Measuring empowerment in rural India using vignettes

Many development projects aim at empowering women. Yet, the measurement of empowerment has proved difficult and is riddled with technical and conceptual problems. Current approaches rely on long questionnaires and, to some extent, on subjective perceptions which are not comparable across groups. In this paper we propose a method for measuring empowerment based on vignettes and provide an application to a sample of rural women in Andhra Pradesh. The method is simple to administer and addresses biases in subjective perceptions. We show how perceptions vary systematically across groups and how they can be corrected for. We also show how the impact of a project on empowerment can be tested. In our application we find that differences in self-reported empowerment are perceptual and that a self-help group intervention does not increase women empowerment.

Introduction

Gender inequalities in living standards are widespread. For example, in India, preference for sons is such that females are under-represented among births and over-represented among deaths (Sen 1990). Girls in India are also less likely to attend secondary school, they marry at an early age, have limited access to media, markets and health care, and do not have major say in household decision-making (Kishor and Gupta 2009). It is therefore not surprising that governments, international organisations and NGOs are trying to readdress gender imbalances through projects that "empower women". Yet it is

1

far from clear what progress these interventions have made in general (Duflo 2012), and more specifically in the context of microfinance programmes (Vaessen et al. 2014), not least because of the difficulty to measure empowerment.

"Empowering women" is the third millennium development goals but the concept of empowerment is poorly defined and difficult to measure. Solava and Alkire (2007) found 29 different definitions of empowerment in the literature and many others were certainly produced since their review. Empowerment is multidimensional and contextually determined so that it defies a unique definition. There are then technical problems in the measurement of empowerment, which have been reviewed, among others, by Malhotra et al. (2002). Problems include multidimensionality (what dimensions of empowerment should we consider?), aggregation (how should we combine different dimensions?), and dynamics (how can we observe changes over time?). To address these problems, researchers have resorted to the use of theory-based indices (see in particular Solava and Alkire 2007; Alsop, Bertelsen, and Holland 2006). These indices aggregate empowerment measures over several dimensions using data collected through carefully designed survey questionnaires. However, despite their attempts to measure empowerment in an objective way, these indices have limitations. First, they are static, while empowerment is a process, a transition from one state to another that does not need to happen gradually (Eyben, Kabeer, and Cornwall 2008). Second, much of the information used to build the indices consists of subjective perceptions. Women responding to survey questionnaires may have 'adaptive preferences', or 'false consciousness' or being influenced by the prevailing 'doxa' (Kabeer 2001). Perceptions of empowerment may vary greatly and systematically across groups and countries so as to make comparisons impossible. For example, marginalised women might perceive themselves as more empowered than they actually are in comparison to other women.

In this paper we propose a method to measure empowerment that addresses many of these methodological limitations. We measure empowerment using anchoring vignettes (King et al. 2004), a method that elicits subjective evaluations of empowerment correcting for systematic bias in responses across groups. The method is easily administered within a household survey; it is able to capture static as well as dynamic dimensions; and can generate measures of empowerment over different dimensions. We provide an application of the method using data from a rural development project in Andhra Pradesh to illustrate its potential in the evaluation of empowerment programmes.

Measuring empowerment using vignettes

Similarly to concepts like 'social capital', 'destitution', and 'resilience' – just to name a few – 'empowerment' defies clear definitions. In response to this, researchers wanting to measure empowerment have relied on the careful design of survey questionnaires in order to cover its many facets. Since the definitions of empowerment, as well as the theory informing the definitions, differ from study to study, different studies have measured different constructs and their results are not comparable. In addition, most empowerment indices are based on some form of subjective self-assessment so that empowerment levels of different people or groups of people are not comparable. People interpret the same question in different ways and attach different meanings to the same construct. Differences in interpretation may be particularly large across countries and socio-economic groups. This is a well-known problem in the interpretation of self-reported morbidity and disability, whereby, for example, people of richer countries tend to report lower health status when we know this is not true (Sen 2002). Similarly, women of high socio-economic status may, for example, report levels of empowerment that are higher than those they actually experience.

