Community-Based versus Statistical Targeting of Anti-Poverty Programs: Evidence from Burkina Faso

Michael Hillebrecht [1,2] *

Stefan Klonner[1]
Klonner@uni-heidelberg.de

Athanase Narangoro Pacere[3]
Athanpace@yahoo.fr

Aurélia Souares[4]
Souares@uni-heidelberg.de

[1] South Asia Institute, Heidelberg University
Im Neuenheimer Feld 330, 69120 Heidelberg, Germany

[2] Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ)
Dag-Hammarskjöld-Weg 1-5, 65760 Eschborn, Germany

[3] Centre de Recherche en Santé de Nouna, Ministère de la Santé
BP 02 Nouna, Rue Namory Keita, Sect 6 Nouna, Burkina Faso

[4] Institute of Global Health, Heidelberg University
Im Neuenheimer Feld 324, 69120 Heidelberg, Germany

Acknowledgements

We thank Alice Mesnard, Ken Harttgen, Kristina Kis-Katos, and two anonymous reviewers for valuable comments and Germain Savadogo for excellent background information. Jan Melcher and Marwa Mahran provided invaluable research assistance. The usual disclaimer applies.

*Corresponding author: Schleicher@uni-heidelberg.de. Present address: Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), Dag-Hammarskjöld-Weg 1-5, 65760 Eschborn, Germany.

Abstract

Targeting of governmental welfare programs in low-income countries commonly relies on statistical procedures involving household-level data, while smaller-scale programs often employ community-based targeting, where village communities themselves identify beneficiaries. Combining original data from community-based targeting exercises in Burkina Faso with a household survey we compare the targeting accuracy of community-based targeting with four common statistical targeting methods when the objective is to target consumption-poor households. We find that community-based targeting is substantially less accurate than statistical targeting in villages, while it is as accurate as the much more costly statistical methods in semi-urban areas. We show that this difference is due to differences in poverty concepts held by rural and urban communities. Its large cost advantage makes community-based targeting far more cost-effective than statistical targeting for common amounts of welfare program benefits.

Keywords: *Targeting, Community-based Targeting, Welfare Programs, Poverty, Proxy-means Testing*

JEL Codes: I13, I38, O15

1 Introduction

Poverty reduction programs designed to directly improve the well-being of the poor have become increasingly popular around the global South (Honorati et al., 2015). When a program's purpose is to maximize poverty reduction under a limited budget, its success depends on the program's individual welfare effects as well as on the accuracy of the underlying targeting method (Ravallion, 1993). In low-income countries, where administrative data on households' incomes ('means') are typically unavailable, targeting of welfare programs tends to rely on statistical procedures processing suitable proxies of households' means in a Proxy-Means Test (PMT). Alternatively, targeting may be decentralized through Community-based Targeting (CBT), where the choice of beneficiaries is delegated to local communities (Ravallion, 2003).

Existing studies mention superior cost-effectiveness (Chambers, 1994b) and higher satisfaction rates (Alatas et al., 2012; Schüring, 2014; Robertson et al., 2014) as two advantages of community-based over statistical targeting methods. One reason for the latter is that community members better understand participative procedures while statistical targeting procedures are often perceived as a black box (Brown et al., 2018). In addition, local participative assessments have been found to consider more poverty dimensions than only consumption (Alatas et al., 2012; Van Campenhout, 2007) and to improve local ownership and sustainability of the underlying program (Robertson et al., 2014). On the downside, decentralization of political decision making may be susceptible to capture by local elites (Conning and Kevane, 2002; Bardhan and Mookherjee, 2006). In contrast, statistical targeting is the more objective and easier-to-replicate approach. With its more centralized organizational structure it hast the potential to reduce potential principal-agent problems (Ravallion, 1993). On the other hand, it may suffer severely when only a small number of officials is corrupt (Niehaus and Atanassova, 2013).

In this paper we investigate which method, community-based or statistical targeting, targets consumption-poor households more accurately. We compare community-based targeting with four frequently used statistical methods. In addition, we study how the targeting accuracy of the two families of methods compares across rural and semi-urban sectors, and across community characteristics. Finally, we evaluate the cost-effectiveness of each procedure. In our empirical analysis, we combine original data from community targeting exercises conducted in 36 villages and 22 semi-urban neighborhoods in north-western Burkina Faso in 2009 with a household survey that includes consumption as well as common proxy-means variables. We also use census data to construct community characteristics and administrative cost data for a cost-benefit analysis.

All statistical targeting methods that we consider have in common that they involve an index which is calculated as a weighted average of transformed proxy-means variables, while they differ along three dimensions, the set of indicators, the way these indicators are transformed into proxy-variables, and the weights used to aggregate the proxy-variables into a single index. The *Econometric PMT* is based on a statistical model and employs a large number of indicators available in census data (Brown et al., 2018). The weights are obtained from a regression of consumption on these indicators, or proxy-means variables (Alatas et al., 2012; Filmer and Scott, 2012; Klasen and Lange, 2014). Alternatively, the weights can be obtained from the joint distribution of the indicators themselves through Principal Component Analysis (Filmer and Pritchett, 2001). This index is often referred to as *Asset Index* (Filmer and Scott, 2012). Third, we consider the *Poverty Scorecard Index*, a targeting tool popular among practitioners around the globe. It comprises only a small number of transformed indicators (Schreiner, 2015). Finally, we calculate a *Multidimensional Poverty Index*. Following Alkire and Santos (2010), in this approach all indicators are first transformed into binary deprivation indicators and the index equals a weighted deprivation count.

We find that community-based targeting is substantially less accurate than any of the statistical methods in villages, where it fails to improve on random targeting. In contrast, CBT is almost as accurate as the much more expensive econometric proxy-means test and more accurate than any of the other, less costly statistical methods in the overall less impoverished semi-urban area of our study site, where it improves on random targeting by about 30 percent. In additional analyses, we show that this difference in CBT's accuracy is due to a difference in poverty concepts held by rural and urban communities: while CBT beneficiaries in both sectors lack assets and education, rural communities put large weights on morbidity and household demographics, in particular the share of elderly household members, both of which are positively correlated with per capita consumption. On the other hand, we find no community-level predictors, such as the degree of economic inequality or ethnic diversity, explaining the rural-urban difference in CBT's accuracy.

In a cost-benefit analysis, we find statistical targeting to be more cost-effective than community-based targeting only for very large program budgets, even in the rural areas. Hence, for the benefits usually encountered in welfare programs in low-income countries, community-based targeting is by far the more cost-effective method in the sub-Saharan African context studied here. Moreover we find that the less expensive statistical methods, which do not require consumption data for calibration, have no cost-effectiveness advantage when community-based and econometric targeting are available to the policy maker.

