. [Instruction: Ask respondent whether there is any other information s/he would like to give to help us understand how to use the ITNs to fight Iba in our community]

67. Time interview ended:

Name and signature of the interviewer:

Appendix 10: The English version of the questionnaire for the last survey

Dear respondent,

As you know, we are trained local workers sent by staff of the University of Nigeria Teaching Hospital (UNTH) Enugu. Remember that last month, we conducted a questionnaire survey on the feasibility of starting a community-based malaria control project here with a new tool called Insecticide-treated nets (ITNs). We have now started the sale of the nets that we told you about. However, we are again interested in your opinion about the whole project and the factors leading up to it. Therefore, we will like to ask you some similar questions as we did before though some may be a little different. Your answers will help us to finally design how the programme can be sustained in the long-term. All information given will be treated with utmost confidentiality. Your participation is voluntary.

Instruction: Please fill the next section at the beginning and at the end of the interview.

What is the name of the respondent' s community?

What is the respondent's home address_____

What is the name of the interviewer

When was the date of interview	**Dateint[ ]**
What time did the interview start	**Timestart[ ]**
What time did the interview end	**Timeend[ ]**

Instruction: Fill in question 26, 36 and 37 from your register before starting the interview.

Instruction: Please, enter the appropriate number representing the answer given in the spaces provided in both the question and code columns.

QUESTION	**CODE**
SECTION A: PERSONAL DATA (of respondent) I would like to again begin by asking a few questions about you and your household.	
1. What is your status in this household? [ ] 1 = household head. 0 = representative of household	**Status[ ]**
2. How many people live in this household? [ ]	**Livein[ ]**
3. How many of household members belong to each of the following age-groups? 1. Less than five years [ ]; 2. five to fifteen years[ ]; 3. More than 15 years	**<5 [ ]** **5-15 [ ]** **>15 [ ]**
4. Was there either any increase or decrease in the number of people living in your household since the last interview? 0 = decrease 1= no change 3 = increase	**Livein2[ ]**
5. What is your name? _____	
6. Sex [ ] 1 = male 0 = female	**sex[ ]**
7. How old are you?	**Age[ ]**
8. Did you go to school? 1 = yes 0 = no	**Schl [ ]**
9. What occupation is your major source of getting money? _____	**Occupation [ ]**
10. Are you married? 1 = yes 0 = no	**Marital [ ]**
11. Was there a birth of a new baby within the past one month? 1 = Yes 0 = No	**Birth [ ]**
12. Did any of your household members become pregnant within the past month? 1 = Yes 0 = No	**Preg [ ]**

SECTION B: Information on malaria and mosquito
Introduction:(Read out to the respondent)

Please, I would like to ask you some questions about Malaria.

Instruction: Please, enter the number representing the answer given in the spaces provided in both the question and code columns.

13. In your own view, do mosquitoes worry your household? 0 = No 1 = Yes	**Mworry [ ]**
13a. Has there been a difference in your opinion of mosquito nuisance since you were last interviewed? 0 = decreased 1 = increased 2= the same	**Mworry2[ ]**
14. What happened to your perception of mosquito nuisance to your household in the past one month? 0 = decreased 1 = increased 2= the same	**PR[ ]**
15. Do you think that anyone in your household stands a chance of getting malaria? 0 = No 1 = Yes 2 = Don't know	**Mchance[ ]**
16. How has your perception of the chance of anyone in your household's getting malaria changed in the past one month? 0 = decreased 1 = increased 2 = the same	**Mchance2[ ]**
17. Has there been a difference in rate of malaria attacks to your household since you were last interviewed? 0 = decrease 1 = Increase 2 = the same	**Moccur [ ]**

Treatment and prevention of malaria:
Now, I would like to ask you some questions on how your household spent money to prevent and treat malaria within the past month.

18. How many people within your household had malaria within the past month?	**Hadmal [ ]**
19. What was/were the age-groups of the people that had malaria? 1 = less than 5 years; 2 = 5 to 15 years; 3 = more than 15 years	**<5 [ ]** **5 - 15 [ ]** **> 15 [ ]**

20. In your own estimate, how much did your household spend within the last month to treat malaria suffered by household member 1 from the sources listed below?

Item	Number of visits	Amount spent on treatment	Amount spent of transportation	Who paid	Where was it bought	**CODE Treatmal1**
20a1.Chemist						
20a2. Hospital						
20a3. Clinic						
20a4. Dispensary						
20a5. Laboratory						
20a6. Herbal home						
20a7.Mixed dispensers						
20a8.Home dispensers						
20a9.Village health workers						
20a10.Spiritual healers						
20a11.Others (specify)						
20a12. Total						[ ]

20b. What was the age of the ill person? **AGEMAL1[ ]**

21. In your own estimate, how much did your household spend within the last month to treat malaria suffered by household member 2 from the sources listed below?

Item	Number of visits	Amount spent on treatment	Amount spent of transportation	Who paid	Where was it bought	**CODE Treatmal2**
21a. Chemist 21b. Hospital 21c. Clinic 21d. Dispensary 21e. Laboratory 21f. Herbal home 21g. Mixed dispensers 21h. Home dispensers 21i. Village health workers 21j. Spiritual healers 21k. Others (specify) 21l. Total						[ ]

21m. What was the age of the ill person? **AGEMAL2**[ ]

22. In your own estimate, how much did your household spend within the last month to prevent malaria/mosquito nuisance on the items listed below?

