Better to be indirect? Testing the accuracy and cost-savings of indirect surveys

Jennifer M. Alix-Garcia, Katharine R.E. Sims, and Laura Costica

Acknowledgements: We are grateful for financial support from The World Bank's i2i program, Mexico's National Evaluation Agency (CONEVAL), the Mexican National Forestry Commission (CONAFOR) the Carnegie Corporation of NY, and our academic institutions. We are also grateful for feedback given in seminars at the International Food Policy Research Institute and the University of California at Davis. Excellent research assistance was provided by Thamanna Vasan. The data used in this paper was collected concurrently with data for a study led by PIs Alix-Garcia, Sims, and the World Bank Development Impact Evaluation (DIME) unit. The Mexican National Forestry Commission and its field offices aided substantially in the implementation of the survey, which was carried out by Chapingo University. The IRB protocol number is: #2016/04/17. This paper was considerably improved by feedback from participants in a seminar at the International Food Policy Research Institute.

Author information: Alix-Garcia: Department of Applied Economics, Oregon State University, jennifer.alix-garcia@oregonstate.edu. Sims: Department of Economics, Amherst College, ksims@amherst.edu. Costica: World Bank, Washington, D.C., lcostica@worldbank.org.

Better to be indirect? Testing the accuracy and cost-savings of indirect surveys

Abstract

We test the validity of indirect surveying as a method to collect household data. We compare household and informant reports of assets, develop poverty indices from both, test their performance as regression covariates, and examine errors in reporting and targeting resulting from using indirectly reported variables. Informant-based targeting indices are highly correlated with household measures and can be reasonable substitutes for self-reported indices in simple regressions. They can also be used to assign a simulated anti-poverty program with similar error rates to related methods. In our setting, eliminating direct household surveys would have reduced survey costs by 50%.

JEL codes: O12; I32; I38; C83
Keywords: indirect survey; survey design; targeting

Highlights: Better to be indirect? Testing the accuracy and cost-savings of indirect surveys

- We compare reports of assets obtained indirectly from knowledgeable local informants to those collected directly from households.
- We test whether these reports are accurate, both for specific characteristics and when aggregated into asset indices.
- There is high correlation between household and informant reports.
- Indirectly reported asset indices perform well in simple regressions and in a very basic targeting exercise.
- In our context, indirect surveying reduces the survey budget by 50%.

Better to be indirect? Testing the accuracy and cost-savings of indirect surveys

1 Introduction

Household surveys provide a foundation for research, policy evaluation, and program implementation. During 2019, nearly 60 percent of over 90 articles published in the *Journal of Development Economics* used data collected directly from households, and over 40 percent of these were collected specifically for the published paper.[1] Household surveys are frequently used to determine eligibility for anti-poverty programs, identifying the communities and households best qualified to receive program benefits. Yet despite the obvious advantages of face to face enumeration, such surveys are very costly for organizations and impose substantial time-costs on households themselves. These high costs reduce funds available for other purposes, including the programs themselves. For example, during the first four years of Mexico's PROGRESA educational subsidy program, the cost of beneficiary identification was 34 percent of total program costs (Caldés et al. 2006). Kenya's Household Safety Net Program, another cash transfer mechanism, cites targeting costs of 7 percent of the operating budget (Bahri & Merttens 2018). Although an average survey cost per household is difficult to calculate, available information suggests a wide range – from $18 per household surveyed in an agricultural household survey in Mozambique (Kondylis et al. 2015) to an implied cost of around $300 per household surveyed in the average LSMS (Sustainable Development Solutions Network 2015).

In this paper we test the validity of "indirect surveying", which provides a possible low-cost way to collect household-level information. In indirect surveying, a knowledgeable informant reports on characteristics of other households, reducing transportation and time costs of surveying. To assess the usefulness of indirect surveying, we report the results of a comparison between household and leader-reported characteristics from a survey in rural Mexico. We assess whether

[1] Authors' manual count.

community leaders report household characteristics accurately and whether errors in reporting or targeting vary systematically across community size, leader education, measures of elite capture, or social capital. We then consider the viability of using informant information in regression frameworks. In addition, we examine if an assets index constructed from leader reports could be used for targeting, and the potential cost savings of indirect surveying.

We find that the leader and household reports of the same characteristic are highly correlated, including when conditioned on community fixed effects. Using the set of variables that are reported by both community leaders and households, we construct binary, proportional weighting, and principal components asset indices. In all cases, the indices constructed using the leader reports are highly correlated with the household indices, with correlations near 0.70 across communities and 0.54 within communities. These indices are not intended to measure poverty perfectly, but rather to compare proxies for poverty constructed from directly versus indirectly reported information. Asset indices are widely used as a proxy for poverty and have been tested in multiple global contexts, including Mexico (McKenzie 2005, Filmer & Scott 2008, Chakraborty et al. 2016).

In addition, we show that leader and household indices produce similar results in simple regressions common to research applications. Then, to assess how indirect surveying would compare with direct surveying as a means to target anti-poverty programs, we consider a hypothetical program where households are identified as poor (eligible for the program) if they fall below the median index value in our sample. We find that if we were to use the leader-constructed indices to identify the poor and non-poor, the correspondence between household and leader-identified poverty as measured by asset indices would be correct between 69 and 77 percent of the time, with highest accuracy for the simplest index. The error from classifying households as poor who are not (error of inclusion), ranges from 10 to 21 percent, while the misidentification of households as non-poor when in fact they are poor (error of exclusion) is from 26 to 46 percent. We also examine whether or not network size, capture of leadership positions, social capital, or education levels affect reporting accuracy. Most informant characteristics are unrelated to accuracy, although we find small associations between social capital and inclusion error and between leader education

and age and exclusion error.

Finally, we construct estimates of the potential cost-savings of indirect surveying using information on staff and travel costs. We find that in the context in which one leader and ten households were interviewed per village, interviewing the leader alone could save nearly 50% of the budget. Back of the envelope calculations using actual conditional cash transfer amounts to proxy for benefits suggest the ratio of benefits provided to survey costs is considerably higher for indirect surveying. Cost-savings of indirect surveying would likely be highest when populations are sparsely distributed and the transport costs of reaching households within communities are high.

This paper contributes to the literature on survey methodology related to alternate methods of collecting household information, such as key informant surveys (Kumar 1989) and rapid rural assessments. Prior literature suggests that covariates like gender, race, age (Frenk et al. 2010), productive capital (Takasaki et al. 2000), landholding and years of schooling (Finan & Schechter 2012), and community characteristics (Young & Young 1961) tend to be more accurately reported by informants outside the household than characteristics like small consumer durables, food consumption, or changes in religious values (although Finan & Schechter (2012) find that knowledgeable informants can also predict others' social preferences). Perhaps unsurprisingly, some indirect informants have been found to report inaccurately, particularly when they do not have much outside contact (Macours 2003), if they believe that their answers could influence the distribution of programs (Rigol et al. 2017), if they do not trust interviewers (Chambers 1994), if enumerators exert undue influence on respondents (Bergeron et al. 1998), or if informants are simply not very well-informed (Phillips 1981, Garrett & Downen 2001, Romney et al. 1986). Our work is similar to this prior literature in that it tests the degree of misreporting of individual assets and whether or not this misreporting is systematic. However, we use a considerably larger and more nationally representative sample than prior tests and we directly test the performance of indirectly-derived measures in regression analysis and targeting. In addition, although previous literature has considered characteristics that might be related to social capital, this is also the first study to examine whether reporting is related to a multi-dimensional index designed specifically to measure com-

munity social capital.

We also contribute to the literature on cost-effective ways to locate the poor. Important progress has been made through methods including poverty mapping (Davis 2003, Elbers et al. 2004), poverty scorecards (Schreiner 2017, Skoufias et al. 2020), and geographic targeting (Bigman & Fofack 2000). Others have examined how to lower the costs of program implementation through community-based methods of beneficiary selection (Alderman 2002, Galasso & Ravallion 2005, Alatas et al. 2012, Karlan & Thuysbaert 2016, Stoeffler et al. 2016, Basurto et al. 2017, Keane et al. 2020) or self-targeting (Alatas et al. 2016, Olken 2016). Some prior research (e.g. Keane et al. (2020)) emphasizes locally appropriate definitions of poverty and uses focus groups with key informants to first define poverty indicators before classifying households as poor.[2] Community-based targeting can be advantageous because it utilizes specific local knowledge about poverty dimensions which may be better observed by those within the community (Alatas et al. 2012, 2016).

Indirect surveying shares with these methods the potential to increase the cost-effectiveness of data collection for targeting purposes. However, it differs from many of these methods in that rather than asking community leaders or members to assess or define poverty directly, informants supply specific characteristics which can then be used for analysis or to facilitate targeting. An assets index based on these responses may be less susceptible to bias than community-based methods because it relies on the collection of observable information, rather than the direct identification of poor households. It may also avoid social desirability bias, where households do not answer or give incorrect information because they are ashamed to admit they do not own a particular item (Stecklov et al. 2006). Like community-based targeting, indirect surveying has the potential to lower survey costs by leveraging the information available to community leaders or members rather than needing to reach each household. Here we show that in the case where policymakers agree that targeting can be done using an asset index approach (Filmer & Pritchett 2001, Rutstein & Johnson

[2]We acknowledge that there is a potential tradeoff between the comparability across villages that is gained by asking about a pre-determined set of asset indicators, versus allowing indicators to be tailored to the specific livelihood strategies of different communities. Appropriate field-testing of surveys in multiple regions would be recommended to ensure a set of indicators that do have relevance across communities.

2004, McKenzie 2005, Filmer & Scott 2008), the benefit-cost ratio of the indirect survey method is higher than that of direct surveying provided the budget savings of the former is greater than 27% . We recognize that asset-based measures of poverty can be flawed and may be controversial (Brown et al. 2018). Our purpose here is not to interrogate this particular targeting method, as we do not have full consumption data with which to compare our indices, but to provide an example of the trade-off between misreporting due to indirect surveying and the potential cost savings of this method.