Within the vast economic literature on targeting of welfare programs, our study contributes to the topic of targeting accuracy.[1] Until recently, the literature on this topic has followed a somewhat narrow approach, where the focus is on one specific targeted anti-poverty program at a time and targeting accuracy is measured by the share of households meeting the program's targeting criteria in all beneficiary households (Ravallion, 2009). These studies mainly consider statistical targeting methods and usually compare a household's self-reported eligibility with hypothetical eligibility calculated from socio-economic household characteristics according to the program's eligibility rules.[2]

[1] Other prominent themes are leakage (Alatas et al., 2013b; Niehaus and Atanassova, 2013), elite capture (Alatas et al., 2012, 2013a; Panda, 2015), agency problems in decentralization (Galasso and Ravallion, 2005; Banerjee et al., 2014), and communities' poverty perceptions (Van Campenhout, 2007; Kebede, 2009; Alatas et al., 2012).

[2] Some prominent examples are Banerjee et al. (2007) for food distribution, housing and employment schemes in India, Skoufias et al. (2001) for Progresa in Mexico, Ahmed and Bouis (2002) for food subsidies in Egypt, Handa et al. (2012) for cash transfer programs in Malawi and Kenya, and Castañeda (2005) for Columbia's SISBEN. We describe this latter program in more detail in section 2. The review by Coady et al. (2004a) summarizes studies of 122 anti-poverty programs in 48 countries.

A more recent set of studies has taken a broader approach by comparing alternative targeting methods in one empirical setting. This small but rapidly growing literature employs consumption as the reference and the variation in targeting methods comes either from alternative treatments (Alatas et al., 2012; Sabates-Wheeler et al., 2015) or hypothetical calculations with household-level survey data (Grosh and Baker, 1995; Filmer and Scott, 2012; Klasen and Lange, 2014; Karlan and Thuysbaert, 2016; Brown et al., 2018; Stoeffler et al., 2016). Within this comparative targeting accuracy literature one may distinguish between a first branch that compares various alternative statistical targeting methods with each other (Grosh and Baker, 1995; Filmer and Scott, 2012; Klasen and Lange, 2014; Brown et al., 2018), and a second branch, whose subject is the comparative assessment of statistical and community-based targeting. Among the latter, the seminal study is Alatas et al. (2012). These authors separately investigate the targeting accuracy of pure community-based targeting, as in the present paper, and a hybrid method, where community-based targeting is combined with econometric targeting to identify the set of beneficiaries. While Karlan and Thuysbaert (2016) and Stoeffler et al. (2016) compare hybrid methods with selected statistical methods, our study is the first comparison of statistical targeting with pure community-based targeting in an African context.[3] Since there is great heterogeneity across the existing studies of hybrid methods regarding the combination of CBT and statistical elements, we think that the comparative study of pure community-based targeting is of particular interest because this method is subject to fewer application-specific details. Hence the insights obtained are more readily comparable across contexts.

Our main methodological contribution is that we merge the two so far disconnected branches of the comparative literature on targeting accuracy. We are first to compare the accuracy of pure community-based targeting with the four most prominent approaches to statistical targeting in one empirical setting. The second major innovation of our study is that we conduct CBTs not only in rural but also in semi-urban areas of a poor African country. CBTs have long been popular

[3] Another related article is Sabates-Wheeler et al. (2015), who compare the targeting accuracy of pure community-based targeting with two forms of categorical targeting, where the target groups are households with high fractions of elderly and dependents, respectively. The empirical context is the Hunger Safety Net Programme in Kenya.

in villages and initially been known as Rapid Rural Appraisals (Chambers, 1994a). To the best of our knowledge, there is only a single study of CBTs, for Indonesia, that also covers urban areas (Alatas et al., 2012). While other research has documented the low targeting accuracy of CBT in rural sub-Saharan African contexts (Stoeffler et al., 2016; Sabates-Wheeler et al., 2015), we find a high targeting accuracy and a superior cost-effectiveness of this method in urban neighborhoods, which parallels existing findings on CBT in Indonesia, a middle-income country (Alatas et al., 2012; Yamauchi, 2010). In addition, we identify differences in communities' poverty concepts as the source of this rural-urban difference in targeting accuracy.

More broadly, we view our results as a reconciliation of several disparate findings in recent studies of CBT's accuracy in different settings. While Stoeffler et al. (2016) and Sabates-Wheeler et al. (2015) are set in severely impoverished rural sub-Saharan African contexts with poverty rates of around 50 percent, Alatas et al. (2012) and Yamauchi (2010) primarily study rural areas of Indonesia, a middle-income country with rural poverty rates of around 20 percent. Our study area comprises rural as well as semi-urban areas with poverty rates of around 60 and 20 percent, respectively, and the rural-urban heterogeneity in our results very closely mirrors the contrasting findings obtained in the two sets of studies just cited. Hence our results support the view that community-based targeting is in general more effective in less impoverished settings because communities' poverty concepts become more congruent with researchers' and policy makers' preferred poverty concept, which is consumption-based, in the process of economic development and poverty reduction. This has important policy implications and highlights an extra challenge faced by targeted welfare and poverty reduction programs with limited budgets in severely impoverished settings, where comparatively inexpensive CBTs are especially popular. Moreover, our exploration of community-based targeting in semi-urban areas is of great practical importance since urbanization rates are rapidly rising in Western Africa, from 9 in 1950 to 45 percent in 2016 (UN-DESA, 2015), and the combating of urban poverty is becoming a pressing issue.

Our third major contribution is a detailed cost-benefit analysis, which includes various statistical methods as well as community-based targeting, an important topic on which empirical evidence has been especially thin.[4] We use comprehensive targeting cost data, consider alternative cost scenarios, and also make a methodological contribution by quantifying the trade-off between the costs and benefits of targeting, in terms of poverty reduction, with explicit formulae without relying on numerical poverty simulations (as in Skoufias et al., 2001; Alatas et al., 2012; Klasen and Lange, 2014).

The remainder of this paper is structured as follows. In the next section we review statistical and community-based targeting. The empirical setting is the subject of section 3. Section 4 introduces the empirical methodology. Section 5 contains the estimation results and section 6 the cost-benefit analysis. In section 7 we discuss our findings and conclude.