Item	Quantity	Price of one	Total amount spent	Who purchased	Where purchased	**CODE Prvntmal**
22a.Injection 22b.Tablets 22c.Syrup 22d.Body cream repellant 22e.Mosquito nets 22f.Insecticide spray 22g.Mosquito coils 22h.Herbs 22i.Others (specify) Total						[ ]

Transitional explanation (Read out to the respondent)

There may be other causes of Iba as you may have already said. However, mosquito bites is the major cause. These mosquitoes commonly bite in the night especially when people are sleeping. When a mosquito bites someone that has the malaria illness and later bites someone without the illness, it will transfer the germ collected from the ill person to the second person, who may now develop the illness. In this way, malaria is spread from person to person. ***(Stop and ask the respondent: Do you have any question about the way malaria is transferred from person to person?).***

SCENARIO (Read out to the respondent)

As you may be aware, mosquito coils, insecticide sprays, herbs etc are some of the ways used to prevent mosquito bites. Another method towards preventing mosquitoes from biting people is through the use of ordinary mosquito nets. A form of these nets is mosquito bed-nets. A problem with these ordinary nets is that mosquitoes still perch on them to bite people sleeping inside. Also, these ordinary nets neither kill nor chase away mosquitoes. However, there is a new system of treating mosquito nets with special chemicals. These treated nets are called Insecticide-treated nets (ITNs). They chase away and kill mosquitoes. They also ensure that the mosquitoes do not perch on the nets to bite somebody sleeping inside. These are as a result of the chemicals used to treat the nets. Studies around the world have shown that these treated nets can reduce the number of times one gets malaria. These nets also get rid of other insects like bed bugs, lice etc. ***(Stop and ask the respondent: Is the difference between simple and ITNs are clear to you, together with the advantages of ITNs?).***

The chemicals used to treat the nets are harmless to both adults and children. However, mild side effects of the insecticides occur mostly in the first week of use of the Insecticide-treated nets. These side effects may make the person to feel that he/she has catarrh. These symptoms occur especially on direct contact with the nets, but thereafter there would be no symptoms after the first week. These side effects do not occur amongst everybody using the nets but only in a few people; about one out of fifty people.

However, when symptoms do occur, simple treatment with a drug like paracetamol relieves them. ***(Stop and ask the respondent: Do you have any question?).***

It is a fact that most households spend a lot of money either treating or preventing malaria. The sufferers and people looking after them also lose valuable work-hours and school-hours. Thus, the treated nets will ultimately reduce the time lost due to malaria and constant money spent to treat it in your household.

The nets however need at least yearly re-treatment with the chemicals in order to maintain their effectiveness. This net re-treatment will be organised by a committee to be selected from your community in conjunction with the Enugu State Malaria Control Unit. Individual households can also buy the chemicals and re-treat their nets after receiving a small training. You will have to pay for the net re-treatment. ***(Stop and ask the respondent: Do you have any question?).***

We now have the treated nets and we hope that you can purchase them for both yourself and other members of your household. This is because the nets will not be given free of charge. The bed-nets will be sold only to those who have offered to pay enough money to cover the price of the nets. Trained community members will sell the treated nets at either central locations in your community and or door-to-door. You can buy the nets now from us. You can make weekly instalment payments, but total payments should be made in a maximum period of four weeks. You shall receive the nets after you have paid fully. However, cash-and-carry system is encouraged. Your answers to the following questions will be used to make decisions about design of the programme.

23. The net is now available and the price that has been set is 350 Naira per net. Are you willing to purchase one for your personal use at this price? (please think of yourself for now, as we will ask you about your willingness to pay for other household members later) **(Show the net to the respondent)** 1 = yes (go to Q26) 0 = no (go to Q24)	**WTPYN1**

(Instruction: If willing to purchase a net at the stated price, go to question 26, but if not willing to buy, go to question number 21a. However, if not willing to pay for self but for other household members go to Q30).

24. What are all the reason(s) you would not buy? **(Instruction: After the answers are given, go to number 26)** Reasons for refusing to buy _____

Code for reasons the respondent refused to buy for him/herself. Codes: 1 = yes 0 = No

24a = Lack of money	No money [ ]
24b = Lack of interest	No interest [ ]
24c = Lack of bed	No bed [ ]
24d = Others (please specify)	Others [ ]

25. Despite the fact that you are unable to pay the 350 Naira per net sale price, but you are still willing to buy the nets for your personal use, what is the maximum amount that you are willing to pay for the net? **(Instruction: start at question 25a to elicit the amount the respondent is willing to pay).**

25a. What if the price of a net is 340 Naira; are you willing to pay? 1 = Yes (Q25b) 0 = No (25c)	**Bid1 [ ]**
25b. What if the price is 350, will you be willing to pay? (go to Q25d)	**Bid2 [ ]**
25c. What if the price is 330, will you be willing to pay? (go to Q25d)	**Bid 3[ ]**
25d. What is the maximum amount you are willing to pay for a net?	**WTP1[ ]**

[Instruction: Fill in question 26 from your register before starting the interview]

26. Fill in the amount the respondent stated that s/he was willing to pay for personal ITNs in the first survey.	**BWTPOrig [ ]**
26a. Code same amount or different. 0 = same 1 = different	**WTPDiff [ ]**

[Instruction: If first and present WTP amounts differ, go to question 27 and ask the respondent why s/he changed his/her mind after reminding the person his/her WTP amount from the first survey. However, if the same go to question 28].

Code for reasons the respondent changing his/her mind 1 = Yes 0 = No

27. The amount you said that you were willing to pay per net when you were interviewed last month is different from the one you just stated. Please, why did you change your mind on the amount you were willing to pay per net?	**Nomoney [ ]** **Nointerest[ ]** **No bed [ ]** **Others[ ]**

28. How would like to pay for the net for your personal use? (I'll read out some options, and you should select only one of them) 1 = weekly instalment payment, with total payments not exceeding 1 month 2 = cash and carry 3 = others (please specify)	**PaymntM** **[ ]**

[Instruction: Ask Q29 to only those willing to purchase the net at 350Naira. For other people, go to Q30].