2 Data and empirical approach

2.1 Data description

The test of indirect surveying was included as a short module to a survey collected for the purposes of evaluating the social and environmental impacts of a payments for ecosystem services (PES) program in Mexico (Alix-Garcia, Sims, Orozco-Olvera & et al. 2018). The design of the evaluation took advantage of a threshold in the enrollment criterion for the program, where beneficiaries were selected based upon a point system assigned by the implementing agency (National Forestry Commission of Mexico – CONAFOR). As a result, the sampling strategy selected treatment and control communities with assignment scores as close to the acceptance threshold as possible while still preserving regional representativeness. In total, 862 communities and 8,413 households were surveyed in 12 Mexican states in 2016. All participants had applied to the PES program between 2011 and 2014.

Figure A1 highlights the states in which interviews took place. Table A1 includes the means and normalized differences of means between the sampled applicants and the universe for the survey (column 2) and universe of all common properties (column 5). The differences between the sampled and not sampled communities within states are generally quite small in magnitude – the main differences are that sampled villages tend to be in municipalities with greater proportions of indigenous people, have slightly higher forest canopy cover, and be somewhat farther from

major cities. The fourth column of the table lists a small set of covariates from all of the agrarian communities or "ejidos" and "comunidades" in the entire country.[3] Compared to this group, the communities in our sample tend to be larger in area, but similar in population size (within 0.25 standard deviations). This makes sense, given that a requirement for being in our survey was to have submitted an application to a land cover conservation program, and communities with larger area tend to also have more vegetation to conserve. In other respects, these communities are likely to be broadly representative of rural agrarian communities.

CONAFOR helped enumerators from Chapingo University in Mexico gain access to communities. Chapingo staff implemented the survey, but no CONAFOR staff were present for interviews. Enumerators identified themselves as coming from Chapingo and conducting academic research. Community leaders and individual households were interviewed separately. Leaders were individuals holding one of the main elected community positions of president, secretary, treasurer, or oversight committee. Surveyors asked first to speak with the president, but if the president was absent, they would speak with the designated representative. For the household survey, ten households were randomly selected to be interviewed, based on lists of members obtained in advance, with responses given by the household head.

The data used here are the questions related to household indicators, participation in community work, and meeting attendance. These questions appeared at the end of the leader survey, and the interviewed leader was only questioned about three households that were randomly selected out of the ten chosen for household surveys. Leaders were also asked about the functioning of the payments for ecosystem services program, community characteristics (infrastructure, productive activities, size, etc.), and social capital (Alix-Garcia, Sims, Orozco-Olvera, Costica, Medina & Monroy 2018).

Households responded to a set of questions about the PES program and their participation in community work, standard household demographic characteristics, income sources, housing

[3]These are collectively managed communities, many of which have origins before the Spanish conquest (Knight 1986). Current titles to the rights in these communities occurred as part of Mexico's Agrarian Reform between 1917 and 1992, and additional "regularization" of holdings occurring to the present. Today over half of the country's land area and half its rural population is held in agrarian communities (De Janvry et al. 1997, Fox 2007).

characteristics, and small assets. Housing characteristics included type of walls, roofing material, floor material, number of rooms, as well as access to public services, including existence of water and sewage connections, and amount of water and electricity availability. In order to construct a broad assets index, questions were asked about the presence or absence of the following items: television set, refrigerator, blender, microwave oven, laptop, car, truck, motorbike, bicycle, gas or electrical cooker, telephone landline, and mobile phone.

Community leaders were asked to report on a smaller set of questions about the three households. The indirect survey questions included: whether the household head had secondary education, type of walls of the household's home, ownership of a television, car, or mobile phone, the number of household members, how many assemblies the household attended in the past year, if someone from the household participated in forest maintenance in the past year, and how many heads of large livestock the household owned.[4] Although limited, these measures have the following useful characteristics. First, they include assets – educational attainment and housing characteristics – that are similar to those used to construct Mexico's marginality index (CONAPO 2019), which is used for targeting welfare programs throughout the country. Second, they include goods which are easier (a car) and more difficult (a mobile phone) to observe. These smaller assets are not included in the Mexican marginality index, but are frequently collected to construct the wealth index used internationally in the Demographic and Health Surveys (DHS 2020). Finally, these characteristics encompass a range of possible indicators from common to uncommon, which helps to capture the variation in wealth in the population.

2.2 Empirical approach and accuracy of reporting

The empirical strategy to assess how well leaders report household characteristics is straightforward. We examine simple correlations and correlations conditioned on fixed effects. The assumption here is that the preferred approach is to interview all households individually, so although

[4]We did not ask about land holding size because, although land sizes can vary within ejidos, all members have access to land.

we are well aware that survey respondents may misreport for a variety of reasons, our interest is in assessing the accuracy of the leader reports relative to those of the households.

Table 1 examines the correspondence between individual characteristics reported directly by households and indirectly by leaders. It shows the means and normalized differences between the two reports. The normalized differences should be interpreted as standard deviations of difference between the two listed means. The normalized difference is a unit-free measurement that gives a sense of the magnitude of a deviation, and is in many cases preferable to a traditional t-test for differences. We use the rule of thumb of a difference of 0.25 standard deviations as measuring a substantial difference (Imbens & Wooldridge 2009).

The table shows that the normalized differences are small – in all cases, less than one-quarter of a standard deviation. Consistent with prior studies, more observable items – housing walls, cars, and number of livestock – have the smallest differences. There are larger differences in the harder to observe characteristics: the number of assembly meetings, whether or not the household head has secondary education, and if the household has a cell phone. It is interesting that in eight out of the nine variables, leaders seem to overestimate – i.e., they are more likely to report positively for the binary variables, and also report slightly larger household and cattle herd sizes.

Table 2 reports correlations between leader and household reports of assets using OLS and regression with fixed effects at the community level. These range from 0.36 and 0.79 across the entire dataset, and from 0.19 to 0.68 with community fixed effects. For the responses that might be integrated into a wealth measure, the lowest overall correlation is for household head's schooling (0.36). We conclude from this table that reporting accuracy is generally high, but again tends to be higher for more observable characteristics.[5]

A potential concern with this method could be survey fatigue. Our application was done in a situation where leaders had already answered about sixty minutes of questions, so the indirect

[5]The estimations on number of assembly meetings is not estimated due to the fact that leaders generally report the same number of assembly meetings attended by everyone in the subsample of households surveyed. While attendance rates are generally high, it may be that reporting for "socially desirable" outcomes (e.g., attending assemblies) is likely to be more problematic. Recalling attendance precisely may also be particularly difficult. This recall difficulty does not affect our results as the assembly meetings are not a measure of poverty and not included in the assets index.

Table 1: Summary of leader and household reports of characteristics for the same household

	Leaders	Households	Norm Diff
Has cell phone	0.443	0.403	-0.057
	(0.497)	(0.491)	(0.000)
Head has secondary education	0.154	0.114	-0.083
	(0.361)	(0.318)	(0.000)
HH size	4.058	3.921	-0.057
	(2.022)	(2.018)	(0.000)
Has improved walls	0.509	0.503	-0.009
	(0.500)	(0.500)	(0.000)
Has television	0.727	0.782	0.090
	(0.446)	(0.413)	(0.000)
Has car	0.403	0.400	-0.004
	(0.491)	(0.490)	(0.000)
# of assembly meetings attended	2.344	2.565	0.226
	(0.732)	(0.637)	(0.000)
# of large livestock owned	1.062	0.989	-0.061
	(1.493)	(1.496)	(0.000)
Participated in forest care	0.796	0.737	-0.099
	(0.403)	(0.440)	(0.000)

Table shows the variables measured in both the direct and indirect surveys. Columns (1) and (2) contain mean values of variables reported by leaders and households, respectively. Standard deviations are in parentheses below the means. Column (3) lists the normalized difference of the two means.

Table 2: Correlations between leader and household reports of individual variables

	Cell phone (1)	Secondary school (2)	HH size (3)	Wall type (4)	TV (5)	Car (6)	# livestock (7)	Forest work (8)
A: Simple correlations								
Household report	0.468***	0.368***	0.676***	0.675***	0.501***	0.717***	0.790***	0.326***
	(0.020)	(0.031)	(0.022)	(0.016)	(0.023)	(0.016)	(0.017)	(0.024)
B: Ejido FE								
Household report	0.317***	0.337***	0.627***	0.493***	0.336***	0.615***	0.686***	0.157***
	(0.021)	(0.023)	(0.019)	(0.019)	(0.021)	(0.019)	(0.019)	(0.017)

Each panel represents a different estimation. Dependent variables are leader reports, and household report indicates the household response for the same variable. Standard errors are clustered at the ejido level.* p <.10, ** p< .05, *** p<.01.

surveying was already happening in conditions of potential fatigue. To further probe this issue, Appendix Table A2 table presents interactions between the household observation and the order of the interview in OLS and fixed effects regressions. The intention is to test whether the correlations between household and leader reports decrease as the leaders have to report on more households. There is no evidence that the correlation decreases with the interview number, although this certainly could be the case if leaders were required to report on greater numbers of households.

3 Asset indices using leader versus household reports

3.1 Index construction

To understand how indirect reporting could be used to save research or program costs, we create parsimonious asset indices from the leader and household reports. Such asset indices have been shown to be useful proxies for the level of well-being and for measuring wealth inequality in Mexico (McKenzie 2005). The DHS wealth index, which uses a principal components method to aggregate household characteristics, is shown to explain at least as much and sometimes more of the differences in household health indicators than consumption expenditures in a sample of 44 countries (Rutstein & Johnson 2004). Our index is not intended to be a perfect reflection of wealth or poverty in Mexico – the index in McKenzie (2005), for example, has thirty components, while our index has only five, similar to a simple poverty scorecard method (Schreiner 2017). Our intention is rather to test whether or not aggregations of variables reported by leaders resemble aggregations of variables reported by households. We aggregate variables using three approaches, then test the utility of the leader-reported indices as both objects in regression settings and as a mechanism for targeting a hypothetical anti-poverty program.