Subjective responses however become comparable if they are linked to a scale which is interpreted in the same way by all respondents. If such a scale is found it is said to provide an *anchor* to subjective responses. Vignettes are an effective method of *anchoring* subjective responses. In the vignette methodology respondents are first asked for a self-assessment of a particular condition, for example happiness, using a scale from high to low. Respondents are then asked to rate hypothetical individuals (the vignettes) using the same scale used in the self-assessment. The goal of the exercise is to pin down what the respondents mean, when they state they are happy or unhappy, by looking at the way they rate a particular happy or unhappy vignette.

In our survey we asked all women in the sample the following question: "How much freedom/opportunity do you have?" using a scale running from none, a little, some, to a lot. Next we presented a sub-sample of women with three vignettes representing typical women with different levels of power (see the Appendix). The vignettes were designed and phrased in Telegu by a local consultant to represent typical cases occurring in the implementation of the project. Each woman is taking a loan from the Self-Help Group (SHG). One woman, obviously empowered, uses the loan to start a successful business. Another woman, obviously powerless, is denied access to the loan by her own husband. A third woman lies in between the other two: she is able to take the loan but is forced to give the money to her husband. Respondents were asked to rate the empowerment level of each vignette using the same scale used in the self-assessment. Finally, women were asked to compare their level of empowerment with that of the women in the vignettes.

4

Two key assumptions are made when using vignettes: response consistency and vignette equivalence (King et al. 2004). Response consistency requires that respondents use the same categories in the self-assessment and in the vignettes ratings. It would be problematic if a woman defined herself as powerful and at the same time identified with a vignette that she defined as powerless. Vignette equivalence requires that the respondents interpret the vignettes in the same way. A problem would arise if some women interpreted the vignettes as referring to agency (as intended in our case) while other women interpreted the vignettes as referring to freedom of movement. In particular, we would like women to give responses in agreement with the vignettes ranking constructed by design. Our vignettes have an implicit 'empowerment' ordering, whereby Chandamma is more empowered than Manemma who is more empowered than Neelamma. By rating the empowerment level of each vignette, respondents are implicitly ranking the three vignettes. The rankings constructed by design and the implicit rankings produced by the respondents can be compared to see to what extent they are similar. Table 1 shows that most respondents implicitly ranked the three women as in the vignette design. However, the table also show that many women ranked Chandamma as the most 'empowered', particularly in the first survey, showing that the design and the administration of the questionnaires could be much improved.

Results from vignettes interviews can be analysed using non-parametric (King and Wand 2007) and parametric methods, but it is the Hopit parametric model (the hierarchical ordered probit model) that has gained more popularity in the literature. The great advantage of the parametric model is that it allows administering the vignettes to a sub-sample of respondents, while non-parametric methods require administering the vignettes to all respondents. The Hopit model has been used to quantify hard-to-measure outcomes such as 'political efficacy' (King et al. 2004), 'work disability'

5

(Kapteyn, Smith, and van Soest 2007), 'job satisfaction' (Kristensen and Johansson 2008), 'health status' (Bago D'Uva et al. 2008). The model has been described several times in the literature and will only briefly summarised here.[i] In words, the Hopit models self-assessments on an ordinal scale after removing systematic variations across groups due to subjective perceptions. Its goal is estimating an ordered probit of self-assessments, after removing from self-assessments all subjective biases that vary across groups.

The Hopit model can be better described as composed of two parts. In the first part, the respondent's reporting behaviour is modelled using covariates that explain systematic variation in reporting patterns. For example, more educated women may rate the same vignette differently from less educated women. This part of the model is used to identify the rating scales used by the respondents. In the second part of the model, self-assessments are modelled using covariates (such as socio-economic status or project status) as in an ordered probit model, but the cut-off points are obtained from the first part of the model and are respondent-specific.