29. How much payment are you making now for the net for your personal use?	**Dwnpay** **[ ]**

30. How many nets are you going to purchase for your other household members now at the price of 350 Naira per net?	**#netsHH [ ]**

[Instruction: if the respondent does not want to buy the nets for other household members at 450 Naira per net ask why (Q31) and go to Q32, but if willing to buy, go to Q33].

31. What are all the reason(s) you would not buy ITNs for your other household members? **(Instruction: After the answers are given, go to number 32)**
Reasons for refusing to buy _____

Code for reasons the respondent refused to buy for other household members. Codes: 1 = yes 0 = No

31a = Lack of money	No money [ ]
31b = Lack of interest	No interest [ ]
31c = Lack of bed	No bed [ ]
31d = Others (please specify)	Others [ ]

32. Despite the fact that you are unable to pay the 350 Naira per nete, but you are still willing to buy the nets for your other household members, what is the maximum amount that you are willing to pay for the net? (

32a. What if the price of a net is 340 Naira; are you willing to pay? 1 = Yes (Q32b) 0 = No (32c) 32b. What if the price is 350, will you be willing to pay? (go to Q32d) 32c. What if the price is 330, will you be willing to pay? (go to Q32d) 32d. What is the maximum amount you are willing to pay for a net?	 **Bid1 [ ]** **Bid2 [ ]** **Bid 3[ ]** **WTP1[ ]**

[Instruction: Fill in question 33 from your register before starting the interview]

33. Fill in the amount the respondent stated that s/he was willing to pay for other household members' ITNs in the first survey. 33a. Code same amount or different. 0 = same 1 = different	**BWTPOrig** **[ ]** **WTPDiff [ ]**

[Instruction: If first and present WTP amounts differ, go to question 34 and ask the respondent why s/he changed his/her mind after reminding the person his/her WTP amount from the first survey. However, if the same go to question 35].

Code for reasons the respondent changing his/her mind 1 = Yes 0 = No

34. The amount you said that you were willing to pay per net for other household members when you were interviewed last month is different from the one you just stated. Please, why did you change your mind on the amount you were willing to pay per net?	**Nomoney [ ]** **Nointerest[ ]** **No bed [ ]** **Others[ ]**

35. How would like to pay for the net for other household members'? (I'll read out some options, and you should select only one of them) 1 = weekly instalment payment, with total payments not exceeding 1 month 2 = cash and carry 3 = others (please specify)	**PaymntM2** **[ ]**

[Instruction: Ask Q36 to only those willing to purchase the net at 350Naira. For other people, go to Q37].

36. How much payment are you making now for the net for your personal use?	**Dwnpay** **[ ]**

[Instruction: Fill in question 37 from your register before starting the interview]

37. In the last survey, you stated that you were willing to pay ------ Naira so that the nets could be purchased for the households or individuals in the community that may be too poor to purchase the insecticide-treated bed nets. ***(instruction: fill in the amount from the result of the first survey)*** Are you willing to pay that amount now? 1 = yes (Q40) 0 = no (Q38)	**AaltWTP** **[ ]**

38. What are all the reason(s) you would not contribute the amount you had promised in the first survey?
(Instruction: After the answers are given, go to number 40)
Reasons for refusing to contribute _____

Code for reasons the respondent refused to contribute. Codes: 1 = yes 0 = No

38a = Lack of money 38b = Lack of interest 38c = Lack of bed 38d = Others (please specify)	**No money [ ]** **No interest [ ]** **No bed [ ]** **Others [ ]**

[Instruction: After question 38 go to question 39 if respondent wishes to contribute but not at the fixed price of the net, but if not willing to but at all, go to question 40]

39. If you are still willing to contribute for the poor, but a different amount, what is the maximum amount that you are willing to contribute now?	**NaltWTP** **[ ]**

[Instruction: Fill in question 40 from your register before starting the interview]

40. Fill in the amount the respondent stated that s/he was willing to pay for altruistic WTP in the first survey. 40a. Code same amount or different. 0 = same 1 = different	**HWTPOrig** **[ ]** **ALTDiff [ ]**

[Instruction: If initial and present amounts differ, go to question 41 and ask the respondent why s/he changed his/her mind. However, if the same go to question 42].

Code for reasons the respondent changing his or her mind. Codes: 1 = yes 0 = No

41. The amount you said that you were willing to contribute when you were interviewed last month is different from the one you just stated. Please, why did you change your mind?	**No money [ ]** **No interest [ ]** **Others [ ]**

42. How would like to pay for the nets in order to help the poor people? (I'll read out some options, and you should select only one of them) 1 = weekly instalment payment, with total payments not exceeding 1 month 2 = cash and carry 3 = others (please specify)	**PaymntA** **[ ]**

43. How much payment are you now making for the nets for the poor people?	**DwnpayA** **[ ]**

42. (Instruction: ask only people not willing to buy the nets). Why did you not buy the nets?

Reasons	Codes
1 = it was finished	**Finished [ ]**
2 = I had no cash	**No cash [ ]**
3 = Its too expensive	**Expensive [ ]**
4 = I did not know where to buy	**Lackof K[ ]**
5 = others (please specifiy)	**Others [ ]**

SECTION C: Other socio-economic factors

44. How much did your household spend to purchase of food from the market in the past one week on the various items that I will read out?