1. "Simple": We calculate a simple average of indicator variables for: a household head with secondary education, high quality walls on the home, ownership of a television, car, or mobile phone.

2. "Inverse proportion weighted" (IPW): We use a summation that weights each asset by the inverse proportion of its presence in the sample (Decancq & Lugo 2013).

3. "Principal components" (PCA): We use principal components to aggregate these same binary responses (Filmer & Pritchett 2001, McKenzie 2005). This is the method of aggregation used for Mexico's national marginality index and for the DHS wealth index. The former is calculated at the locality level and the latter at the household level.

Table 3: Summary of asset indices

	Leaders	Households	Norm Diff
Simple index (Simple)	0.492	0.444	-0.131
	(0.270)	(0.248)	(0.000)
Inverse proportion weighted index (IPW)	1.412	1.204	-0.227
	(0.689)	(0.604)	(0.000)
Principle components index (PCA)	0.000	0.032	0.018
	(1.336)	(1.249)	(0.000)
Observations	1764	1764	1764

Table shows the constructed indices. Columns (1) and (2) contain mean values of variables reported by leaders and households, respectively. Standard deviations are in parentheses below the means. Column (3) lists the normalized difference of the two means.

Table 3 reports the mean values of the indices calculated with each methodology using both the leaders' and the households' responses to the survey questions. The normalized difference of the mean reports is the largest for the IPW index, and smallest for the PCA index. Table 4 shows the simple and fixed effects correlations between leader and household-derived indices. There is a strong and similar correlation for all indices, with values of the correlation coefficient close to 0.70 and the coefficients from the fixed effects estimation greater than 0.50 in all cases. As with the individual variables, there is no evidence that indices calculated from the third household reported

by the leader exhibit lower correlation with the household reported index than in those calculated from the first household (Appendix Table A3).

Table 4: Correlations between leader and household reported indices

	PIB	IPW	PCA
	(1)	(2)	(3)
A: Simple correlations			
Household report	0.692***	0.693***	0.690***
	(0.018)	(0.019)	(0.018)
B: Ejido FE			
Household index	0.551***	0.529***	0.540***
	(0.024)	(0.024)	(0.024)

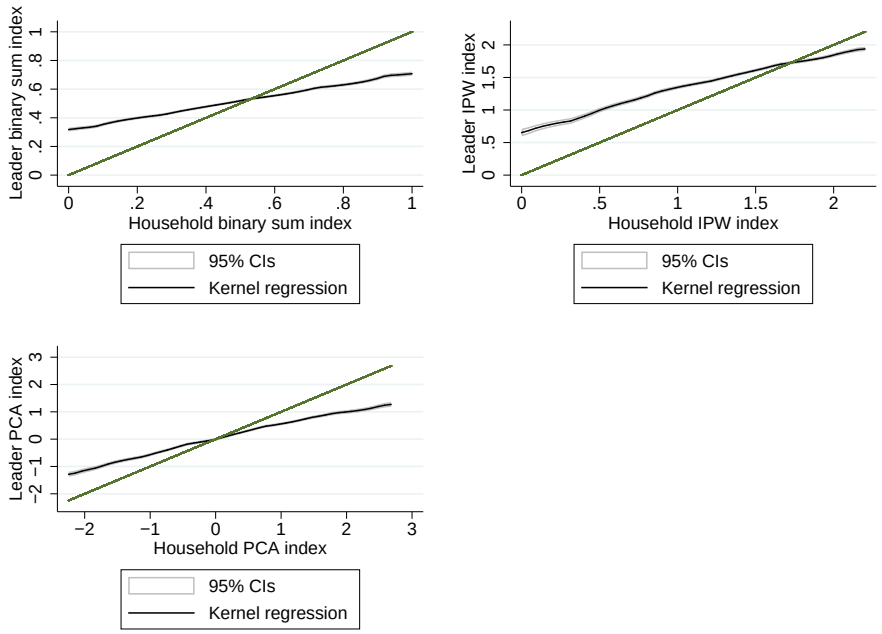
Each panel represents a different estimation. Dependent variables are leader indices, and household index indicates the index constructed from the household responses for the same variable. Standard errors are clustered at the ejido level.* p <.10, ** p< .05, *** p<.01.

Appendix Table A4 shows these same values for the indices aggregated up to village averages. The aggregation slightly increases the correlation between household and leader reports. In Appendix Appendix C we conduct an exercise where we vary the assumption that the household data is closer to the "truth". Under the assumption that the household data is a true measure, the reliability ratio of the simple leader-reported index is 0.44. If it is assumed that either could be equally correct, the ratio is 0.86.

Figure 1 visualizes the correlations between the leader reports (on the y-axis) and the household reports (on the x-axis) of each index. The 45-degree line shows hypothetical perfect correlation. In both cases, the correlations between the leader and household indices are strongly positive. However, they are clearly imperfect – leaders tend to report households as wealthier than households self-report, while the opposite is true for wealthier households (this pattern may also indicate some simple reversion to the mean). The problem of over-estimating household wealth appears worst in the IPW index. The variance of the raw indices differs across aggregation methods – the sim-

ple index has the lowest standard deviation and the principal components index the highest. For subsequent analysis, we normalize each index to facilitate comparability.

Figure 1: Asset indices calculated using leader versus household reports



Figures show kernel regressions of leader reported indices on household reported indices. The y-axis are leader reports and the x-axis household reports. The green line follows the 45 degree path from the origin.

An important limitation of this exercise is that we cannot assess these metrics relative to true consumption levels, since we did not collect detailed consumption data. We do have a rather simple index comprised of responses to a series of questions about whether or not a household consumed particular foods in the previous week. These foods include beans, tomato, chicken, milk, cheese, pork, and beef. The correlation coefficients between the simple food index and our constructed wealth indices for households are 0.41, 0.42, and 0.42 for the Simple, IPW and PCA indices respectively. They are also statistically different from zero. This indicates reasonable correspondence between assets and this available simple measure of consumption.

3.2 Using noisy measures in regressions

Because data is frequently collected to conduct empirical academic studies, we use this section to test whether using the leader rather than household indices makes a difference in a simple OLS or fixed effects regression. We are cognizant that they could vary substantially depending upon the outcome or context analyzed. In the interest of examining at least two cases, we estimate the correlation between school enrollment and migration and the poverty indices. Specifically, we generate variables measuring the proportion of children between 12 and 18 enrolled in school in each household, and if the household has at least one migrant in another country. We run simple regressions limiting our sample to households that have children in these age ranges in the first case, and on all the households in the second case. We include only households that have full data for both indices and the outcomes. We also test whether results are sensitive to the type of index.

Table 5 shows correlations calculated using both leader (odd columns) and household (even columns) indices, and with (columns (5)-(8)) and without (columns (1)-(4)) fixed effects. The point estimates for the relationship between household wealth and outcomes using direct versus indirect measures are similar in magnitude and significance, with slightly larger differences for the migration outcome. Chi-squared tests of difference across leader and household reported indices for each outcome all fail to reject equivalence of coefficients at the 5 percent level.

This table also confirms two standard results from the literature on measurement error. First, classical measurement error attenuates coefficients in a manner proportional to the signal to noise ratio. This explains why the coefficients in column (1) are slightly smaller than those in column (2). The differences are quite similar across indices, which is consistent with the very similar correlations between leader and household reported indices across aggregation methods. However, the difference in magnitude in the fixed effects estimation is larger than that of the simple regressions, confirming that a panel type structure can exacerbate this problem, particularly where there is high correlation within the panel. These results suggest that the indirect method of surveying would increase measurement error, but would still be useful for research purposes based on a realistic survey collection and application.

Table 5: Sample regressions using constructed poverty indices

	Secondary		Migration		Secondary		Migration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Simple index								
Simple index, leader responses	0.081***		0.014**		0.062*		0.015*	
	(0.016)		(0.006)		(0.037)		(0.009)	
Simple index, hh responses		0.095***		0.022***		0.092**		0.028***
		(0.016)		(0.006)		(0.037)		(0.009)
P-value diff.		0.207		0.081		0.314		0.150
B: IPW index								
IPW index, leader responses	0.073***		0.013**		0.054		0.011	
	(0.016)		(0.006)		(0.037)		(0.009)	
IPW index, hh responses		0.087***		0.021***		0.079*		0.028***
		(0.016)		(0.006)		(0.042)		(0.010)
P-value diff.		0.205		0.099		0.448		0.056
C: PCA index								
PCA index, leader responses	0.077***		0.014**		0.057		0.015*	
	(0.016)		(0.006)		(0.035)		(0.009)	
PCA index, hh responses		0.093***		0.021***		0.092**		0.028***
		(0.016)		(0.006)		(0.037)		(0.009)
P-value diff.		0.165		0.135		0.235		0.147
Ejido FE	no	no	no	no	yes	yes	yes	yes
Observations	614	614	1764	1764	614	614	1764	1764
Mean DV	0.703	0.703	0.071	0.071	0.703	0.703	0.071	0.071

Estimate shown are coefficients from OLS and FE OLS regressions of the proportions of children in a household attending secondary school or higher, or on having an international migrant on the binary wealth indices. School age is defined as 12-18 years old. Regressions only use households who have children in one or both of these age categories. Standard errors are robust and clustered at the ejidal level. Only households with complete information for both leader and household indices are used in the regressions. Indices are normalized before using in regression. The p-value in the last row of each panel is a chi-squared test for the difference in coefficients between the leader and the household reported measure. * p <.10, ** p< .05, *** p<.01.