2 Statistical versus community-based poverty targeting

Statistical targeting is a relatively recent but increasingly popular targeting tool in low-income countries (Coady et al., 2004a). In Latin America, statistical targeting has been used for large-scale cash-transfer programs in Mexico (Progresa/Oportunidades), Colombia (Familias en Acción), and Chile (PASIS and SUF). National food-subsidization programs such as those in Indonesia and Egypt use statistical targeting as well (Coady et al., 2004a; Ahmed and Bouis, 2002). Statistical methods are also popular among small-scale poverty reduction programs, where often only a small set of indicators is used.

Statistical targeting typically relies on self-reported, and sometimes validated, information on a household's demographic, occupational, and asset structure to calculate for each household a score, the approximate 'means' of a household.[5] A household is targeted if its score falls short of

[4] In their recent book Del Ninno and Mills (2015, p.12) point out that "Trade-offs between the administrative costs of targeting and lower program costs are not well documented; further research is needed in this area." Karlan and Thuysbaert (2016) compare the costs, but not the cost-effectiveness, of a hybrid targeting method to the costs of two statistical targeting methods.

[5] Such information is usually preferred over self-reported income or expenditures for several reasons. First, collecting detailed income or consumption data for an entire population is very costly. Second, both measures leave more room for strategic misreporting and can be hardly verified by the enumerator. Finally, income often suffers from considerable short-term fluctuations (Alatas et al., 2012).

a pre-specified cutoff, which may be defined in absolute terms or as a population quantile. This score is calculated as a weighted average of potentially transformed proxy-means variables and, in general, involves three choices; first, the set of indicators: given the high cost of data collection for entire populations, often indicators available from existing census data are used (Ravallion, 2009); second, the transformation of each indicator into a proxy-means variable, and third the index weights. We shall discuss four eminent statistical methods, which we categorize along these lines, and contrast them with community-based targeting.

Econometric proxy means testing

This method typically uses a large set of proxy-means variables. The indicators are obtained from census data and may or may not be transformed (Filmer and Pritchett, 2001; Klasen and Lange, 2014; Alatas et al., 2012). Weights are obtained from a regression of per capita consumption on the proxy-means variables with a smaller, typically a survey data set (Filmer and Scott, 2012). Regression coefficients are then used as weights for the entire population. Hence, for a given household, the resulting score essentially equals its predicted consumption in a linear regression sense. In addition to census data, this approach requires consumption data for at least a subset of households. When a program's purpose is to reduce consumption poverty, this method is motivated by the fact that the resulting index is the best linear predictor of household consumption given the information available in a census. Most of the recent comparative targeting accuracy literature involves this statistical targeting method, and the large-scale cash transfer programs in Mexico (Progresa) and Indonesia (BLT) are two prominent applications.

Asset index

For the asset index weights are obtained from the joint distribution of the proxy-means variables themselves. Specifically, principal component analysis (PCA) is used to reduce a large set of proxy-means indicators to a small set of orthogonal linear combinations of the variables that best capture the variation in the original indicators. Following Filmer and Pritchett (2001), the first principal component is used as score and its so-called factor loadings as weights.

The PCA-based index is most frequently used to proxy a household's socio-economic status in the absence of consumption data and has been particularly popular in health-related studies that rely on data from Demographic and Health Surveys (DHS). In this literature the index has been called "wealth index" (Howe et al., 2009) or "index of socio-economic position" (Wagstaff and Watanabe, 2003). No study in the recent targeting accuracy literature considers a comprehensive asset index.[6] A large-scale application is Columbia's *Sistema de Selección de Beneficiarios para Programas Sociales* (SISBEN), which has been in effect for more than twenty years. In this system, eligibility for various social programs relies on a wealth index with 13 proxy-means variables and PCA-based weights (Castañeda, 2005).

The Poverty Scorecard Index

In comparison to the just discussed two statistical methods, scorecards typically rely on a smaller set of indicators. By means of a scorecard each indicator is transformed into an indicator score, which typically takes only integer values. The sum of indicator scores gives the score, here called wealth score. The mapping of indicator realizations into indicator scores simultaneously delivers the transformation of each indicator and the weighting between indicators. A property the scorecard approach has in common with all other statistical methods considered here is that the final score is additively separable in the individual indicators.

We consider the Poverty Scorecard Index (PSI). It was initially developed by a microlender in Bosnia-Herzegovina and primarily used to measure the microfinance institution's outreach to the poor and the institution's impact on customers' welfare. It has subsequently been managed on a global scale by Grameen Foundation and, lately, the non-governmental organization Innovations for Poverty Action (IPA), where it is called Progress out of Poverty Index (PPI). The index has

[6] Karlan and Thuysbaert (2016) consider a PCA-based index where five housing variables are aggregated into a housing index.

also been used for targeting of anti-poverty interventions and is increasingly used in contexts other than microfinance, such as health and education (Schreiner, 2015; Alkire et al., 2015). According to IPA's 2014 report, the PSI is being used by more than 200 organizations for anti-poverty programs around the global South. Among them are the Bangladesh Rural Advancement Committee (BRAC), the Grameen Bank, the Ford Foundation, and the International Finance Corporation (Innovations for Poverty Action, 2014). Customized scorecards for 46 countries are available as of 2016. The selection of indicators is based on "statistics and judgment" and, similar to the econometric PMT, indicator scores are obtained from national expenditure surveys through regression techniques with consumption poverty as the dependent variable (Schreiner, 2015, p.556). In our analysis, we use the 2011 version of the PPI scorecard for Burkina Faso, which we have retrieved from IPA's website in January 2016.

The Multidimensional Poverty Index

In this approach all indicators are first transformed into binary deprivation indicators and the relevant poverty index is a weighted deprivation count. The so-called Global MPI (Alkire et al., 2015) comprises ten indicators from three different dimensions of well-being, education, health, and standard of living. Its weights are equal across and within the three dimensions of well-being, such that the sum of all indicator weights within a dimension always equals one third. The MPI and scorecards have in common that they involve normative judgments regarding the selection and the transformation of indicators, as well as the choice of weights, while the latter is fully data driven for both the econometric proxy-means test and the asset index.

The Global MPI has been developed for the United Nations Development Programme. It is annually reported in the Human Development Report and is calculated for more than one hundred countries (Alkire and Santos, 2014). Its primary purpose is the measurement of multidimensional poverty in the developing world given common data constraints. In addition, Alkire and Santos (2010) argue that the methodology underlying the global MPI may also serve as a tool for the targeting of anti-poverty programs (see also Alkire et al., 2015). Along these lines, four recent studies compare targeting based on the MPI methodology with other targeting approaches, such as the Below the Poverty Line (BPL) scorecard for the context of India (Thomas et al., 2009; Alkire and Seth, 2013; Azevedo and Robles, 2013; Robano and Smith, 2013). We are aware that the MPI intends to capture a different, more multidimensional concept of poverty than our consumption benchmark. Due to its popularity in policy applications and its ambitions for targeting, we find it nonetheless of interest to include it in our comparative analysis to see how suited (or not) it is for proxying consumption poverty.