Formally, in the first part of the model, the latent rating (latent is indicated here by a star) of a vignette *j*, as perceived by the respondent *i*, simply depends on random measurement error (*e*) (the *v* superscript indicates this is a vignette rating):

$$Y_{ij}^{v*} = a_j + e_{ij}^{v} \qquad (1)$$

The observed vignette rating is related to the latent rating by:

$$y_{ij}^v = k \; if \; \tau_i^{k-1} \leq Y_{ij}^{v*} < \tau_i^k \qquad (2)$$

where the thresholds (τ) are modelled as function of X covariates:

$$\tau_i^k = X_i \gamma^k \qquad (3)$$

In the second part of the model, the latent self-assessment is a function of Z covariates (with s superscript indicating this is a self-assessment rating):

$$Y_i^{s*} = Z_i \beta + e_i^s \qquad (4)$$

and the observed self-assessment is related to the latent self-assessment by:

$$y_i^s = k \; if \; \tau_i^{k-1} \leq Y_i^{s*} < \tau_i^k \qquad (5)$$

The first part of the model identifies the thresholds that are imposed in the estimation of the second part of the model. In the second part of the model self-assessments are explained by sample characteristics after correcting for heterogeneity in responses. The model can then be used to explain, for example, how empowerment varies across women of different socio-economic status or across women exposed and not exposed to a particular project.

The first part of the model, explaining reporting behaviour, is also of interest because it detects the sources of variations in reporting behaviours. For example, to what extent, and for which groups, reported outcomes differ from actual outcomes? Differences in

reporting patterns across groups are better investigated through statistical tests. In particular, we can test whether reporting behaviour is homogeneous across women characteristics (whether for example education has any impact on reporting behaviour) by testing the joint significance of the coefficients of that factor across the thresholds. In what follows we will refer to this test as the 'homogeneity test'.

The data and the Rural Livelihood Project

The data used in this paper were collected in Andhra Pradesh in two waves from June to July of 2005 and from June to July 2007. They form a panel of households designed to evaluate the impact of the Andhra Pradesh Rural Livelihood Project (APRLP) implemented by the government of Andhra Pradesh with the support of DFID. A baseline survey was conducted in 2005 before project implementation in 15 project villages and 15 control villages matched to project villages using census data. A follow-up survey was carried out after two years to assess the impact of the intervention. The project had the final objective of reducing poverty by building social capital and empowering women. Self-help groups (SHGs), community-based organisations of poor women, were a key element of the project. They were given the task of allocating project funds and of channelling loans to their members.

The surveys interviewed 1,482 households from thirty villages in the districts of Anantapur, Kurnool, Mahaboobnagar, Nalgonda, and Prakasam. The surveys collected data on demographic characteristics, agricultural incomes, employment, water use, access to credit and participation to SHGs. In order to assess the impact of the intervention on empowerment an experimental vignette module was included in both survey rounds. In 2007, self-assessment questions were asked in all interviews and the vignette module was administered to a sub-sample of 236 households from 5 villages

randomly selected from each district in the study. In 2005, vignettes were administered to 243 women but for an oversight in the printing of the questionnaires, the self-assessment question was not asked to all 1,482 respondents. For this reason we are not able to look at changes in self-reported empowerment across the two survey rounds. However, in both surveys women selected for the vignette module were also asked to position themselves in relation to the three vignettes (see question 3 in the Appendix). We will use this self-assessment in the following sections to test the impact of the project on the sub-sample of women taking the vignette module.