4 Targeting using leader indices

4.1 Basic results

In this section we examine the additional error introduced by using the indirectly-derived leader index, rather than the directly-generated household index, for targeting of an anti-poverty intervention. Table 6 examines the level of error that would be introduced in a program that used the leader indices for targeting of an anti-poverty intervention. The targeting criterion is that a household is identified as "poor" if it falls below the median value of the poverty index calculated over the entire sample of household responses. This is meant to proxy for what might occur if an index were used for targeting – although the actual cutoff value might be established using a different set of data, such as the census. As shown in Table 6, in our sample using the simple index of household responses, 41.5% were non-poor and 58.5% were classified as poor.[6] Leaders correctly identified households as poor or non-poor 77% of the time (34.6%+42.6%). Conditional on not being poor, leaders correctly identified households as non-poor 83% of the time (34.6%/41.5%) and conditional on being poor, leaders correctly identified households as poor 73% of the time (42.4%/58.5%). Under standard error classifications, this means that for the simple index, inclusion error – where the non-poor were classified as poor – was 17% (6.9%/41.5%). Exclusion error – where the poor were classified as non-poor – was 27% (15.9%/58.5%). Exclusion error was largest for the IPW index (46.5%) and lowest for the PCA index (26%), with the magnitudes for inclusion error reversed. The overall targeting error rates range from 23% to 31%.

We also test to see if the targeting outcomes are improved by using those variables that are most observable and have the highest correlation in the simple regression tests – wall type, presence of car, and presence of television – to construct the indices. Appendix Table A5 shows that for the simple index, the classifications of leaders and households concur about 83% of the time (relative to 77% of the time in the 5-component simple index). The classifications corresponded 70 and 82% of the time for these restricted IPW and PCA indices, which also constituted a small improvement

[6]The numbers are not exactly 50-50 due to lumpiness in the indices.

Table 6: Classification of poverty using household versus leader reporting

	Household responses		
	Non-poor	Poor	Total
Leader responses			
Simple index			
Non-poor	34.6	15.9	50.5
Poor	6.9	42.6	49.5
Total	41.5	58.5	100.0
	Non-poor	Poor	Total
Inverse proportion weighted index			
Non-poor	37.5	27.4	64.9
Poor	3.6	31.5	35.1
Total	41.1	58.9	100.0
	Non-poor	Poor	Total
Principle components index			
Non-poor	36.3	13.9	50.2
Poor	10.4	39.4	49.8
Total	46.7	53.3	100.0

Cells show percentages in each category. In all cases, the poor are those with less than the median level of the index calculated across the entire sample using the index developed from household reports.

over 69 and 75% of the cases for the broader indexes.

Recent work reports error rates relative to consumption measures. Alatas et al. (2012) report that proxy means testing results in an overall targeting error of 30 percent in an experiment in Indonesia, with an inclusion error of 20% and an exclusion error of 53%, based on a poverty classification using detailed consumption information. The overall error increased by 10 percent (3 percentage points) using ranking by community leaders as a targeting method, with the increase coming from errors of inclusion. Other work on hybrid household and community targeting also suggests comparable error rates (Karlan & Thuysbaert 2016, Stoeffler et al. 2016). Our numbers are not directly comparable to these, since we are not focusing on a comparison to consumption measures. However, they do give measures of comparable types of targeting error generated by using leader rather than household reports of assets as a measure of poverty.

Additionally, Table 7 shows how leader and household indices would rank individuals from poorest to richest in a community. For the simple index, leader indices yield the same rank as households 55% of the time. Correspondence in ranking tends to be higher for poorer than for wealthier households. Confounding occurs most commonly between the two poorest households, rather than between the wealthiest and the second ranked. For example, the poorest and second-poorest household are confused in about 23 (11.1 + 12) percent of the cases, whereas this occurs between the second-poorest and the wealthiest household only about 15 (7.3+8.1) percent of the time using the simple index.

4.2 Are errors systematic?

This section examines whether or not village or leader characteristics influence the magnitude of misclassification. There are a variety of reasons why this might occur. First, leaders might be misreporting if they believe that such behavior might result in more resources being sent into the community or to particular households in the community. Reporting accurately might also be easier where there is more community coherence and stronger networks; conversely, it might be more difficult in settings where populations are larger – as larger networks are well-known to

Table 7: Classification by rank within community, leaders versus households

	Household responses			
	1	2	3	Total
Leader responses				
Simple index				
1	27.2	12.0	2.7	41.8
2	11.1	16.0	8.1	35.2
3	3.5	7.3	12.2	23.0
Total	41.8	35.2	23.0	100.0
	1	2	3	Total
Inverse proportion weighted index				
1	28.1	10.6	3.1	41.8
2	10.1	18.1	7.0	35.2
3	3.5	6.5	13.0	23.0
Total	41.8	35.2	23.0	100.0
	1	2	3	Total
Principle components index				
1	27.5	11.1	3.2	41.8
2	10.7	17.5	7.0	35.2
3	3.6	6.6	12.8	23.0
Total	41.8	35.2	23.0	100.0

Cells show percentages in each category. Rankings are unique, with the lowest value going to the poorest household. In the event that there are ties – which occurs for about 30 percent of the households – the ties are broken arbitrarily.

create complications in the dissemination of information (Genicot & Ray 2003). We also examine whether leaders who have been in power longer tend to report differentially. Individuals in community leadership positions are only allowed to hold a post for 3 years before there is a new election, and consecutive re-elections to the same position are not allowed. However, individuals may move into other positions within the community, return to a leadership position after a respite of 3 years, or simply ignore the rule. To try to capture this behavior, we take the proportion of years within the past ten that the secretary or president has held a community leadership position, and create an indicator for when this exceeds three years. Our intention is that this variable proxy for capture of the leadership, although it might also simply capture good leaders who are frequently elected.

Because community cohesion may affect the ability of a leader to report accurately on the characteristics of other households, we also test the relationships between error and an index measuring social capital. This index is based on participation in community assemblies, self-reported measures of trust, the presence of local public goods (a village vehicle, community house, etc.), the institutional structures that support social capital, such as the range of decisions made in assemblies, and the ability of non-community members to participate in them. The index comes from Alix-Garcia, Sims, Orozco-Olvera, Costica, Medina & Monroy (2018), which tests the impact of payments for ecosystem services on social capital using a regression discontinuity approach. Appendix D presents this measure in more detail. Finally, we examine if the education or age of leaders affects their ability to report accurately on the characteristics in which we are interested. These are examined individually, and then together in the same regression.

Table 8 shows regressions of binary indicators of exclusion or inclusion error for each of the indices. We do not find that community size or whether leaders were in power for more than three years is associated with error. While we do not observe significant differences in reporting bias as our sample approaches its maximum number of community members (over 3,000), we do find that the prevalence of "don't know" answers increases with community size (Appendix Table D3). Higher social capital is correlated with greater inclusion error for the simple and PCA indices, however, the magnitudes of this effect are small. In both cases, one standard deviation higher so-

cial capital results in an increase in inclusion error of 4 percentage points. Secondary education and greater age of leaders are significantly associated with greater exclusion error. A one standard deviation increase in the probability of having an educated leader leads to an increase in the probability of exclusion error of 7 percentage points, and an increase of parallel magnitude in leader age leads to a 5 percentage point increase in exclusion error. Table A6 shows these correlations in separate regressions for each covariate. The point estimates in these separate regressions are nearly equivalent to those in the estimation including all covariates simultaneously.

Table 8: Correlates of errors of inclusion and exclusion

	Exclusion error			Inclusion error		
	Simple	IPW	PCA	Simple	IPW	PCA
IHS(Community members w/ full rights)	-0.001	0.003	0.001	-0.006	-0.005	0.006
	(0.015)	(0.018)	(0.016)	(0.016)	(0.012)	(0.017)
President/secretary in office $>$ 3 yrs	0.077*	0.053	0.065	-0.008	-0.006	-0.003
	(0.046)	(0.052)	(0.048)	(0.046)	(0.036)	(0.050)
Social capital index	0.136	0.077	0.149	0.327**	0.150	0.322**
	(0.146)	(0.158)	(0.153)	(0.139)	(0.109)	(0.143)
Leader completed secondary sch.	0.094***	0.136***	0.105***	-0.025	0.030	-0.056
	(0.036)	(0.040)	(0.038)	(0.033)	(0.024)	(0.034)
Leader age	0.003	0.005***	0.003**	-0.000	0.001	-0.000
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Observations	1032	1039	941	732	725	823
Mn dep. var.	0.271	0.465	0.261	0.167	0.088	0.222

Regressors are binary variables indicating errors of inclusion or exclusion. Standard errors are clustered at the community level and regressions are simple OLS with a constant. Regressions are run on the sub-sample of households identified as poor (exclusion error) or non-poor (inclusion error) according to the household data. All panels contain only the covariates reported and a constant. * p $<$.10, ** p$<$ .05, *** p$<$.01.

When we examine the drivers of erroneous reporting of specific characteristics (Appendix Table B1), the only asset misreporting that we observe associated with social capital is television – a higher social capital index is associated with marginally statistically significant underreporting of the presence of televisions, which may explain the inclusion error associated with social capital. More educated leaders and older leaders are also significantly more likely to report that household

heads have secondary education, which may partially explain increased exclusion error associated with these characteristics.

Finally, there are a significant number of households for which leaders state that they "do not know" at least one of the characteristics that they are requested to recall on the indirect survey. 68% of households have indirect observations on all five characteristics used in the poverty index. However, 89% are missing only one (usually secondary education). Appendix D examines patterns of missing observations and examines two possible corrections. We show that poorer communities are likely to have more missing information but that there is no correlation between poverty and misreporting within villages. Furthermore, we consider two simple corrections for "don't know" responses: replacing these responses with the average of the variable from the village level and using a modified index that is weighted by the inverse of the covariances between variables, as suggested by Anderson (2008). Neither of these substantially increases targeting error.