Community-based targeting

In community-based targeting the choice of beneficiary households is delegated to local communities (Ravallion, 1993). The approach usually involves a so-called community wealth ranking and has earlier often been called rapid rural appraisal, or RRA for short. According to Chambers (1994a), RRAs were pioneered in the late 1970's because of a growing discontent with statistical poverty assessments and, in particular, their relatively high costs. Since then, community wealth rankings have not only been used for poverty assessments (see, for instance, Devereux and Sharp, 2006; McGee, 2004; Van Campenhout, 2007) but have also emerged as a targeting tool.

Recent examples include small to medium-scale asset creation programs geared at the ultra-poor and funded by the Consultative Group to Assist the Poor (CGAP). Karlan and Thuysbaert (2016) analyze two such programs in Honduras and Peru. Banerjee et al. (2007) investigate CBTs within the context of a similar asset-creation program in rural India. Community-based targeting is sometimes also used on a larger scale. In their cross-country analysis of targeted anti-poverty interventions, Coady et al. (2004a) state that community-based targeting is similarly often used as proxy-means testing, equally popular on all continents and especially wide-spread in very poor countries.

[Table 1 about here]

To the best of our knowledge, there is no structured summary of the procedural details of community wealth rankings and community-based targeting exercises in the extant literature. Therefore, in Table 1, we review eighteen studies, inclusive of an intervention preceding the one studied in this paper (Souares et al., 2010), which are sufficiently explicit regarding procedural details. Eleven of them are situated in sub-Saharan Africa. All eighteen instances have in common that the entire community is involved, at least at an initial stage. They differ along five characteristics, which are set out in columns 1 to 5 of Table 1. First, most exercises start with a public focus group discussion to elicit wealth and poverty perceptions, and sometimes also to define wealth brackets. Second, in most of the instances summarized in Table 1, all households of the community are assigned to wealth brackets. Third, in ten of the studies, a complete wealth ranking of households is undertaken by sorting households within each wealth bracket subsequently. Fourth, the outcomes of the wealth ranking exercise are used for targeting of a welfare program in two thirds of the cases. Finally, there is variation regarding agency. In particular, the assignment of households to wealth brackets as well as the comprehensive ranking may be carried out either by the community as a whole or by a small number of elected local informants, which is the more common scenario in the studies reviewed here.

3 Empirical setting and data

3.1 The community-based targeting exercise

The empirical setting of our study is the administrative department of Nouna in the northwest of Burkina Faso, which is part of Kossi province (see Figure A1 in the Online Appendix for a map). At the time of the national census preceding our field campaign, it was inhabited by a population of about 70,000 of which 72 percent resided in 59 villages and the rest in Nouna Town, the department's only urban settlement. With this makeup, the study area is similar to the country as a whole, where 28 percent of the population have resided in urban settlements in

2014 (UN-DESA, 2015). Moreover, small urban settlements (according to the United Nations' definition) with less than 300,000 inhabitants host more than 40 percent of Burkina Faso's urban population (UN-DESA, 2015). To make clear that the urban area of our study site is not a city, we will refer to Nouna Town as 'semi-urban' in the sequel, even though it is 'urban' by the definition of Burkina Faso's census.[7]

According to the country's national statistical office, consumption poverty evaluated at the national poverty line, which is close to the World Bank's dollar a day (in 1990), stood at 47 percent in the department of Nouna in 2009, which equals precisely the national figure for that year (INSD, 2015). A continuous collection of vital statistics and later also of sample survey data by a local, government-funded health research center, the *Centre de Recherche en Santé de Nouna*, and Heidelberg University's Institute of Public Health has been ongoing in 41 of the department's 59 villages as well as in Nouna Town since 1993 (De Allegri, 2006).[8]

With the objective of developing a nation-wide voluntary health insurance scheme, the Burkinabé Ministry of Health decided to explore the potential of voluntary community-based health insurance during the early 2000s (Fink et al., 2013). The said 41 villages and Nouna Town were chosen as the pilot site because of the statistical monitoring systems already in place. The roll-out of the insurance scheme commenced in 2004 and since 2006 all households in the study area have had the opportunity to purchase community-based health insurance from a formal not-for-profit provider, the *Assurance Maladie á Base Communautaire* (AMBC), sponsored by the central government. Despite the seemingly affordable insurance premium, overall health insurance enrollment rates had remained far below expected levels. As enrollment rates were especially low among poor households (Souares et al., 2010), from 2007 onward a fifty percent discount on the premium, essentially a voucher, was offered to the poorest quintile of households in each village and urban neighborhood. To be precise, our proposal to the ethical review committee of Burkina Faso stated the intention to "identify the twenty percent poorest households (...) such that they could benefit

[7]The national definition of 'urban locality' is a "locality with 10,000 inhabitants or more and with sufficient socio-economic and administrative infrastructures" (United Nations, 2006; INSD, 2008).

[8]See Figure A2 of the Online Appendix for a detailed map.

from health insurance at lower prices."

For the identification of beneficiary households, CBT exercises were carried out in 2007, 2009 and 2011. In this paper we focus on the 2009 campaign, where CBTs took place in 36 villages of the study area as well as in 22 neighborhoods of Nouna Town.[9] In what follows, we shall refer to both villages and semi-urban neighborhoods as 'communities'. In each community, the procedure started with a publicly convened community meeting, where the facilitators first informed about the purpose of the meeting. Detailed transcripts of these meetings confirm that the official targeting objective of the insurer stated above was also communicated on the ground. The facilitators then initiated a focus group discussion to elicit criteria regarding poverty and wealth. The two most often stated criteria for characterizing poverty, "has insufficient food" and "has nothing", directly or indirectly relate to consumption (Savadogo et al., 2015).

The community was then instructed to use these criteria for defining three or four wealth brackets. In a third step, the community assembly elected three local key informants by acclamation. Physically separated from the assembly and each other, each key informant first assigned each household to one of the previously defined wealth brackets and, second, ordered all households within each bracket. While the number of households eligible in the respective community, m say, was fixed in advance by us and set equal to 0.2 times the number of ranked households, neither the community nor the key informants were informed about it before the completion of the rankings.