Heterogeneity in self-reported empowerment

We start by looking at how self-reported empowerment varies with women's characteristics. Table 2 shows the result of an ordered probit regression of self-assessments (responses to the question "How much freedom/opportunity do you have?" on a scale from 1 to 4) on characteristics normally associated with women's empowerment. Table 2 also shows the result of the same regression using a more tranditional 'agency index' as the dependent variable. This agency index was built as the sum of the positive responses to the following questions: "Who accompanies you when you travel outside the village?" (no one); "Who decides on how to spend the money you earn?" (only me); "Who decides on whether or not you should apply for a loan?" (only me). Self-assessments are positively correlated with SHG membership, husband's education, female-head household status and wage employment. The correlations are not strong and there is little systematic variation across groups. Interestingly, the agency index produces very similar results: the signs of the coefficients in the two models are similar, almost the same variables are found to be associated with the indices and little overall systematic variation across groups is found. The similarity of the results suggests that the two indices are describing a similar conceptual construct.

Women who are member of SHGs, whose husbands are more educated, who are head of household, and who are wage labourers report higher levels of empowerment. These associations can reflect true empowerment differences across groups or a different understanding of empowerment across groups. In order to remove perception biases from self-reported assessments we run a Hopit model that simultaneously estimates reporting patterns and determinants of self-reported empowerment. We do this separately for the full sample of 1,431 women and for the restricted sample of 236 women that were administered the vignettes. The results are shown in Table 3 together with the result of a standard ordered probit. The Hopit model finds smaller coefficient estimates than the ordered probit and the coefficients are not statistically significant. This suggests that the empowerment differences found across groups by the ordered probit, with the exception of SHG membership, are more perceptual than real. We further investigate this hypothesis by testing homogeneity in vignettes responses using the coefficients of the Hopit model (Table 4). We conduct this test for all coefficients together and for each separately. The tests show that homogeneity in responses is rejected in the full sample for all variables together and that, in particular, illiterate women, women with educated husbands and head of households tend to report higher levels of empowerment than those actually experienced. The analysis conducted on the sub-sample of women that were administered the vignettes does not provide any new insight. The small size of the sample inflates the standard errors and few variables are statistically significant. The differences in coefficients between the unadjusted ordered probit and the Hopit model are similar to those observed using the full sample.

The analysis conducted so far shows that there are some systematic differences in the way women report their empowerment status. In particular, women whose husband is

highly educated, women head of households, agricultural labourers and members of self-help group report higher levels of empowerment. However, associations between these characteristics and empowerment vanish once they are corrected by perception biases, with the exception of a weak association between SHG membership and empowerment. This suggests that women in the sample are similarly disempowered regardless of their background characteristics, and that membership of self-help groups is the only association that holds after correcting for perception biases.

Impact of SHGs on women's empowerment

The weak association between SHG membership and empowerment observed in the previous section, though unaffected by perception bias, does not have a causal interpretation. It is equally possible that SHGs empower women as well as that more empowered women join SHGs. The evaluation of APRLP was framed as a difference-in-difference design of matched project and control villages in order to assess programme impact. However, after the baseline survey the government of Andhra Pradesh started to form SHGs in all communities in the state following the implementation of a World Bank-supported project (the *Indira Kranti Pratham*), which was very similar to the APRLP. As a result, SHGs are present in all communities surveyed and in some villages the establishment of SHGs by NGOs and by other projects pre-dates APRLP. In the absence of a valid control group we follow two approaches to establish causality between SHG membership and empowerment: 1) an instrumental variable approach and 2) a difference-in-difference analysis applied to the small panel of women that were administered the vignettes.

In the first approach we instrument SHG membership using 2-stage least square by running a regression of SHG membership on the same covariates used in the models of

Table 3 and variables representing shocks. Crop shocks and demographic shocks (illness and deaths) are the selected instruments, as they affect SHG participation but are uncorrelated with empowerment. The predicted SHG participation is then included in the Hopit model. The results of this approach for the full sample and for the sub-sample are reported in Table 5. The coefficients have the expected positive sign but standard errors are very large and estimates are not statistically significant.