5 Cost-benefit

The cost of directly surveying households rather than using indirect informants is the extra personnel and time need to visit households separately. This cost is a function of the number of enumerators needed and the travel costs to get additional team members to each community, as well as the time that households spend taking the surveys. The survey implemented here occurred over an 11-week period, with survey teams of three per community, with one of these enumerators acting as a "manager" and earning a slightly higher salary. Each team visited 1 to 2 communities per day, and transportation generally occurred via rented trucks. In total, 30 enumerators and 15 managers were hired for this project. The approximate cost of the entire survey was $551,000 USD to reach 862 communities and survey 10+ households in each one (around 8,000 total households). Transportation costs were a substantial portion of this budget even though teams were assigned to specific regions to minimize these costs. The total leader survey took about one hour to implement, and the total household survey took about 45 minutes per household. This does not include the

considerable transition time that it takes to arrive in a community, meet people, explain the survey process, etc.

By eliminating the household surveys (or only interviewing one or two in order to incentivize truth-telling by leaders), we estimate that the number of team members required to interview leaders and include questions about 10 households would be at most 10 managers and 10 enumerators, traveling in teams of 2. We arrived at this number of field staff by assuming people must always travel in pairs and in vehicles with 4x4 traction; this is for safety reasons and due to the poor state of rural roads and remoteness of communities. The reduced time needed to survey in each community means that the teams are now able to visit one additional community per day, further reducing transportation and time costs.

Under these assumptions about staffing and travel time, the approximate cost of surveying only leaders is $276,023 USD – a decrease of nearly 50% in the survey budget. Surveying also includes the cost of time on the part of respondents. If we assume that the survey was designed only to collect the information used in this paper – community characteristics and the small battery of asset and participation questions for each household – then the actual survey time for respondents is quite low. For the version of the survey where households and leaders are interviewed to collect the indicators and key community characteristics, we estimate a half hour for each survey. This totals 4,431 participant hours. Valued at the average agricultural wage reported in the communities of $1.28/hour, this equals $5,672 in terms of value of time. The indirect survey requires interviews with 862 leaders and includes questions about 10 households. Our field experience suggests that the indirect survey questions take about 5 minutes per household, which would add 50 minutes to each leader survey, for a total of one hour and 20 minutes. This yields a time value of the indirect survey of $1,471. These figures are likely to underestimate the cost-savings, as they do not account for the value of the time taken by the households to travel to the central site, or for leaders to assist enumerators in locating households. The total change in survey and time costs from direct to indirect surveying is from $556,672 to $275,471, a decrease of 50.5%.

This decrease in expenditures could make a substantial difference in most research budgets. It

can also be compared to the potential change in value for an example case of the tradeoffs between benefits of transfers and the costs of targeting. We use Mexico's conditional cash transfer program for schooling and health to help define some possible parameters. The annual transfers for a family with three children ages 6 months, 6 years and 9 years from Mexico's Prospera program (formerly Oportunidades/Progresa) are currently around $700 per year, depending on the selection of health and education supports (Dávila Lárraga 2016). Under our binary, household-reported measure of poverty, 58.3 percent of the 8,000 households in the survey should be classified as poor.[7] If all households defined as poor by the household responses were to receive payments, then the Prospera budget for our sample would be $3,264,800 for this group. In order to value benefits, we assume that the benefit of this transfer is that it induces children to attend school at the same rate as Progresa did. Assuming no spillover effects, an 8% return to schooling, a 5% discount rate, and rather low additionality, the benefit-cost ratio for Progresa has been calculated to be about 2.32 (Behrman et al. 2011). In our calculation, we therefore assume that accurate targeting multiplies the benefit of accurate transfers by 2.32 minus the value of the transfer itself, which yields 1.32. Because the transfer is itself is not a true resource cost, we include only the survey cost in the denominator. We assume that inaccurate transfers produce zero benefits. Under these assumptions, the ratio of benefits properly transferred to survey costs for the household survey is 7.7 ((3,264,800*1.32)/(556,672)). The simple cost of correctly transferring $1 of benefits is $0.17, and the net benefits are around $3.8 million.

If instead we engaged in targeting based upon the leader binary index, the required budget would be $2,766,400 (0.494*$700*8,000). This means that the ratio of benefits to survey costs using the indirect survey would be 11.4 ((2,766,400*1.32)/(275,471)). In order for this benefit-cost ratio to be equal to that of the household survey, the cost of the survey would have to increase by $129,381. This implies that the benefit-cost ratio of the leader survey exceeds that of the household survey if the savings in survey costs is greater than 27%. The simple cost of directly transferring $1 of benefits under this scenario is $0.11. However, because the multiplier effect is substantial

[7] We do these calculations for the full sample of households, rather than for the subsample that were used for the indirect survey.

and the transfers are rather large, the net benefits in the indirect survey case are smaller than in the household survey approach, at only $2.9 million.

Calculating cost-benefit for the case where indirect surveying is used for research and not program targeting is more difficult, as the benefits of research vary significantly across projects and are difficult to quantify.

6 Conclusion

Indirect surveying provides a simple methodology that can lower survey costs by collecting information from knowledgeable community leaders about observable characteristics of households in their community. We showed that community leaders report reasonably accurately on basic characteristics that would commonly be used to generate asset indices. These indirect informant-generated indices appear to work well as regressors in fairly standard estimation frameworks. In addition, using the leader-generated indices results in targeting error that is within the usual range for other related targeting methodologies to assess poverty using community knowledge. The survey cost savings in our application is around 50% of the budget.

The utility of indirect surveying is likely to depend upon the configuration of households and indirect informants, as well as the roles those informants play. First, indirect surveying is more likely to implemented successfully where communities are small. In our setting of rural Mexico, communities are rural and have an average number of members of 170, so most are small enough for households to know each other well.

Second, indirect surveying is more likely to work if leaders are in representative positions and do not have an incentive to misreport. In our setting, community leaders are elected and had an incentive to report correctly, as they knew that we would also be asking similar questions to the same households. One could replicate this incentive by surveying one household at random (rather than many), or by also interviewing a second knowledgeable informant, such as a women's group head or labor leader to triangulate information, and still benefit from significant cost reductions.

These features are not unique to Mexico, but are indeed likely to be present in many parts of the globe where there are sparsely dispersed populations. Our findings are likely most relevant for rural areas, particularly those with poor road quality or rugged terrain, where transport costs will be high. However, given the known relationships between remoteness and poverty itself, these are likely the types of areas where poverty alleviation programs themselves may matter most in a range of countries.

Finally, we acknowledge that our study provides an important preliminary test of a method, but also raises additional questions that should be answered before trying to scale up this methodology. In particular, it would be useful to probe the limits of indirect surveying, including the number of households and covariates about which informants can accurately respond. In terms of the targeting exercise, it would also be useful to compare this method more directly with other lower-cost methods of targeting, including geographic approaches or subjective leader measures, and ideally with detailed consumption or health data to assess crucial short-run measures of poverty and well-being. In addition to the technical accuracy of targeting methods, their perceived legitimacy may play an important role in the ultimate success of the program, particularly for programs that require behavioral changes. This will require attention to both targeting errors and the local relevance of the indicators used to assess poverty. Furthermore, in our application leaders had no reason to think that we would be using this information to target a program. We cannot assess whether or not their responses might be different if they believed this to be the case, but think this could be an interesting behavior to analyze before widely applying a method of this nature. We hope that future research will illuminate more information about these parameters in the interest of lowering survey costs to make more funds available for vital anti-poverty programs or research endeavors.

References

Abowd, J. M. & Stinson, M. H. (2013), 'Estimating measurement error in annual job earnings: a comparison of survey and administrative data', *Review of Economics and Statistics* **95**(5), 1451–1467.

Alatas, V., Banerjee, A., Hanna, R., Olken, B. A. & Tobias, J. (2012), 'Targeting the Poor: Evidence from a Field Experiment in Indonesia', *American Economic Review* **102**(4), 1206–1240. **URL:** *https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4156293/*

Alatas, V., Purnamasari, R., Wai-Poi, M., Banerjee, A., Olken, B. A. & Hanna, R. (2016), 'Self-targeting: Evidence from a field experiment in Indonesia', *Journal of Political Economy* **124**(2), 371–427.

Alderman, H. (2002), 'Do local officials know something we don't? Decentralization of targeted transfers in Albania', *Journal of Public Economics* **83**(3), 375–404.

Alix-Garcia, J. M., Sims, K. R., Orozco-Olvera, V. H., Costica, L., Medina, J. D. F. & Monroy, S. R. (2018), 'Land conservation payments conserve communal social capital', *Proceedings of the National Academies of Science* **115**, 7016–7021.

Alix-Garcia, J. M., Sims, K. R., Orozco-Olvera, V. & et al. (2018), 'Evaluation of Mexico's Payments for Environmental Services Program: 2011-2014, Report for CONAFOR-CONEVAL'.

Anderson, M. L. (2008), 'Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects', *Journal of the American Statistical Association* **103**(484), 1481–1495.

Bahri, S. & Merttens, F. (2018), 'Evaluation of the Kenya Hunger Safety Net Programme Phase 2: Cost-efficiency analysis final report'.
URL: *https://www.opml.co.uk/files/Publications/a0013-evaluation-kenya-hunger-safety-net-programme/hsnp-cost-efficiency-report.pdf?noredirect=1*

Basurto, P. M., Dupas, P. & Robinson, J. (2017), 'Decentralization and efficiency of subsidy targeting: Evidence from chiefs in rural Malawi', *NBER Working Paper No. 23383* .

Behrman, J. R., Parker, S. W. & Todd, P. E. (2011), 'Do conditional cash transfers for schooling generate lasting benefits? A five-year followup of PROGRESA/Oportunidades', *Journal of Human Resources* **46**(1), 93–122.