Item	Quantity	Amount	Who purchased	Buyfood
44a.Gari				
44b.Beans				
44c.Cassava (akpu)				
44d.Fio-Fio				
44e.Cocoyam (ede)				
44f.Yam				
44g.Rice				
44h.Corn				
44i.Fish				
44j.Meat				
44k.Vegetables				
44l.Others (specify)				
44m.Total				[ ]

45. If the food items that your household produced that it also consumed in the past one week were bought from the market, how much will they cost?

Item	Quantity	Amount	Who purchased	Profd
45a.Gari				
45b.Beans				
45c.Cassava (akpu)				
45d.Fio-Fio				
45e.Cocoyam (ede)				
45f.Yam				
45g.Rice				
45h.Corn				
45i.Fish				
45j.Meat				
45k.Vegetables				
45l.Others (specify)				
45m.Total				[ ]

46. Did you your household make major expenditures within the past one month? 1 = yes 0 = no	**Mexpend** [ ]

47. If yes, on what items and what was the total amount in the itema that I will read out to you?

Item	Total amount	Who paid
47a.Marriages		
47b.Funeral		
47c.Healthcare		
47d.Birth		
47e.School fees		
47f.Hiring of labourers		
47g.Naming ceremony		
47h.Festivals		
47i.Robbery		
47j.Natural disasters		
47k.Court/ police cases		
47l.Land disputes		
47m.Others (please specify		
Total (Majorexp)		

SECTION D: Special change factors

Please, I would like to find out the effect that each of the following factors that I will ask you about have on the amount that you are now willing to pay per net compared to the amount you stated during the last survey?

I. Income effect

48. What changes occurred in your total household income within the past month? 0 = decreased (Q49) 1 = increased (Q49) 2 = remained the same (Q50).	**IncomCh** [ ]	
49. What effect did the change in your household income have on the cash availability have on your willingness to pay for ITNs now as compared to one month ago? 0 = decreased 1 = increased 2 = no effect	**Incomef** [ ]	

II. Bed effect

50. Did you buy a new bed within the past month? 1 = Yes (Q51)) 0 = No (Q52)	**Beds** [ ]	
51. What effect does on your willingness to pay for ITNs now as compared to one month ago? 0 = decreased 1 = increased 2 = no effect	**Bedeff** [ ]	

III. Effect of WTP technique, knowledge and attitude

Factor	Effect	Please explain
52. You have had about one month to think about your willingness to pay for the ITNs. So what effect has this time to think had on the amount you are now willing to pay? (I'll read out some options, and you should select only one of them) 0 = decreased 1 = increased 2= no effect	**Think** [ ]	
53. How has the fact that you have been interviewed more than once influenced the amount that you are now willing to pay for the ITNs? (I'll read out some options, and you should select only one of them) 0 = decreased 1 = increased 2= no effect	**Interv** [ ]	
54. In your opinion, what is the attitude of a majority of the community leaders to the ITNs project? (I'll read out some options, and you should select only one of them) 0 = negative (Q55) 1 = positive (Q55) 2 = do not know (Q56)	**AttitdCL** [ ]	
55. How has the attitude community leaders to the ITNs project affected your of WTP for the ITNs? 0 = decreased 1 = increased 2 = no effect	**Attidef** [ ]	
56. What nature of information have you been receiving regarding the ITNs? (I'll read out some options, and you should select only one of them). 0 = negative (Q57) 1 = positive (Q57) 2 = no information (Q58)	**Einfo** [ ]	**(here Q56 ask what have you heard)**
57. How has the information you have been receiving regarding the ITNs affected your WTP for the ITNs? (I'll read out some options, and you should select only one of them) 0 = decreased 1 = increased 2 = no effect	**Einfoef** [ ]	
58. Have/will your neighbours buy the nets? (I'll read out some options, and you should select only one of them) 0 = No (Q59) 1 = Yes (Q59) 3 = do not know (Q60)	**Neib** [ ]	

59. How has the effect of whether or not your neighbours will buy the nets influenced your actual WTP? 0 = decreased 1 = increased 2 = no effect	**Neibef [ ]**	
60. Could you tell me whether there were other things that happened between the 1st and 2nd survey that affected the amount that you are willing to pay now?	**Others [ ]**	

61. How long does it take you to get to the sales point using your normal means of transport (in minutes)?	SALESD [ ]

Assets of the household

62. From the assets that I am going to read out to you, could please tell me the ones that your household has? Please answer yes or no. 1 = yes 0 = No	
62a. Radio	**Radio[ ]**
62b. Bicycle	**Bicycle[ ]**
62c. Manual grinding machine	**Gmachine[ ]**
62d. Motor cycle	**Mcycle[ ]**
62e. Motor car	**Mcar[ ]**
62f. Upholstery chairs/Cane chairs	**Chairs[ ]**

42. What type of housing does the household live in? **(interviewer to fill in after observing the house)**	
63a. What type of roof? [ ] 1 = Thatch 2 = Zinc 3 = Cement 4 = others (please specify)	**Roof [ ]**
63b. What type of wall are the rooms made of [ ] 1= mud 2 = plastered mud 3 = block 4 = others (please specify)	**Wall [ ]**

65. **(Instruction: ask only people not willing to buy the nets).** Why did you not buy the nets?