In a final step, the set of beneficiary households was determined according to the following rule: First, households which had been ranked among the m poorest by all three informants were automatically eligible (about 40 percent of beneficiary households; see Table 2). Second, all households which had been ranked among the m poorest by exactly two informants were included, provided that the resulting number of beneficiaries did not exceed m, and, in a consultation among the key informants, the remaining beneficiary households were chosen from the set of households which had been ranked among the m poorest by exactly one of the three informants previously.

[9] For the purpose of the targeting exercise, each of the seven administrative sectors of Nouna Town was divided into up to four neighborhoods with similar numbers of households (see Figure A1 in the Online Appendix for a map).

Otherwise, only a subset of the households which had been ranked among the m poorest by exactly two informants were selected in a consultation among the key informants. This occurred in only four of the 58 communities. In 13 communities, the number of households ranked among the m poorest by at least two informants equaled m and no consultation had to be initiated. On average, the entire exercise took half a day per community.

3.2 Data

Our CBT data set contains all three key informants' wealth rankings as well as the final eligibility status of 5,708 households. Table 2 contains summary statistics of the targeting exercise, pooled as well as separate for rural and semi-urban communities. The average community size is 106 households and 20 percent of households in each community were targeted. In rural and urban communities alike, there is a substantial positive correlation of 0.65 between the three informants' rankings as measured by the Spearman rank correlation coefficient. Nonetheless, if we define an individual key informant's target group by his m lowest-ranked households, a unanimous agreement occurs for no more than 40 percent of beneficiary households (8 percent relative to 20 percent). Finally, educational attainments of CBT informants reflect education and literacy levels among household heads in Nouna, where literacy is about twenty percent more common in the semi-urban areas (see also Table 3).

We employ two additional data sources to calculate community-level measures of heterogeneity, an economic census and a demographic census of all households in the study area, which were both fielded in the spring of 2009. As a measure of wealth inequality, we calculate a Gini coefficient from an asset index involving 25 census variables, where the weights are obtained in a principal component analysis. The Gini index for consumption in Table 2, in contrast, is calculated from a smaller sample survey data set discussed in detail in the next paragraph. According to these figures, wealth inequality is similar in rural and urban communities, while consumption inequality is larger, by about one third, in urban communities. This pattern is well in line with national figures, where consumption Ginis have stood at 0.33 and 0.46 in 2009, respectively (Odusola et al., 2017). For a measure of ethnic heterogeneity, we calculate an index commonly known as ethno-linguistic fractionalization, or ELF for short Alesina et al. (1999), which equals the probability that two randomly drawn household heads in a community belong to different ethnic groups. According to the sector-wise figures, the likelihood of encountering a household head of a different ethnicity in one's own urban neighborhood is twice as large as in the villages, which reflects the fact that the semi-urban area is much more ethnically diverse, even within neighborhoods.

[Table 2 about here]

For our main empirical analyses we match a household survey data set, the Nouna Household Survey, with the CBT data. The household survey is based on a cluster-stratified random sampling methodology, where clusters are villages and seven urban administrative sectors, and covers 655 households in all 58 communities where CBTs were carried out. Data collection took place between September and November 2009. Among the studies set out in Table 1, ours is the second largest regarding the number of communities and the only one reaching beyond villages. On the other hand, the number of sample households per community, 11.3 on average (see Table 3), is small in comparison. We will address this issue in detail below.

For the merged dataset summary statistics are set out in Table 3.[10] Targeted households are slightly overrepresented, albeit not in a statistically significant fashion, with a targeting share of 0.22 in comparison to 0.20 in the population. Households are relatively large and literacy rates low. Children and adolescents younger than 16 years of age account for two fifths of the population, while only ten percent are 55 and older. Agriculture is the predominant occupation in villages and also for half of the semi-urban households. There is a sizable morbidity risk with 21 percent of household heads reporting an illness during the month preceding the interview.

[10] While our previous explanations suggest that the Nouna department is similar to Burkina Faso regarding demographic and poverty characteristics, it should be clear that our study area as a whole is not because it does not include some of the department's villages. Therefore we focus on sector-wise analyses and alert the reader that our pooled results refer to a population in which urban dwellers are about 15 percent more frequent than in the Nouna department.

Livestock possession is wide-spread in rural and semi-urban communities alike. Bicycles are the dominant means of transportation and close to a quarter of households owns a motorbike.

[Table 3 about here]

The reference variable for our targeting accuracy analysis is monthly per capita consumption (MPC), which we measure as follows. The Nouna Household Survey contains a short questionnaire on consumption expenditures on regular as well as lumpy consumption items, which is administered to all household members aged 15 and older. It records market purchases with two recall periods, one and six months, for each of sixteen expenditure categories. We partition these sixteen categories into five high and eleven low-frequency items according to the questionnaire of the 2014 Burkina Living Standard Measurement Survey (LSMS) and calculate monthly household market purchases from the one-month and six-months recalls, respectively.[11]

Non-market consumption in the form of self-produced food items is common in our study area and, according to our own calculations with household-level data from Burkina's 2014 LSMS, accounts for 15 and 37 percent of the value of food consumption and 8 and 23 percent of total consumption among urban and rural households in Kossi province. As our survey questionnaire does not include home-produced consumption, we approximate home-produced food consumption by combining a detailed questionnaire on the household's last harvest with household demographics, which are both part of the household survey. For each household, we calculate the value of annual home-produced food consumption as the market value of all food crops in the household's last harvest (around December 2008) and reduce the resulting figure whenever the calories contained in the harvest of food crops exceed the household's annual calorie requirement. We determine these calorie thresholds endogenously through a calibration exercise, in which we match the resulting average rural and urban food consumption shares in our data with the ones in the LSMS data from Kossi province, which are 66.0 and 49.5 percent.[12]

[11] In the 2014 Burkina LSMS, the recalls for the two categories are one week and three months, respectively.