Bergeron, G., Morris, S. S. & Banegas, J. M. M. (1998), 'How reliable are group informant ratings? A test of food security ratings in Honduras', *World Development* **26**(10), 1893–1902.

Bigman, D. & Fofack, H. (2000), 'Geographical targeting for poverty alleviation: An introduction to the special issue', *World Bank Economic Review* **14**(1), 129–145.

Brown, C., Ravallion, M. & Walle, D. V. D. (2018), 'A Poor Means Test? Econometric Targeting in Africa?', *Journal of Development Economics* **134**, 109–124.

Caldés, N., Coady, D. & Maluccio, J. A. (2006), 'The cost of poverty alleviation transfer programs: a comparative analysis of three programs in Latin America', *World Development* **34**(5), 818–837.

Chakraborty, N. M., Fry, K., Behl, R. & Longfield, K. (2016), 'Simplified asset indices to measure wealth and equity in health programs: a reliability and validity analysis using survey data from 16 countries', *Global Health: Science and Practice* **4**(1), 141–154.

Chambers, R. (1994), 'Participatory rural appraisal (PRA): Analysis of experience', *World Development* **22**(9), 1253–1268.

CONAPO (2019), 'Índices de Marginación'. Consejo Nacional de Población.
URL: *http://www.conapo.gob.mx/es/CONAPO/Indices_de_Marginacion*

Dávila Lárraga, L. G. (2016), How does Prospera Work?: Best Practices in the Implementation of Conditional Cash Transfer Programs in Latin America and the Caribbean, Technical report, Inter-American Development Bank.

Davis, B. (2003), *Choosing a method for poverty mapping*, Food & Agriculture Org.

De Janvry, A., Gordillo, G. & Sadoulet, E. (1997), *Mexico's Second Agrarian Reform: household and community responses, 1990-1994*, Center for US-Mexican Studies.

Decancq, K. & Lugo, M. A. (2013), 'Weights in multidimensional indices of well-being: An overview', *Econometric Reviews* **32**(1), 7–34.

DHS (2020), 'DHS Wealth Index Construction'.
URL: *https://www.dhsprogram.com/topics/wealth-index/Wealth-Index-Construction.cfm*

Elbers, C., Fujii, T., Lanjouw, P., Yin, W. & Özler, B. (2004), *Poverty alleviation through geographic targeting: How much does disaggregation help?*, The World Bank.

Filmer, D. & Pritchett, L. H. (2001), 'Estimating wealth effects without expenditure data – or tears: an application to educational enrollments in states of India', *Demography* **38**(1), 115–132.

Filmer, D. & Scott, K. (2008), *Assessing asset indices*, The World Bank.

Finan, F. & Schechter, L. (2012), 'Vote–buying and reciprocity', *Econometrica* **80**(2), 863–881.
URL: *https://www.jstor.org/stable/41493836*

Fox, J. (2007), *Accountability Politics: Power and Voice in Rural Mexico*, Oxford University Press, New York, NY.

Frenk, S. M., Anderson, S. L., Chaves, M. & Martin, N. (2010), 'Assessing the validity of key informant reports about congregations' social composition', *Sociology of Religion* **72**(1), 78–90.

Galasso, E. & Ravallion, M. (2005), 'Decentralized targeting of an anti-poverty program', *Journal of Public Economics* **89**(4), 705–727.

Garrett, J. L. & Downen, J. (2001), *Rapid assessments in urban areas: lessons from Bangladesh and Tanzania*, International Food Policy Research Institute (IFPRI).

Genicot, G. & Ray, D. (2003), 'Group formation in risk-sharing arrangements', *The Review of Economic Studies* **70**(1), 87–113.

Imbens, G. W. & Wooldridge, J. M. (2009), 'Recent developments in the econometrics of program evaluation', *Journal of Economic Literature* **47**(1), 5–86.

Inglehart, R., Haerpfer, C., Moreno, A., Welzel, C., Kizilova, K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E. & Puranen, B. (2014), 'World Values Survey: All Rounds - Country-Pooled Datafile'.
URL: *http://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp*

Karlan, D. & Thuysbaert, B. (2016), 'Targeting ultra-poor households in Honduras and Peru', *World Bank Economic Review* **33**(1), 63–94.

Keane, A., Lund, J. F., Bluwstein, J., Burgess, N. D., Nielsen, M. R. & Homewood, K. (2020), 'Impact of Tanzania's Wildlife Management Areas on Household Wealth', *Nature Sustainability* **3**(3), 226–233.

Knight, A. (1986), *The Mexican Revolution*, University of Nebraska Press, Lincoln, NE.

Kondylis, F., Mueller, V. & Zhu, S. (2015), 'Measuring agricultural knowledge and adoption', *Agricultural Economics* **46**, 449–462.

Kumar, K. (1989), *Conducting key informant interviews in developing countries*, Agency for International Development Washington DC.

Macours, K. (2003), 'Comparing a Direct with an Indirect Approach to Collecting Household Level Data: Who tells the truth about what?'. Unpublished Manuscript.

McKenzie, D. J. (2005), 'Measuring inequality with asset indicators', *Journal of Population Economics* **18**(2), 229–260.

Merino, L. & Martinez, A. E. (2014), *A vuelo de pájaro: las condiciones de las comunidades con bosques templados en México*, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO).

Olken, B. A. (2016), 'Hassles versus prices', *Science* **353**(6302), 864–865.

Phillips, L. W. (1981), 'Assessing measurement error in key informant reports: A methodological note on organizational analysis in marketing', *Journal of Marketing Research* pp. 395–415.

Rigol, N., Hussam, R. & Roth, B. (2017), 'Targeting high ability entrepreneurs using community information: Mechanism design in the field'.
URL: *https://economics.mit.edu/files/14389*

Romney, A. K., Weller, S. C. & Batchelder, W. H. (1986), 'Culture as consensus: A theory of culture and informant accuracy', *American Anthropologist* **88**(2), 313–338.

Rutstein, S. & Johnson, K. (2004), 'Dhs comparative reports no. 6: The dhs wealth index.', *Calverton: ORC Macro* .

Schreiner, M. (2017), 'Simple Poverty Scorecard Poverty Assessment Tool Mexico'.
URL: *http://www.simplepovertyscorecard.com/MEX_2014_ENG.pdf*

Skoufias, E., Diamond, A., Vinha, K., Gill, M. & Dellepiane, M. R. (2020), 'Estimating poverty rates in subnational populations of interest: An assessment of the simple poverty scorecard', *World Development* **129**, 104887.

Stecklov, G., Winters, P., Todd, J. & Regalia, F. (2006), 'Demographic externalities from poverty programs in developing countries: Experimental evidence from latin america'. Working paper: Washington DC: Department of Economics, American University.

Stoeffler, Q., Mills, B. & Del Ninno, C. (2016), 'Reaching the Poor: cash transfer program targeting in Cameroon', *World Development* **83**, 244–263.

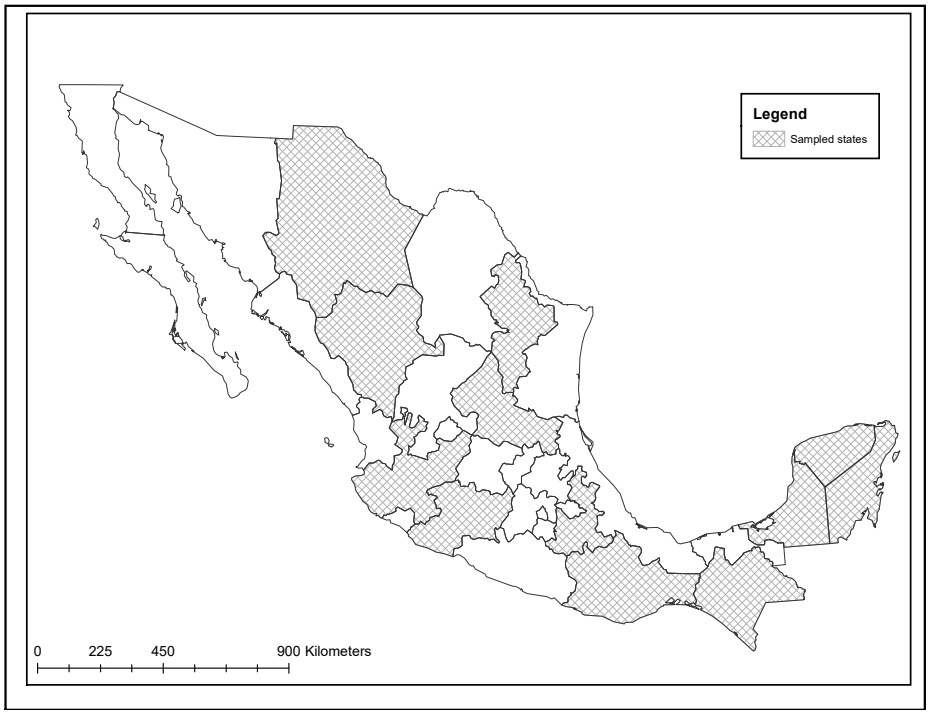
Sustainable Development Solutions Network (2015), *Data for development: a needs assessment for SDG monitoring and statistical capacity development*, Sustainable Development Solutions Network.

Takasaki, Y., Barham, B. L. & Coomes, O. T. (2000), 'Rapid rural appraisal in humid tropical forests: an asset possession-based approach and validation methods for wealth assessment among forest peasant households', *World Development* **28**(11), 1961–1977.

Young, F. W. & Young, R. (1961), 'Key informant reliability in rural Mexican villages', *Human Organization* **20**(3), 141–148.
URL: *https://www.jstor.org/stable/44124921*

Appendix A Supplementary tables and figures

Figure A1: Sampled states



States within which surveys occurred are cross-hatched.