Reasons	Codes
1 = It was finished	**Finished [ ]**
2 = I had no cash	**No cash [ ]**
3 = Its too expensive	**Expensive [ ]**
4 = I did not know where to buy	**Lackof K[ ]**
5 = others (please specify)	**Others [ ]**

66. [Instruction: Ask respondent whether there is any other information s/he would like to give to help us understand how to use the ITNs to fight Iba in our community]

67. Time interview ended

Name and signature of interviewer:

Appendix 11: Structured haggling technique (after content validation)

1. The price of a net is 550; are you willing to pay? 1 = Yes (Q7) 0 = No (Q2) Do not know (Q2)

2. What is the maximum amount you are willing to pay? (Interviewer: if more or equal to 450 Naira go to Q3, but if less than 450 Naira, go to Q4)

3. What if the price is 540 Naira, will you be willing to pay? 1 = yes 0 = No (Interviewer: no matter the answer, go to Q7).

4. What if the price is 490 Naira, will you be willing to pay? 1 = yes (Q7) 0 = No (Q5)

5. What really is the maximum amount you are willing to pay for a net? (Interviewer: If more or equal to 450 Naira go to Q7, but if less than 450 Naira go to Q6)

6.The amount that you have quoted is too low, and cannot cover the cost of the net, and so you will have to increase the amount if you really want to buy the net. So what is the final maximum amount you are willing to pay for a net? (Interviewer: No matter the answer, go to Q7)

7. If due to inflation or other uncertainties, the price of the net increases, what is the maximum amount you are very certain to pay?

Appendix 12: The Binary with follow-up format

1. Are you willing to pay 550 for a net? 1 = Yes (Q2) 0 = No (Q2)

2. What is the maximum amount you are willing to pay for a net? (Interviewer: if the amount is less than 450 Naira, go to Q3).

3. The actual price of the net is 450 per one, are you willing to buy at this price? 1 = yes 0 = No

Appendix 13: Full and reduced Truncated regression models

	BG		BWFU		SH	
Variables	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff.(SE)	Reduced Coeff. (SE)
Status in the household	-.09 (.10)		-23.33 (24.44)		.22 (.11)*	.27 (.08)***
Number of household residents	-.02 (.02)	-.01 (.02)	-.96 (3.47)	.75 (3.21)	.02 (.02)	.02 (.02)
Sex	.18 (.10)*	.17 (.08)**	19.66 (20.96)		.06 (.14)	
Age	-.0002 (.003)		-.24 (.63)		.002 (.003)	
School years	-.01 (.01)		7.92 (2.30)***	8.12 (2.11)***	.01 (.01)	.01 (.01)
Marital status	.24 (.13)*	.31 (.12)**	-13.71 (24.31)		-.22 (.12)*	-.22 (.11)*
Perceived incidence of malaria	-.03 (.11)		.53 (19.73)		.03 (.09)	
Actual incidence of malaria	.08 (.10)		16.95 (17.28)	20.30 (16.29)	.14 (.06)**	.14 (.06)**
Dummy of exp. treat malaria	-.09 (.15)		-28.25 (30.91)	-37.93 (28.99)	-.17 (.13)	-.15 (.12)
Dummy of exp. to prevent malaria	.30(.09)***	.30 (.08)***	-18.79 (16.53)	-22.79 (16.30)	.07(.11)	
Actual incidence of other illnesses	.001 (.10)		-9.66 (12.66)	-8.84 (8.14)	.01 (.07)	
Dummy of exp. to treat other illnesses	-.22(.15)	-.24(.08)***	-.19 (23.54)		.02 (.13)	
Previous purchase of nets	.03(.10)		38.75 (20.83)*	34.93 (20.50)*	-.04 (.11)	
Ownership of radio	.08 (.10)		31.70 (21.83)	38.58 (21.32)*	.15 (.11)	.17 (.10)*
Ownership of bicycle	.02(.10)		17.60 (17.58)		-.08(.08)	-.10 (.08)
Ownership of grinding machine	.27 (.12)**	.29 (.11)***	-43.22 (24.21)*	-45.48 (22.58)**	.10 (.11)	
Ownership of motorcycle	.14(.10)	.16 (.10)	-33.68 (18.87)*	-25.75 (18.35)	-.04(.10)	
Ownership of car	-.31 (.17)*	-.29 (.16)*	24.74 (27.73)		.09(.14)	.19 (.10)
Log of Food cost	.08 (.07)	.13 (.06)**	13.51 (12.33)	13.53 (11.90)	.10 (.06)*	.10 (.06)*
Occup2	.07 (.22)		-66.63 (44.31)	-75.42 (44.20)*	-.16 (.16)	
Occup3	.20 (.22)		-57.96 (44.43)	-62.19 (44.81)	-.12 (.17)	
Occup4	.27 (.24)		-152.25 (52.75)***	-154.68 (52.63)***	-.13 (.19)	
Occup5	.29 (.31)		-151.82 (60.82)**	-153.91 (61.21)**	.18 (.23)	
Constant	4.36 (.53)***	4.05 (.43)***	200.03 (93.27)**	175.79 (88.51)**	4.42 (.46)***	4.34 (.38)***
Number of Obs.	198	198	214	215	204	204
Truncated observations	60	60	52	52	63	62
Chi square	60.57***	53.11***	47.21***	42.46***	66.98***	58.16***
Log likelihood	-146.137	-149.033	-1279.314	-1286.989	-154.725	-158.099
Sigma			100.867 (5.555)***	101.766 (5.607)***		
Normality		11.25**		6.51*		6.12*
Predicted values		5.18 (ref 5.21)		232.14 (ref 195.07)		5.29 (ref 5.36)
RESET		0.312		0.86		1.304

Significance of parameters * <.10, **<.05, ***<.01

The sign and magnitudes on the variables were similar to that from the OLS models (Appendix 13). The statistically significant variables were also similar across the three methods to the OLS model. Amongst the few differences with the OLS models, the actual malaria incidence and food cost were no longer statistically significant in the BWFU group. Also, the previous purchase of nets and occupational group 2 became statistically significant here and also had the expected signs. Similarly, ownership of

radio became statistically significant and with the right sign in the haggling group. The Bidding group remained unchanged. The reduced models for the three elicitation methods were all statistically significant ($p<0.01$).