In the second approach we use women's comparison to the three vignettes. Women participating in the vignettes exercise were asked to position themselves in relation to the ordered scale provided by the vignettes. This self-assessment is not affected by perception bias and is available for a panel of 236 women in the 2005 and 2007 surveys, so that a difference-in-differences approach is feasible. We estimate the difference-in-differences model in first differences. Changes in self-assessments are regressed over changes in SHG status as some women are joining SHGs between the two surveys while other are leaving SHGs over the same period. Other control variables included in the regression are: changes in household size, changes in wage employment status, changes in head of household status and baseline values of land ownership, caste membership, age and literacy. The observed impact of SHG on empowerment is very small considering that the dependent variable is valued from 1 to 7 and is not statistically significant (bottom row of Table 5). These results suggest that the association between SHG membership and empowerment found in the previous section does not have a causal interpretation: SHGs do not empower this sample of women.

Conclusions

Several projects are aimed at empowering women and many empowerment indices have been proposed to assess the impact of these interventions. These indices rely on

long questionnaires and subjective assessments. As a result, they are costly and not fully reliable. In this paper we propose to measure empowerment using anchoring vignette. The method is easy to implement and addresses the problem of incomparability of subjective perceptions.

We present an application of this methodology using data from the evaluation of a rural livelihood project in Andhra Pradesh. We find that women of different socio-economic backgrounds perceive empowerment in different ways. In particular, women members of SHGs, and women of higher socio-economic status tend to overstate their empowerment status. However, after adjusting for perception biases, none of the variables traditionally associated with empowerment (such as education, wage earning status and household headship) appears to be correlated with empowerment, with the exception of a weak association with SHG membership.

We further investigate the impact of SHG membership on women's empowerment using an instrumental variable approach and difference-in-difference analysis. In both cases, the results show that membership of SHGs has a small impact on empowerment and it is never statistically significant. These results suggest that the observed association between SHG membership and empowerment is partly perceptual, as SHG members tend to overstate their empowerment status, and partly running in the opposite direction: more empowered women are more likely to join SHGs.

Table 1 Comparison of vignettes rankings and respondents' rankings

	Survey 2005			Survey 2007		
	Ranked 1st	Ranked 2nd	Ranked 3rd	Ranked 1st	Ranked 2nd	Ranked 3rd
Nelamma	**67.4**	12.5	20.2	**81.8**	9.8	8.5
Manemma	13.2	**77.3**	9.5	18.2	**77.5**	4.2
Chandamma	29.7	14.7	**55.7**	15.7	18.6	**65.7**

Table 2 Characteristics associated with self-reported assessment and with the *agency* index

	Self-reported assessment	Agency index
SHG member	0.097*	0.92
	(0.099)	(0.145)
Scheduled tribe	0.070	-0.114
	(0.499)	(0.310)
Other backward caste	0.091	0.050
	(0.225)	(0.527)
Other caste	0.049	-0.122
	(0.622)	(0.259)
Age	0.024	0.007
	(0.123)	(0.674)
Age squared	-0.001	-0.001
	(0.296)	(0.856)
Illiterate	0.063	-0.021
	(0.457)	(0.814)
Husband's education	0.100***	0.022
	(0.000)	(0.334)
Female-headed household	0.285**	1.258***
	(0.003)	(0.008)
Household size	-0.023	-0.004
	(0.134)	(0.814)
Land size	0.006	-0.012
	(0.439)	(0.127)
Agricultural labourer	0.189**	0.166**
	(0.004)	(0.019)
Observations	1,428	1,428

Note: P-value of t statistical tests in parentheses. *** is 1% significance, ** is 5% significance and * is 10%