Table A1: Sample versus universe of PSA applicants and universe of ejidos

	Sampled	Not sampled		Not in universe	
	(1)	(2)	(3)	(4)	(5)
			(1) v (2)		(1) v (4)
Ln(ejido area, ha)	8.21	8.27	-0.04	6.84	0.79
Ln(common property area, ha)	7.81	7.76	0.02	4.08	1.02
Ln(ejidatarios)	4.60	4.60	-0.00	4.30	0.25
Ln(non-ejidatarios)	1.35	1.42	-0.03	1.73	-0.17
Ln(area submitted, ha)	6.70	6.68	0.01		
Average slope (deg)	11.00	12.28	-0.13		
Average elevation (m)	1309.40	1388.84	-0.06		
Ln(km to any road)	7.64	7.77	-0.09		
Ln(km to major city)	4.60	4.43	0.20		
Ln(km to city $>$ 5,000)	3.23	3.23	-0.00		
Mean canopy cover	50.91	42.43	0.20		
% indigenous in municipality	0.33	0.16	0.29		
Deforestation risk (INE)	0.06	0.06	-0.02		
Obs	862	1977		27497	

Columns (1) and (2) show mean values of variables for communities that were sampled versus those that were not sampled but did apply for PSA. Column (3) presents the normalized difference between these two. The fourth column shows means for variables available in Mexico's PHINA database, which gives information on all agrarian communities existing as of November 2016. The last column presents the normalized differences in means between the sampled communities and nationwide PHINA database.

Table A2: Correlations between leader and household reports

	Cell phone (1)	Secondary school (2)	HH size (3)	Wall type (4)	TV (5)	Car (6)	# livestock (7)	Forest work (8)
A: OLS with household order								
Household report	0.496***	0.383***	0.697***	0.655***	0.485***	0.737***	0.782***	0.340***
	(0.030)	(0.052)	(0.031)	(0.026)	(0.038)	(0.024)	(0.024)	(0.037)
Second household	-0.063	-0.033	-0.050	0.003	0.050	-0.042	0.024	-0.022
	(0.041)	(0.069)	(0.051)	(0.034)	(0.049)	(0.032)	(0.032)	(0.047)
Third household	-0.025	-0.007	-0.019	0.057*	-0.001	-0.020	-0.002	-0.020
	(0.040)	(0.077)	(0.044)	(0.033)	(0.051)	(0.031)	(0.035)	(0.044)
B: Ejido FE with household older								
Household report	0.329***	0.322***	0.635***	0.481***	0.321***	0.643***	0.687***	0.184***
	(0.031)	(0.040)	(0.030)	(0.027)	(0.032)	(0.028)	(0.027)	(0.027)
Second household	-0.027	0.000	-0.036	-0.002	0.030	-0.047	0.014	-0.039
	(0.041)	(0.054)	(0.041)	(0.034)	(0.042)	(0.035)	(0.033)	(0.036)
Third household	-0.018	0.055	0.009	0.034	0.017	-0.038	-0.015	-0.041
	(0.041)	(0.058)	(0.040)	(0.034)	(0.043)	(0.035)	(0.033)	(0.034)

Each panel represents a different estimation. Dependent variables are leader reports, and household report indicates the household response for the same variable. Second and third household indicates the interaction between the household reported variable and the order of questioning. Therefore, second household measures the difference in correlation between the leader and the household report for the second household relative to the first one. Standard errors are clustered at the ejido level.* p <.10, ** p< .05, *** p<.01.

Table A3: Correlations and dependence upon interview order

	PIB	IPW	PCA
	(1)	(2)	(3)
A: OLS with household order			
Household index	0.681***	0.674***	0.684***
	(0.030)	(0.032)	(0.030)
Second household	-0.020	0.011	-0.018
	(0.041)	(0.044)	(0.041)
Third household	0.045	0.039	0.028
	(0.038)	(0.041)	(0.039)
B: Ejido FE with household older			
Household index	0.527***	0.501***	0.522***
	(0.034)	(0.035)	(0.034)
Second household	-0.020	0.003	-0.026
	(0.043)	(0.043)	(0.043)
Third household	0.072*	0.066	0.060
	(0.043)	(0.042)	(0.043)

Each panel represents a different estimation. Dependent variables are leader indices, and household index indicates the index constructed from the household responses for the same variable. Second and third household indicates the interaction between the household reported index and the order of questioning. Therefore, second household measures the difference in correlation between the leader and the household index for the second household relative to the first one. Standard errors are clustered at the ejido level.* p <.10, ** p< .05, *** p<.01.

Table A4: Summary of aggregated community indices

	(1) Leaders	(2) Households	(3) Norm diff	(4) ρ
Simple index	0.489	0.442	-0.152	0.733
Inverse proportion weighted index	1.405	1.204	-0.259	0.741
Principle components index	-0.021	0.012	0.022	0.735
Observations	735	735	735	735

Columns (1) and (2) report mean reported values for indices calculated using assets reported by leaders and households, respectively, aggregated to the level of the whole ejido. Column (3) lists the normalized difference of the two means. ρ is the correlation coefficient between the indices constructed with household and leader reports. The total number of communities included is less than 862 due to missing data.

Table A5: Classification of poverty using household versus leader reporting with abbreviated indices

	Household responses		
	Non-poor	Poor	Total
Leader responses			
Simple index			
Non-poor	53.5	10.1	63.5
Poor	6.6	29.8	36.5
Total	60.1	39.9	100.0
	Non-poor	Poor	Total
Inverse proportion weighted index			
Non-poor	16.8	26.9	43.8
Poor	3.1	53.2	56.2
Total	19.9	80.1	100.0
	Non-poor	Poor	Total
Principle components index			
Non-poor	51.5	11.3	62.8
Poor	6.7	30.5	37.2
Total	58.2	41.8	100.0

Cells show percentages in each category. In all cases, the poor are those with less than the median level of the index calculated across the entire sample using the index developed from household reports.

Table A6: Correlates of errors of inclusion and exclusion

	Exclusion error			Inclusion error		
	Binary	IPW	PCA	Binary	IPW	PCA
A: Community size						
IHS(Community members w/ full rights)	-0.001	0.004	0.002	-0.005	-0.002	0.005
	(0.015)	(0.018)	(0.016)	(0.015)	(0.012)	(0.016)
B: Leadership capture						
President/secretary in office $>$ 3 yrs	0.069	0.053	0.054	-0.020	-0.008	-0.021
	(0.045)	(0.051)	(0.047)	(0.044)	(0.035)	(0.048)
C: Social capital						
Social capital index	0.041	-0.040	0.050	0.334**	0.126	0.334**
	(0.138)	(0.152)	(0.145)	(0.140)	(0.115)	(0.143)
D: Education/age						
Leader completed secondary sch.	0.092**	0.135***	0.103***	-0.034	0.025	-0.062*
	(0.036)	(0.039)	(0.038)	(0.033)	(0.025)	(0.034)
Leader age	0.002	0.005***	0.003*	-0.001	0.001	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
Observations	1032	1039	941	732	725	823
Mn dep. var.						

Regressors are binary variables indicating errors of inclusion or exclusion. Standard errors are clustered at the community level and regressions are simple OLS with a constant. Regressions are run on the sub-sample of households identified as poor (exclusion error) or non-poor (inclusion error) according to the household data. All panels contain only the covariates reported and a constant. * p $<$.10, ** p$<$ .05, *** p$<$.01.

Appendix B Correlates of inaccurate reporting

This section examines whether or not community characteristics influence the magnitude of misreporting, using the same type of analysis as is contained in the main text of the paper.

Table B1 reports regressions of the difference between leader and household reports of the variables of interest. The table shows only those households and leaders where there were no missing values for any of the items. In these regressions we do not examine number of meetings attended, since the results above suggest that leaders largely report the same value for all households within their community.

We find that accurate reporting of individual characteristics can be influenced by community size – having more potential members upon whom to report decreases accuracy for reporting on the presence of televisions and mobile phones, household size, participation in collective forest work and household size. Higher social capital (panel (c)) results in lower reporting of television sets and household size, which explains the correlation with exclusion error of this covariate. Leaders with secondary education tend to be more likely to misreport that other household heads have secondary education, and that they have larger households. Older leaders over-report secondary schooling of other household heads and participation in forest work. The over-reporting of secondary education increases the asset indices and illuminates the source of the exclusion error associated with these characteristics.

Table B1: Difference between leader and household report

	Sec. sch. (0/1)	Walls (0/1)	Television (0/1)	Car (0/1)	Mobile (0/1)	Forest wk. (0/1)	HH size (num.)	Cattle (num.)
A: Community membership size								
IHS(Community members)	0.013	-0.002	0.019*	-0.011	-0.029**	0.027*	0.023**	0.028
	(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.010)	(0.018)
B: Leadership capture								
President/secretary in office $>$ 3 yrs	-0.021	-0.013	0.016	0.023	0.062	0.037	0.006	-0.061
	(0.036)	(0.027)	(0.021)	(0.027)	(0.041)	(0.042)	(0.024)	(0.049)
C: Social capital								
Social capital index	0.026	0.132	-0.136*	-0.022	-0.191	0.115	-0.113*	0.131
	(0.100)	(0.084)	(0.081)	(0.078)	(0.120)	(0.110)	(0.067)	(0.161)
D: Education/age								
Leader completed secondary sch.	0.085***	-0.020	0.028	0.014	-0.002	0.005	0.049**	-0.059
	(0.023)	(0.024)	(0.019)	(0.021)	(0.031)	(0.028)	(0.020)	(0.039)
Leader age	0.002*	-0.001	0.000	0.000	0.001	0.002*	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
E: All								
IHS(Community members)	0.009	-0.003	0.019*	-0.010	-0.027*	0.029**	0.022**	0.029
	(0.010)	(0.010)	(0.010)	(0.010)	(0.014)	(0.014)	(0.010)	(0.019)
President/secretary in office $>$ 3 yrs	-0.018	-0.005	0.014	0.018	0.040	0.058	0.006	-0.043
	(0.035)	(0.028)	(0.022)	(0.029)	(0.041)	(0.043)	(0.025)	(0.051)
Social capital index	0.049	0.113	-0.131	-0.004	-0.149	0.176	-0.094	0.090
	(0.100)	(0.085)	(0.083)	(0.083)	(0.125)	(0.117)	(0.072)	(0.167)
Leader completed secondary sch.	0.085***	-0.017	0.021	0.016	0.000	0.002	0.042**	-0.062
	(0.023)	(0.024)	(0.020)	(0.022)	(0.031)	(0.029)	(0.020)	(0.040)
Leader age	0.002*	-0.001	0.000	0.000	0.001	0.003*	0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	1741	1741	1741	1741	1741	1734	1690	1646
Mean dep. var.	0.077	0.011	0.040	0.016	0.094	0.058	0.042	0.122

Notes: Regressors are the value of the difference between the leader and household report of the same measure. Positive values indicate leader values greater than household values. Standard errors are clustered at the community level and regressions are simple OLS with a constant. Only observations with all variables non-missing are included.* p $<$.10, ** p$<$ .05, *** p$<$.01.