RESET test showed that all the models were correctly specified. However, the residuals were not normally distributed. The models predicted raw WTP best in the bidding game model, while the BWFU had the worst predicted raw WTP. Plots of the residuals versus fitted values showed the absence of heteroscedasticity in the models.

Appendix 14: Full and reduced Tobit models

	BG		BWFU		SH	
Variables	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff. (SE)	Reduced Coeff.(SE)	Full Coeff.(SE)	Reduced Coeff. (SE)
Status in the household	.05(.42)		12.76 (27.44)		.54 (.45)	.66(.32)**
Number of household residents	.004 (.09)	-.01 (.08)	4.87 (4.27)	4.02 (3.98)	.01 (.07)	.01 (.06)
Sex	-.08 (.44)		18.81 (23.97)	23.93(19.42)	.32 (.46)	
Age	.003 (.011)		.03 (.77)		-.01 (.01)	
School years	-.01 (.04)		10.88(2.82)***	10.58 (2.56)***	.06 (.04)	.09(.04)**
Marital status	.03 (.59)		-28.28 (29.33)		-.62 (.49)	-.66 (.49)
Perceived incidence of malaria	-.54 (.48)		-13.98 (24.32)		.17 (.40)	
Actual incidence of malaria	.11 (.43)		48.74(21.99)**	23.45 (10.79)**	.24 (.28)	.19 (.18)
Dummy of exp. to treat malaria	.15 (.65)		-46.32 (37.80)		-.18 (.53)	
Dummy of exp. to prevent malaria	1.35 (.40)***	1.43(.39)***	-18.07 (20.40)		-.56 (.46)	-.47 (.43)
Actual incidence of other illnesses	-.20 (.45)		-16.53 (16.38)		.62 (.32)*	.61 (.31)*
Dummy of exp. to treat other illnesses	.32 (.65)		42.59 (28.77)		-.65 (.56)	-.67 (.54)
Previous purchase of nets	.57 (.46)	.54 (.46)	14.62 (26.37)		.16 (.47)	
Ownership of radio	.26 (.43)		47.13 (25.09)*	53.14 (24.27)**	.96 (.40)**	1.06 (.39)***
Ownership of bicycle	-.24 (.41)		10.22 (21.39)		.17 (.35)	.17 (.34)
Ownership of grinding machine	.93 (.54)*	.95 (.50)*	-50.61 (29.88)*	-46.37 (28.17)	.65 (.47)	.65 (.46)
Ownership of motorcycle	.53 (.47)	.54 (.46)	-35.97 (23.30)	-31.83 (22.56)	-.92 (.42)**	-.87 (.42)**
Ownership of car	-2.19 (.68)***	-2.11 (.67)***	40.47 (34.73)		.45 (.61)	
Log of Food cost	1.40 (.28)***	1.39(.27)***	41.53(13.96)***	41.76 (13.67)***	1.07 (.22)***	1.13 (.22)***
Occup2	1.33 (1.05)	1.59(1.01)	-39.99 (51.26)	-34.37 (50.38)	-.07 (.69)	
Occup3	1.06 (1.02)	1.23(1.00)	-15.63 (50.68)	-12.57 (50.35)	.27 (.71)	
Occup4	-2.10 (1.10)*	-1.90(1.08)*	-119.53 (59.23)**	-114.18 (57.90)*	.44 (.81)	
Occup5	.95 (1.42)	1.43 (1.36)	-157.93 (68.38)**	-169.11 (68.08)**	.45 (.99)	
Constant	-7.35(2.19)***	-7.28(2.02)***	-174.04 (105.52)	-188.10 (98.78)*	-4.11 (1.77)**	-5.03 (1.39)***
Number of Obs.	259	259	266	267	266	267
Censored Obs.	61	61	52	52	62	63
Uncensored Obs.	198	198	214	215	204	204
Chi2	68.77***	65.62***	79.37**	72.97**	108.59***	108.79***
Log likelihood	-529.070	-530.645	-1413.5416	-1422.7773	-530.209	-532.381
Normality		2.44		4.85		1.33
Predicted values		3.65 (ref 3.97)		178.77 (ref 195.07)		3.79 (ref 4.10)
RESET		3.278*		1.89		4.842*

Significance of parameters * <.10, **<.05, ***<.01

Majority of the statistically significant variables had the hypothesised signs (Appendix 14). There were three statistically significant variables in the bidding game that had the right signs and four in both the BWFU and the Haggling group. The log of food cost was uniformly statistically significant across the three elicitation methods, and it also

had the right positive sign. The number of years of schooling and ownership of radio were statistically significant in both the BWFU and Haggling groups, and with the right signs. However, the actual incidence of malaria was only statistically significant in the BWFU group.

The dummy of the expenditure to prevent malaria and ownership of grinding machine were also statistically significant in the bidding group, and had the right signs. Household heads were willing to pay more for ITNs in the Haggling group, and the sign was according to expectation. However, ownership of motorcar and occupational group 4 though statistically significant in the bidding group had the wrong hypothesised signs. Occupational groups 4 and 5 though statistically significant had the wrong hypothesised signs in the BWFU group. The same applies to actual incidence of other illnesses and ownership of motorcycle for the haggling group.

The coefficients showed that the larger the food cost, the more the level of WTP. In the BWFU group where the un-transformed WTP was used as the dependent variable, the coefficient of food cost at 41.53 shows that the level of WTP for an ITNs will increase by that amount if the log of food cost increase by one unit. The coefficients in both the bidding and haggling groups were more than one showing the considerable economic effect of food cost on WTP.