Table 3 Ordered probit and Hopit model of women's empowerment

	Full sample		Restricted sample	
	Ordered probit	Hopit	Ordered probit	Hopit
SHG member	0.081*	0.037	0.298**	0.284*
	(0.099)	(0.712)	(0.014)	(0.061)
Scheduled tribe	0.058	-0.173	0.277	0.111
	(0.499)	(0.502)	(0.397)	(0.786)
Other backward caste	0.076	-0.054	0.390*	0.259
	(0.225)	(0.675)	(0.013)	(0.181)
Other caste	0.041	0.091	0.279	0.334
	(0.622)	(0.607)	(0.193)	(0.213)
Age	0.020	0.022	-0.001	0.001
	(0.123)	(0.354)	(0.987)	(0.974)
Age squared	-0.001	-0.001	0.001	0.001
	(0.296)	(0.321)	(0.567)	(0.850)
Illiterate	0.052	0.202	0.067	-0.162
	(0.457)	(0.165)	(0.698)	(0.455)
Husband's education	0.083***	0.012	0.071*	0.004
	(0.000)	(0.749)	(0.099)	(0.948)
Female-headed household	0.237**	0.140	0.064	0.013
	(0.003)	(0.307)	(0.675)	(0.946)
Household size	-0.019	-0.052*	0.031	-0.001
	(0.134)	(0.060)	(0.344)	(0.986)
Land size	0.005	0.001	-0.001	-0.003
	(0.439)	(0.993)	(0.990)	(0.872)
Agricultural labourer	0.157**	0.099	0.125	0.067
	(0.004)	(0.387)	(0.369)	(0.688)
Observations	1,431	1,431	236	236

Note: P-value of t statistical tests in parentheses. *** is 1% significance, ** is 5% significance and * is 10% significance. In order to compare the results, the coefficients of the ordered probit were multiplied by the variance of the error term estimated by the Hopit model

Table 4 Test of homegeneity in self-reported empowerment across groups

Homogeneity test	Full sample	Restricted sample
All covariates	0.002**	0.552
SHG member	0.096*	0.584
Caste	0.568	0.702
Age	0.514	0.695
Illiterate	0.040**	0.287
Husband's education	0.041**	0.141
Female headed household	0.000***	0.060*
Household size	0.452	0.524
Land size	0.253	0.912
Agricultural labourer	0.835	0.716
Parallel shift	0.001***	0.614

Note: the cells report P-values of statistical tests of joint significance of the coefficients across thresholds in the Hopit model for each covariate separately and for all the covariates together.
 *** is 1% significance, ** is 5% statistical significance and * is 10% significance.

Table 5 Impact of SHGs on women empowerment (IV and difference-in-differences models)

	Coefficient	observations
Hopit model with SHG membership instrumented[a] (full sample)	0.238 (0.844)	1,431
Hopit model with SHG membership instrumented[a] (restricted sample)	0.199 (0.794)	236
Difference-in-difference model[b] (restricted sample)	0.202 (0.313)	232

Note: standard errors in parentheses

[a] Instruments consists of: a dummy variable for households affected by a crop shock and a dummy variable for households affected by the death of a household member in addition to caste, age, woman's literacy, husband's education, head of household status, household size, land size and agricultural labour status. The dependent variable is self-reported assessment on a scale from 1 to 4 as in the models of Table 3.

[b] The dependent variable is a self-assessment based on a comparison to the three vignettes on a scale from 1 to 7 (see question 3 of the vignette module in the Appendix). Explanatory variables are: change in SHG membership status, change in agricultural employment status, change in household size, and baseline values of caste, literacy, husband's education, and land size.

Appendix

Women's empowerment vignettes

1. How much freedom/opportunity do you have? (none, al little, some, a lot)

2. I am going to describe the situation of three different women. For each of them I would like you to answer the question "How much freedom/opportunity does she have?" (none, a little, some, a lot)

    I. Neelamma takes a loan from the Self-Help Group and starts a grocery shop. Despite his initial opposition, her husband is won round to the idea when he sees the good returns from the business.

    II. Manemma takes a loan from the Self-Help Group to start a small home business. But her husband argues with her that the money is needed for land improvement, and she reluctantly parts with the loan money.

    III. Chandamma wants to take a loan from the Self-Help Group of which she is a member to start a tea stall. She could not take the loan because her husband and mother-in-law did not agree.