Appendix C Alternative assumptions on the "truth"

Assuming that both measures have reasons to be imperfect, Table C2 shows a series of weighted reliability ratios following Abowd & Stinson (2013). The rows of the table vary what is taken as "truth" in the comparison. For example, in the first row it is assumed that the leader index is true, and the reliability ratio is calculated relative to this. In the last row the household index is assumed to be true, and each of the rows in between give some weight to each. Considering first the simple index, we observe that the range of reliability is wide for both the leader and household aggregations. In the former case the ratios run from 0.44 to 1, and in the latter from 0.33 to 1. It is interesting that the range is wider for the household index than for the leader index, a pattern which holds up across aggregation techniques. When the "true" index is taken to be an equally weighted average of both indices, the reliability ratios are well above 0.80.

Table C2: Summary of weighted reliability ratios by index

True Model	Simple		IPW		PCA	
Weight(Leader,HH)	Leader	Household	Leader	Household	Leader	Household
	(1)	(2)	(3)	(4)	(5)	(6)
1,0	1.000	0.325	1.000	0.282	1.000	0.333
0.9,0.1	0.994	0.453	0.995	0.419	0.994	0.460
0.5,0.5	0.860	0.831	0.864	0.821	0.857	0.833
0.1,0.9	0.545	0.993	0.560	0.993	0.536	0.993
0,1	0.439	1.000	0.457	1.000	0.428	1.000

The table contains the weighted reliability ratios following Abowd & Stinson (2013). The formula for the statistic is $1 - \mathrm{V}ar(\mu)/\mathrm{V}ar(y^*)$ where $y^* = \omega leader + (1-\omega)HH$ and $\mu = y - y^*$ is measurement error.

Appendix D Missing data

Leaders rather frequently responded that they were unsure of their responses to the indirect survey questions (Table D1). The most common problem was being able to identify if the household head had completed secondary schooling (Table D2).

"Don't know" responses are, of course, ubiquitous in survey settings. For the majority of

Table D1: Frequency of "don't know" responses in indirect survey

Number missing	Frequency	Percent
0	1764	68.29
1	537	20.79
2	216	8.36
3	55	2.13
4	8	0.31
5	3	0.12
Total	2583	100.00

Cells show frequency and percentages for each number of missing responses.

Table D2: Proportion "don't know" responses by question

	Mean	SD	Obs
Car, don't know	0.017	0.131	2583
Wall type, don't know	0.012	0.107	2583
Television, don't know	0.128	0.334	2583
Mobile phone, don't know	0.111	0.314	2583
Secondary education, don't know	0.189	0.392	2583

Table shows proportions of missing responses for each variable in the indirect survey.

this paper, we used only the set of households for which leaders responded to all indirect survey questions. However, we are concerned about two issues. The first is that those about whom leaders know less may be influenced by the household's poverty level, and the second is the potential error generated by using other index constructions that allow for missing data. Regarding the first issue, Table D3 shows the correlations between having at least one missing observation, household level poverty as measured by a full household wealth index, and leader characteristics.

Table D3: Correlates of "don't know" responses, community level

	OLS	Ejido FE
Full assets index, hh survey	-0.113**	0.051
	(0.048)	(0.060)
Leader age	0.001	
	(0.001)	
Leader completed secondary sch.	-0.039	
	(0.024)	
Social capital index	-0.184*	
	(0.107)	
IHS(Community members)	0.033***	
	(0.012)	

Notes: Regressor is an indicator equal to one if there is at least one ' response in the indirect survey. The asset index is calculated from all the household data reported at the household level. Standard errors are clustered at the ejidal level and regressions are simple OLS with a constant. The OLS estimations include state fixed effects.

We note that across villages (column (1)), wealthier households are less likely to have missing information. Communities with larger numbers of members are also more likely to have "don't know" responses, while higher levels of social capital are weakly associated with fewer missing responses. Inclusion of fixed effects eliminates the correlation between misreporting and the assets index. This suggests that it is more likely that poorer villages will have more misreporting, not that poorer households within a particular village are more likely to suffer from missing responses.

Next we examine targeting error in the event that we (1) replace "don't know" responses with the average of the variable from the village level and (2) use a modified index that is weighted by

the inverse of the covariances between variables, as suggested in Anderson (2008).

Simply replacing the "don't know" responses with the community and then the state average of the covariates does not substantially change the percentage of households that are correctly classified. For the simple index, the correct poverty classification decreases from 77.2% of households to to 73.2%, for the inverse proportion weighted index it increases from 68.2 to 68.9%, and for the principal components index it decreases from 75.7 to 71.4%. Using the index based on Anderson (2008) yields correct classifications in 71.4% of cases.

Table D4: Classification of poverty using household versus leader reporting

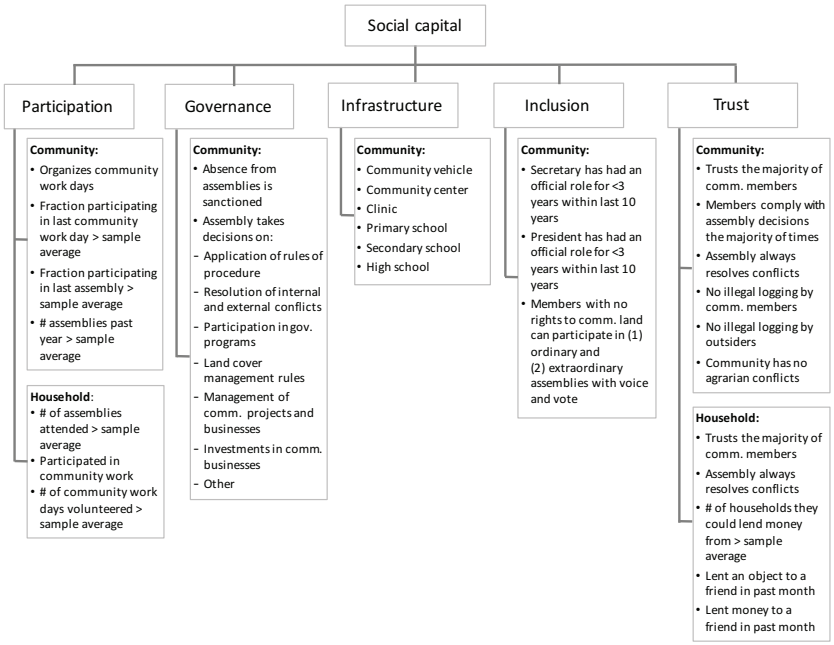
	Household responses		
	Non-poor	Poor	Total
Leader responses			
Simple index			
Non-poor	34.8	21.0	55.7
Poor	5.8	38.4	44.3
Total	40.6	59.4	100.0
	Non-poor	Poor	Total
Inverse proportion weighted index			
Non-poor	36.2	27.1	63.3
Poor	4.0	32.7	36.7
Total	40.2	59.8	100.0
	Non-poor	Poor	Total
Principle components index			
Non-poor	35.1	14.6	49.7
Poor	10.6	39.6	50.3
Total	45.7	54.3	100.0
	Non-poor	Poor	Total
Anderson index			
Non-poor	33.9	19.8	53.7
Poor	8.8	37.5	46.3
Total	42.7	57.3	100.0

Cells show percentages in each category. In all cases, the poor are those with less than the median level of the index calculated across the entire sample using the index developed from household reports. Missing values in leader observations for binary, IPW and PCA indices replaced with community mean value for that variable. The Anderson index is calculated as described in Anderson (2008)

Appendix E Social capital index

This paper uses the aggregate social capital index described in Alix-Garcia, Sims, Orozco-Olvera, Costica, Medina & Monroy (2018). The measurement of social capital is challenging, as definitions of its content can vary depending upon discipline and context. We built upon a social capital index developed for Mexican forest communities by Merino & Martinez (2014) and modified it so that it would also be relevant for areas without a long history of forest management and reflecting a broader definition of social capital. Furthermore, we included questions about trust and infrastructure from the World Values Survey (Inglehart et al. 2014). A full list of the variables can be seen in figure E1.

Figure E1: Variables included in social capital index



This figure originally appeared in Alix-Garcia, Sims, Orozco-Olvera, Costica, Medina & Monroy (2018).

In preliminary analysis we used a variety of methods common in the measurement literature to aggregate these values, including simple summation of the presence or absence of particular characteristics, principal components, polychoric principal components (which takes into account categorical variables), and inverse proportion weighting (which applies higher weight to very scarce

items). Because results vary little depending upon these aggregations, we present here the simple summation approach. For this approach, we created sub-indices for each of the branches of the figure (participation, infrastructure, etc.). We calculated the proportion of positive indicators over the total possible. We then created a total social capital index by summing up the sub-indices and dividing by five. For survey questions that were not originally binary, we computed a variable equal to one if the original value was above the mean.