The economic effects of number of years of schooling, actual incidence of malaria and ownership of radio were considerable, using the BWFU group to illustrate. The result showed that one unit increase in the value of any of these variables would lead to a substantial increase in the level of WTP. The result of the bidding game showed that those with expenditures to prevent malaria and those with grinding machines were more willing to pay than their comparative groups. The haggling result showed that household heads were willing to pay more than their representatives.

The regression statistics showed that the Tobit models for the Bidding and Haggling groups were statistically significant at $p<0.01$, while the BWFU was statistically significant at $p<0.05$. The RESET test showed that only the Tobit model in the BWFU group was correctly specified.

Appendix 15: Demographic data of the respondents (inter-rater reliability)

	BIDDING		BWFU		HAGGLING	
Variables	1st Interviewer n (%)	2nd Interviewer n (%)	1st Interviewer n (%)	2nd Interviewer n (%)	1st Interviewer n (%)	2nd Interviewer n (%)
Status						
1 = Head	76 (69.7)	83 (76.1)	85 (77.3)	89 (80.9)	73 (71.6)	81 (79.4)
0 = Rep	33 (30.3)	26 (23.9)	25 (22.7)	21 (19.1)	29 (28.4)	21 (20.6)
Total	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	102 (100)
Household residents: Mean (S.D)	4.19 (2.15)	4.28 (2.21)	4.46 (2.50)	3.91 (2.23)	4.80 (2.84)	4.34 (2.40)
Sex						
1= Male	(45.9)	50 (45.9)	52 (47.3)	52 (47.3)	52 (51.0)	52 (51.0)
0 = Female	(54.1)	59 (54.1)	58 (52.7)	58 (52.7)	50 (49.0)	50 (49.0)
Total	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	102 (100)
Age (years): Mean (S.D.)	52.57 (16.43)	54.98 (15.99)	51.25 (14.47)	53.20 (16.40)	49.49 (14.52)	50.66 (14.29)
Attended school	51 (46.8)	56 (51.4)	62 (56.4)	66 (60.0)	63 (61.8)	63 (61.8)
1 = Yes	(53.2)	53 (48.6)	48 (44.6)	44 (40.0)	39 (38.2)	39 (38.2)
0 = No	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	102 (100)
Total						
Years of schooling: Mean (S.D)	2.80 (3.69)	2.91 (3.81)	3.87 (4.35)	4.00 (4.18)	4.47 (4.49)	5.38 (8.10)
Marital status						
1=Ever married	96 (88.1)	98 (89.9)	94 (85.5)	92 (83.6)	88 (86.3)	88 (86.3)
0 = Not married	13 (11.9)	11 (10.1)	16 (14.5)	18 (16.4)	14 (13.7)	14 (13.7)
Total	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	103 (100)
Radio						
1 = Yes	80 (73.4)	67 (61.5)	87 (79.1)	87 (79.1)	66 (64.7)	67 (65.7)
0 = No	29 (26.6)	42 (38.5)	23 (21.9)	23 (20.9)	36 (35.3)	35 (34.3)
Total	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	102 (100)
Bicycle						
1 = Yes	69 (63.3)	74 (67.9)	69 (62.7)	69 (62.7)	61 (59.8)	59 (57.8)
0 = No	40 (36.7)	35 (32.1)	41 (37.3)	41 (37.3)	41 (40.2)	43 (42.2)
Total	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	102 (100)
Grinding machine						
1 = Yes	14 (12.8)	13 (11.9)	9 (8.2)	14 (12.7)	13 (12.7)	16 (15.7)
0 = No	95 (87.2)	96 (88.1)	101 (91.8)	96 (87.3)	89 (87.3)	86 (84.3)
Total	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	102 (100)
Motor cycle						
1 = Yes	18 (16.5)	19 (17.4)	28 (25.5)	28 (25.5)	17 (16.7)	23 (22.5)
0 = No	91 (83.5)	90 (82.6)	82 (74.5)	82 (74.5)	85 (83.3)	79 (77.5)
Total	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	102 (100)
Motor car						
1 = Yes	9 (8.3)	9 (8.3)	8 (7.3)	9 (8.2)	9 (8.8)	8 (7.8)
0 = No	100 (91.7)	100 (91.7)	102 (92.7)	100 (91.8)	93 (91.2)	94 (92.2)
Total	109 (100)	109 (100)	110 (100)	110 (100)	102 (100)	102 (100)

Appendix 16: Demographic information (test-retest reliability)