[12] Allowing for different daily thresholds for individuals younger than 16 years of age and for rural and urban inhabitants, we obtain implied daily 'autoconsumption' calorie thresholds of 1800 (1400) for adult and 900 (700) for adolescent rural (urban) inhabitants. These thresholds very likely understate actual levels of home-produced

Descriptive statistics for the resulting MPC measure are set out in Table 3. Per capita consumption is considerably higher, but also much more dispersed, in the urban neighborhoods. The consumption levels underlying the figures in Table 3 very likely understate actual consumption for two reasons. First, a short recall period of one week rather than one month as in our survey for high frequency consumption items has been found to result in consumption figures that are 25 and 13 percent higher for rural and urban households in India, respectively (Deaton and Kozel, 2005). Second, in several regional contexts, short and unspecific questionnaires like the one in our survey have been shown to understate consumption by around 30 percent relative to a detailed questionnaire, which is used, for example, in Living Standard Measurement Surveys (Beegle et al., 2012; Jolliffe, 2001). Taking these insights to our data suggests that the consumption levels in Table 3 may understate actual consumption by 40 to 50 percent in urban and rural areas, respectively. An unpublished study by World Bank staff in Indonesia (World Bank, 1993) discussed in Beegle et al. (2012) has found, however, that such downward biases in survey consumption tend to apply similarly to all households and largely preserve consumption ranks, which is essential for our subsequent targeting accuracy analysis.[13]

4 Empirical methodology

4.1 Sample target sets

In our main analysis we construct and compare five different sets of target households to a reference set of consumption-poorest households (T^{CON}). First, the set of households actually targeted by the communities. We denote the corresponding set of targeted households in our sample of 653

consumption by 30 to 40 percent because of understated market consumption in our survey (see the next paragraph) and our calibration strategy, which matches food expenditure *shares* with those in the LSMS. While we cannot assess how well our measure of MPC approximates actual MPC in our data, applying the same methodology to the LSMS data from Kossi province, which also includes a harvest module, gives very high rank correlations of actual survey MPC (survey market consumption plus survey autoconsumption) and the resulting proxy measure of MPC (survey market consumption plus imputed autoconsumption), 0.94 and 0.99 in rural and urban primary sampling units, respectively.

[13] We also conduct a robustness check with an alternative consumption measure including only monthly per capita market expenditures in our consumption survey. All main results remain qualitatively unchanged. See table A1 in the Online Appendix for details.

households in total by T^{CBT}. The remaining four statistical targeting sets are constructed from the household survey data and denoted by $T^{ECON}, T^{PCA}, T^{PSI}$ and T^{MPI}, respectively. Table 4 provides an overview of the four statistical targeting methods regarding indicators and weights and section 1 of the Online Appendix discusses in detail how we construct the respective statistical targeting sets for our sample.[14]

[Table 4 about here]

4.2 Targeting accuracy

We assess the targeting accuracy of each targeting method, m say, in terms of the overlap of T^m with T^{CON} by means of the mean targeting error, which is defined as the proportion of households erroneously classified as either poor or non-poor (Alatas et al., 2012).[15] Denoting by n the total number of sample households, the mean targeting error (MTE) for method m is calculated as

$$
\begin{aligned}
MTE_m &= \frac{1}{n}\sum_{i=1}^{n}\left[\begin{array}{c} \mathbb{1}\{\text{ household } i \text{ is in } T^{CON} \text{ and not in } T^m\}+ \\ \mathbb{1}\{\text{ household } i \text{ is in } T^m \text{ and not in } T^{CON}\ \} \end{array}\right] \qquad (1)\\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\ \mathbb{1}\{\text{wrongly excluded}\}_i\ \right] + \frac{1}{n}\sum_{i=1}^{n}\left[\ \mathbb{1}\{\text{wrongly included}\}_i\ \right], \\
& m = \{ECON, PCA, PSI, MPI, CBT\},
\end{aligned}
$$

where $\mathbb{1}\{\}$ denotes the indicator function. The second line of equation (1) illustrates that the the mean targeting error is the sum of two types of errors. First, an exclusion error (false negative) occurs when any of the consumption-poorest households is not targeted by the targeting

[14] An obvious limitation of our research design is that the econometric PMT, as well as the other three statistical methods that we consider, is hypothetical. To the best of our knowledge, Alatas et al. (2012) is the only study where both methods, community-based targeting and econometric proxy-means testing, have been implemented in the field simultaneously. Hence our results for the statistical methods should be taken with a grain of salt, reflecting an ideal case where implementers are honest, non-corrupt technocrats.

[15] The Targeting Differential, or TD for short (Galasso and Ravallion, 2005), and the Coady-Grosh-Hoddinott (CGH) index (Coady et al., 2004a) are two additional popular targeting accuracy measures. When benefits are identical across beneficiary household, both measures can be expressed as monotone transformations of the mean targeting error (MTE) within our framework because the share of beneficiary households equals the share of consumption-poorest households. First, the Targeting Differential is the difference between the share of the poor and the non-poor participating in the program, $TD = 1 - \frac{MTE}{q} \in [-1, 1]$, where q is the share of targeted households. Second, the Coady-Grosh-Hoddinott index is the amount of resources transferred to the poor over the total amount transferred by the program, $CGH = \frac{2q-MTE}{2q^2} \in [0, 1]$.

method under consideration. Conversely, non-poor households which are targeted by the method under consideration contribute to an inclusion error (false positive). As a benchmark for comparison, we also calculate the mean targeting error when households are targeted at random. For the sample targeting share of 22 percent, the probability of erroneous targeting under random targeting is $0.78 \cdot 0.22 + 0.22 \cdot 0.78 = 0.34$, which is calculated as the random exclusion error times the probability of being among the consumption-poorest plus the random inclusion error times the probability of not being among the consumption-poorest.

When we compare two alternative targeting procedures, A and B, the object of interest is the difference in the mean targeting error. To conduct statistical inference, we estimate the regression equation

$$Err_{im} = \gamma + \delta \cdot \mathbb{1}\{m = B\} + u_{im},$$

where the dichotomous variable Err_{im} is the targeting error of observation i with procedure m, the term in brackets in the first line of equation 1, and u is a stochastic error term. Procedure A defines the reference category and the least squares estimate of δ equals the difference between the mean targeting errors of procedures B and A. The data set for this exercise has $2n$ observations as every household appears twice, once with procedure A and once with procedure B. We cluster standard errors at the household level.

5 Results

5.1 Targeting accuracy

Table 5 reports mean targeting errors as well as exclusion and inclusion errors for all households, pooled as well as separately for rural and urban communities. Mean targeting errors range from 23.5 to 32.4 percentage points in the rural and from 22.9 to 32.6 in the semi-urban sector (columns 4 and 7), which amounts to a reduction of the random MTE between 35 and two percent. The econometric PMT has by far the lowest MTE with just 23.5 and 22.9 percent of households wrongly classified. The difference to the next-best method, the asset index, is 3.7 percentage points in the pooled data (column 1), which is statistically significant at conventional levels. On the other hand, the Poverty Scorecard Index and the MPI deliver only marginal and statistically insignificant (at the five percent level) improvements relative to random targeting. All statistical methods perform similarly across sectors and the asset index is the best-performing statistical method not involving consumption data for calibration. The differences in MTEs between the asset index, the PSI and the MPI are not statistically significant at the five percent level, however, while the econometric PMT performs significantly better (see Appendix Table A8 for significance levels).