3. How would you compare your freedom/opportunity with that of the women described above? (less than Chandamma, same as Chandamma, more than Chandamma but less than Manemma, same as Manemma, more than Manemma but less than Neelamma, same as Neelamma, more than Neelamma)

References

Alsop, R., M. Bertelsen, and J. Holland. 2006. *Empowerment in Practice: From Analysis to Implementation*. Washington DC: The World Bank.

Bago D'Uva, T., E. Van Doorslaer, M. Lindeboom, and O. O'Donnel. 2008. Does Reporting Heterogeneity Bias the Measurement of Health Disparities? *Health Economics* 17:351-375.

Duflo, E. 2012. Women Empowerment and Economic Development. *Journal of Economic Literature* 50 (4):1051-1079.

Eyben, R., N. Kabeer, and A. Cornwall. 2008. Conceptualising empowerment and the implications for pro-poor growth. A paper fro the DAC Poverty Network. In *Report to DAC POVNET on empowerment*. Falmer: IDS.

Jones, A. M., N. Rice, T. Bago D'Uva, and S. Balia. 2013. *Applied Health Economics*. New York: Routledge.

Kabeer, N. 2001. Reflections on the Measurement of Women's Empowerment - Theory and Practice. In *Discussing Women's Empowerment - Theory and Practice*. Stockholm: Novum Grafiska AB.

Kapteyn, A., J.P. Smith, and A. van Soest. 2007. Vignettes and Self-Reports of Work Disability in the United States and the Netherlands. *The American Economic Review* 97 (1):461-473.

King, G., C.J.L. Murray, J.A. Salomon, and A. Tandon. 2004. Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review* 98 (1):191-207.

King, G., and J. Wand. 2007. Comparing Incomparable Survey Responses: New Tools for Anchoring Vignettes. *Political Analysis* 15:46-66.

Kishor, S., and K. Gupta. 2009. *Gender Equality and Women's Empowerment in India*. Deonar, Mumbai: International Institute for Population Sciences.

Kristensen, N., and E. Johansson. 2008. New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Economics* 15:96-117.

Malhotra, A., S. R. Schuler, and C. Boender. 2002. Measuring Women's Empowerment as a Variable in International Development. World Bank.

Rabe-Hesketh, S., and A. Skrondal. 2014. *Estimating chopit models in glamm: Political efficacy example from King et al. (2002)* 2002 [cited 9 October 2014]. Available from http://www.gllamm.org/chopit.pdf.

Sen, A. 1990. More than 100 Million Women are Missing. *The New York Review of Books*.

———. 2002. Health: Perception versus Observation. *British Medical Journal* 324:860-861.

Solava, I., and S. Alkire. 2007. Agency and Empowerment: A Proposal for Internationally Comparabe Indicators. *Oxford Development Studies* 35 (4):379-403.

Tandon, A., C.J.L. Murray, J.A. Salomon, and G. King. 2003. Statistical models for enhancing cross-population comparability. Health systems performance assessment: debates, methods and empiricism. In *Health Systems Performance Assessment: Debates, Methods and Empiricism*, edited by C. J. L. Murray and D. B. Evans. Geneva: World Health Organisation.

Vaessen, J., A. Rivas, M. Duvendack, R. Palmer Jones, F.L. Leeuw, G. Van Gils, R. Lukach, N. Holvoet, J. Bastiaensen, J.G. Hombrados, and H. Waddington. 2014. The Effects of Microcredit on Wome's Control over Husehold Spending in Developing Countries: A Systematic Review and Meta-Analysis. *Campbell Systematic Reviews* 2014:8.

[i] For a description of the original formulation of the Hopit model (also called chopit) see Tandon et al. (2003) and King et al. (2004). The code for running the model in stata can be found in Jones et al. (2013). Alternatively, researchers can use the *anchor* software in R or the user-written stata program *gllamm* (Rabe-Hesketh and Skrondal 2002). See Kristensen et al. (2008) for an application of the latter software to vignettes. Codes and data used in this paper are available upon request with the corresponding author.