	BIDDING		BWFU		HAGGLING	
Variables	1st Interview n (%)	2nd Interview n (%)	1st Interview n (%)	2nd Interview n (%)	1st Interview n (%)	2nd Interview n (%)
Status						
1 = Head	106 (72.6)	100 (68.5)	126 (78.3)	139 (86.3)	94 (67.6)	102 (73.4)
0 =	40 (27.4)	46 (31.5)	35 (21.7)	22 (13.7)	45 (32.4)	37 (26.6)
Representative	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)
Total						
Sex						
1= Male	63 (43.2)	63 (43.2)	77 (47.8)	77 (47.8)	67 (48.2)	67 (48.2)
0 = Female	83 (56.8)	83 (56.8)	84 (52.2)	84 (52.2)	72 (51.8)	72 (51.8)
Total	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)
Age (years):	51.88	50.45	50.05	50.40 (14.85)	48.61	50.53
mean (S.D.)	(16.82)	(14.18)	(14.46)		(14.64)	(15.04)
Attended school						
1 = Yes	(52.7)	73 (50.0)	102 (63.4)	95 (59.0)	83 (59.7)	90 (64.7)
0 = No	(47.3)	73 (50.0)	59 (36.6)	66 (41.0)	56 (40.3)	49 (35.3)
Total	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)
Marital status						
1= Ever married	130 (89.0)	118 (80.8)	138 (85.7)	142 (88.2)	116 (83.5)	113 (81.3)
0=Not married	16 (11.0)	28 (19.2)	23 (14.3)	19 (11.8)	23 (16.5)	26 (18.7)
Total	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)
Radio						
1 = Yes	104 (71.2)	102 (69.9)	127 (78.9)	131 (81.4)	102 (73.4)	96 (69.1)
0 = No	42 (28.8)	44 (30.1)	34 (21.1)	30 (18.6)	37 (26.6)	43 (30.9)
Total	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)
Bicycle						
1 = Yes	95 (65.1)	96 (65.8)	97 (60.2)	96 (59.6)	82 (59.0)	76 (54.7)
0 = No	51 (34.9)	50 (34.2)	64 (39.8)	65 (40.4)	57 (41.0)	63 (45.3)
Total	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)
Grinding machine						
1 = Yes	16 (11.0)	12 (8.2)	17 (10.6)	23 (14.3)	29 (20.9)	23 (16.5)
0 = No	130 (89.0)	134 (91.8)	144 (89.4)	138 (85.7)	110 (79.1)	116 (83.5)
Total	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)
Motor cycle						
1 = Yes	27 (18.5)	27 (18.5)	29 (18.0)	29 (18.0)	26 (18.7)	30 (21.6)
0 = No	119 (81.5)	119 (81.5)	132 (82.0)	132 (82.0)	113 (81.3)	109 (78.4)
Total	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)
Motor car						
1 = Yes	12 (8.2)	18 (12.3)	12 (7.5)	15 (9.3)	13 (9.4)	9 (6.5)
0 = No	134 (91.8)	128 (87.7)	149 (92.5)	146 (90.7)	126 (90.6)	130 (93.5)
Total	146 (100)	146 (100)	161 (100)	161 (100)	139 (100)	139 (100)

Appendix 17: Binary distribution of the two divergent WTP groups for test-retest reliability

	Bidding		BWFU		Haggling	
Variables	Did not change mind N=104 Mean (SD)	Changed mind N= 54 Mean (SD)	Did not change mind N=73 Mean (SD)	Changed mind N= 93 Mean (SD)	Did not change mind N=93 Mean (SD)	Changed mind N=51 Mean (SD)
Status in the household	.69 (.47)	.62 (.49)	.88 (.33)	.85 (.36)	.73 (.45)	.75 (.44)
Household residents	4.01 (2.63)	3.85 (1.96)	4.27 (2.76)	4.33 (2.58)	3.85 (2.06)	3.84 (1.77)
Sex	.45 (.50)	.33 (.48)	.47 (.50)	.49 (.50)	.46 (.50)	.55 (.50)
Age	51.55 (14.49)**	45.76 (13.12)	52.97 (13.76)*	48.71 (15.80)	51.15 (15.04)	49.22 (16.45)
Attended school	.50 (.50)	.56 (.50)	.55 (.50)	.60 (.49)	.59 (.49)	.73 (.45)
Marital status	.82 (.39)	.80 (.41)	.88 (.33)	.88 (.32)	.82 (.39)	.82 (.39)
Occup1	.02 (.14)*	.07 (.26)	.05 (.23)	.01 (.10)	.04 (.20)	.06 (.24)
Occup2	.65 (.48)	.57 (.50)	.66 (.48)	.62 (.49)	.46 (.50)	.45 (.50)
Occup3	.24 (.43)	.28 (.45)	.21 (.41)	.24 (.43)	.38 (.49)	.41 (.50)
Occup4	.04 (.19)	.06 (.23)	.01 (.12)**	.11 (.31)	.05 (.23)**	.08 (.27)
Occup5	.05 (.21)	.02 (.14)	.07 (.25)	.04 (.20)	.08 (.27)	.00 (.00)
Actual malaria incidence	.42 (.63)	.50(.64)	.73 (.89)	.87 (1.07)	.32 (.68)	.45 (.70)
WTP self	144.33 (124.33)***	199.07 (119.15)	165.41 (129.43)	255.70 (129.86)	165.00 (138.49)	253.14 (142.98)
Total expenditure	3635.74 (7925.31)	3888.89 (7481.69)	4370.68 (9816.22)	5799.57 (15141.16)	5814.95 (15454.50)	3985.10 (10444.06)
Sales distance	15.79 (11.99)	16.80 (12.79)	14.34 (14.88)	14.03 (13.13)	16.58 (12.10)	13.55 (10.53)
Food cost	770.08 (454.50)	824.07 (980.53)	1344.45 (1680.23)	1009.18 (665.71)	894.52 (532.77)	928.73 (696.28)
Radio	.68 (.47)	.78 (.42)	.75 (.43)	.85 (.36)	.67 (.47)	.75 (.44)
Bicycle	.68 (.47)	.61 (.49)	.62 (.49)	.58 (.50)	.52 (.50)	.63 (.49)
Grinding machine	.09 (.28)	.07 (26)	.12 (.33)	.15 (.36)	.17 (.38)	.14 (.35)
Motorcycle	.19 (.40)	.13 (.34)	.19 (.40)	.17 (.38)	.19 (.40)	.27 (.45)
Car	.13 (.33)	.11 (.32)	.07 (.25)	.11 (.31)	.10 (.28)	.02 (.14)

Significance of differences between groups * <0.10, **<0.05, ***<0.01