Regarding community-based targeting, there are two main findings emerging from Table 5. First, averaged across rural and urban communities (columns 1 to 3), CBT performs about as accurate as the three statistical methods that are not calibrated with consumption data and reduces the random targeting error by 14 percent, which is a statistically significant improvement at the one percent level. The more remarkable finding is the large difference in targeting performance between rural and semi-urban communities. In the villages, CBT performs poorly and does not deliver a tangible improvement over random targeting. This finding confirms existing evidence on CBT's low targeting accuracy in other rural contexts in sub-Saharan Africa (Stoeffler et al., 2016, for Cameroon; Sabates-Wheeler et al., 2015, for Kenya) and Latin America (Karlan and Thysbaert, 2013, for Honduras and Peru). In contrast, in our setting, CBT achieves a high accuracy in the semi-urban areas (columns 7 to 9). As the second-best performing method, it reduces the random targeting error by 33 percent (significant at one percent), from 35.4 to 23.8 percentage points. This order of magnitude is similar to the one reported by Alatas et al. (2012) for urban areas in Indonesia, for which their results imply an improvement of 24 percent over random targeting. The improvement is greater than the 12 percent reported by the same authors or the 17.5 percent improvement in Yamauchi (2010) for Indonesian villages.

Regarding CBT and PMT in comparison, the rural-urban pattern in our data closely parallels the findings of Alatas et al. (2012), where the difference in error rates between an econometric

PMT and CBT equals seven percentage points in villages but only two percentage points in urban neighborhoods. This double difference of five percentage points is not statistically significant in their estimations, however.

[Table 5 about here]

Table 5 also contains exclusion and inclusion errors, and the corresponding random targeting errors as reference. By construction, the number of erroneously included households always equals the number of erroneously excluded households. Accordingly, the values in columns 5 and 8 equal the mean targeting error divided by two times the corresponding sample targeting share set out in Table 3. The mean inclusion errors are a multiple of the respective exclusion errors, where the factor of proportionality is the sample targeting share, s say, divided by one minus s. We will return to the exclusion errors in our cost-benefit analysis.

In Table 6 we decompose targeting errors along the consumption distribution. In particular, we calculate exclusion errors separately for extremely poor and moderately poor households, and inclusion errors for households around the distribution's median as well as for relatively affluent households. We define the expenditure classes such that the shares of extremely and moderately poor households are roughly equal and sum up to the sample targeting shares of the community-based targeting exercises, which are set out in Table 3. The other two expenditure brackets contain the complementary sets of households and are defined such that they are roughly of equal size; for example in the pooled sample with a sample CBT targeting share of 22 percent, the affluent and around median expenditure brackets roughly contain the first and second 39 percent wealthiest households as measured by consumption, respectively. As a consequence, the mean exclusion and inclusion errors in Table 5 are the arithmetic means of the respective consumption-bracket-wise errors in Table 6.[16]

[16] The numbers of households in columns 1 and 2, 3 and 4, 5 and 6, and 7 and 8, respectively, are not equal because of communities where the CBT sample target set or its complement contains an odd number of households. In that case, we have chosen to allocate the median household of the consumption sample target set to the moderately poor group and the median household of the complementary group to the around-median group. Changing this rule does not affect the results substantively.

[Table 6 about here]

Only the Econometric PMT performs consistently well across sectors and expenditure classes. In the semi-urban neighborhoods, CBT and the asset index share the favorable feature of identifying correctly households in the lowest and highest expenditure brackets, where they improve on random targeting by around 60 percent. In contrast, and unlike the econometric PMT, neither of those methods provides a sizable accuracy improvement for households around the poverty threshold.

5.2 Community characteristics and targeting accuracy

The main innovation of the field work underlying our study are the CBTs in semi-urban areas alongside the rural ones. We have found a considerably better targeting performance in the urban neighborhoods compared to the villages, while the econometric PMT performs similarly well in both settings. In this and the next section we aim to identify the sources of the difference in CBT's targeting performance across rural and semi-urban sectors. We explore three possibilities in turn: first, differences in observable community and key informants' characteristics, such as community-level economic inequality or key informants' education; second, differences in revealed poverty concepts, which we identify through the implicit weights rural and urban CBT key informants put on various observable household characteristics when identifying beneficiaries. Finally, we test for the possibility that rural key informants' decisions are more idiosyncratic than those of their urban counterparts. Toward this, we decompose CBT outcomes into a predictable and an unpredictable component and assess whether the latter is larger in the rural sector.

In this section, our focus is on whether observable community characteristics can explain the difference in CBT's performance between rural and urban communities. For the empirical implementation, we regress the targeting error of an observation, Err_{cim}, on characteristics of the community where the respective household resides, X_c. We cluster standard errors at the level of a community. Probit estimation results are set out in Table 7.

[Table 7 about here]

Consistent with our core results set out in Table 5, there is no statistically significant difference in targeting errors between rural and urban areas for any of the four statistical methods that we have considered (first four columns of Table 7). On the other hand, for CBT, the difference is large, close to nine percentage points, and statistically highly significant (column 5). Column 6 shows that this difference is not driven by observable community characteristics: the point estimates of all explanatory variables, community size as well as measures of economic inequality and ethnic diversity, are very small and statistically insignificant, while the rural dummy even increases slightly in magnitude with the inclusion of these control variables. In column 8, we add two summary measures of CBT key informants, their education level and degree of agreement - with no effect on the rural-urban difference in the first row. Since information on council members is available for only 25 of the 36 villages, column 7 replicates column 6 with the restricted sample of column 8 - with no mentionable effect on any of the coefficients of interest. Qualitatively, these findings are in accordance with Alatas et al. (2012), who also find no significant community-level predictors of CBT errors. On the other hand, in a study set in rural Indonesia, Yamauchi (2010) finds a slightly greater targeting accuracy of CBTs in smaller communities and ones with greater consumption inequality.

5.3 Communities' poverty concepts

Our results in the previous section suggest that there are no community-level predictors of CBT's accuracy. As a second possible source of the difference in CBT's accuracy between rural and urban areas we examine differences in revealed poverty concepts across the two sectors. Suppose for example that urban communities have a more consumption-oriented poverty concept, while rural communities put more weight on components of multidimensional poverty, such as health, education or asset possession, which may be little correlated with survey consumption.

We elicit key informants' revealed poverty preferences by regressing a household's beneficiary status on observable household characteristics which key informants may take into account when identifying beneficiaries. As predictors, we consider the educational attainment and the current health status of the household head as well as demographic characteristics of the household. For assets, to save degrees of freedom, we use two principal-component-weighted summary measures of twelve consumption assets and nine productive assets from agriculture, which is the modal occupational activity in both sectors. We also include an indicator for whether the household head belongs to an ethnic minority in his or her community to test for the possibility of favoritism along ethnic lines. We carry out probit estimations separately for rural and urban households, whose marginal effects are set out in columns 1, 2, 4 and 5 of Table 8. We also estimate a model that nests the rural and urban estimation equations to test for differences in the rural and urban coefficients (columns 3 and 6 of Table 8).

[Table 8 about here]

The first two columns of Table 8 confirm our core finding regarding CBT's accuracy in rural and urban communities: decreasing consumption by ten percentage points increases the probability of being targeted by about five percent (or one percentage point) in urban neighborhoods, while the same probability decreases, albeit insignificantly, by a little more than one percent for rural households.

Including additional household characteristics (columns 4 and 5) suggests that urban and rural communities put similar weights on asset possession: a simultaneous reduction of assets in both asset categories by one standard deviation increases the likelihood of being targeted by 60 percent (or twelve percentage points) in rural and urban communities alike. Similarly, illiteracy and female headship are important predictors of beneficiary status in both sectors: each of them increases the likelihood of being targeted by one third or more. We also find some evidence for ethnic favoritism in both sectors as minority households are less likely to be beneficiaries.

Turning to differences across rural and urban sectors, unlike their urban counterparts, rural key informants put a large weight on household demographic characteristics and health. According to the point estimates in column 5, a rural household in which half of the members are 55 and older has an eighty percent (or 17 percentage points) higher chance of being targeted than a household with no elderly members. Moreover, rural household heads who reported an illness during the month preceding the interview are one third more likely to be beneficiaries - conditional on all other characteristics. In contrast, the corresponding effects are small and insignificant in the urban sample and, according to column 6, also significantly different from the rural ones (at the ten percent level).

Overall the results set out in Table 8 imply that several observable household characteristics, such as assets, literacy and female headship, play a stronger role in community targeting decisions than survey consumption. Even in our urban sub-sample, the consumption coefficient almost vanishes when other characteristics are conditioned on (first entry in column 4). Importantly, however, all predictors of CBT eligibility that are significant in the semi-urban areas according to column 4 are negatively conditionally correlated with per capita consumption: in a regression of logarithmic MPC with community fixed effects, both asset variables, literacy and male headship have positive signs, and the consumption asset index is the most important predictor of MPC. Somewhat analogous, in the rural sub-sample agricultural assets are the most important predictors of MPC and literacy as well as male headship have positive, albeit very small, effects on consumption. In contrast, the other two significant predictors of CBT eligibility in villages, the share of elderly household members and illness, are strong positive predictors of consumption, which explains, at least in part, the positive coefficients in the first row of columns 2 and 5.

We finally turn to the possibility that rural key informants' decisions are more idiosyncratic than those of their urban counterparts in the sense that they depend on factors unobserved by the researcher to a greater extent. Toward this, we construct rural and urban targeting sets from the predicted CBT eligibility scores obtained from the regressions whose results are set out in columns 4 and 5 of Table 8. Essentially these are hypothetical beneficiary sets resulting from weighted averages of household characteristics where the weights are calibrated by CBT outcomes. Parallel to our previous analysis, we quantify the idiosyncratic component of CBT outcomes by the MTE between actual and predicted CBT eligibility. For the 22 urban and 36 rural communities, we obtain MTEs of 17.9 and 19.0 percent, respectively, with standard errors of 2.5 and 1.9 percent. We hence conclude that rural and urban targeting outcomes incorporate observable household characteristics to very similar extents and that rural key informants are no more idiosyncratic than their urban counterparts.

The pattern that assets and other household characteristics are more important predictors of CBT than consumption parallels Alatas et al. (2012), who also find no or only a small effect of consumption on CBT eligibility when other observable factors are conditioned on. Moreover, a negative consumption coefficient in multiple regressions of community targeting outcomes (as in column 5) has also been found in studies set in rural Kenya and Cameroon (Sabates-Wheeler et al., 2015; Stoeffler et al., 2016). The main novelty of our results is that the disappointing targeting performance of CBTs in sub-Saharan African contexts does not appear to be due to idiosyncrasies in communities' decisions. Instead, poverty concepts that are, at least partially, at odds with conventional consumption poverty seem to drive CBT's poor targeting accuracy. Moreover, in the less impoverished urban sector of our study area, communities' poverty concepts are more congruent with consumption poverty, perhaps because of greater market development and competition, and CBT's targeting precision is qualitatively more similar to the accuracy rates reported in Alatas et al. (2012) and Yamauchi (2010), which are both set in a middle-income country.

6 Cost-benefit analysis

Given the superior targeting accuracy of the econometric proxy-means test and partly also the asset index over community-based targeting we compare costs and benefits of these three methods. Mayoux and Chambers (2005, p.283) state that "a key advantage of participatory methods is their cost-benefit in rapidly bringing together information and knowledge from many participants." In the same vein, the meta-studies of Coady et al. (2004b, p.61) and Conning and Kevane (2002) attribute the lower administration costs of CBT to the wage differential between external enumerators and community agents. When a welfare program's intention is to reduce poverty and CBT is cheaper but at the same time less accurate than statistical targeting, there is a trade-off and the relatively inexpensive CBT will be more cost-effective than statistical targeting for programs with relatively small budgets, while the opposite holds for relatively generous transfer programs. This is precisely what Alatas et al. (2012) find in their Indonesian context (Table 5, columns 1 and 2, of their Online Appendix).

In this section we will calculate cost thresholds for the competing targeting methods. We make two innovations. First, on the cost side, we consider alternative scenarios regarding the availability of data for statistical targeting. Second, on the benefit side, we derive explicit formulae linking exclusion errors from the estimations to poverty reduction instead of relying on numerical poverty simulations (as in Ravallion, 2009; Alatas et al., 2012; Klasen and Lange, 2015).

We use cost information from the 2009 community-based targeting intervention and implementation costs for the statistical methods based on data collection campaigns in 2010 (Lietz et al., 2015).[17] All figures are inflated to 2014 CFA (African Financial Community Francs) using the consumer price index of Burkina Faso and converted to 2014 US dollars using the 2014 average market exchange rate of 526 Francs per dollar.

Total implementation costs of our CBTs amount to $2,373. For the two statistical methods we consider three cost scenarios. First, we assume that census and household survey information are freely available and only data processing costs of $5,761 for the econometric PMT and $2,665 for the asset index accrue. The difference between the two amounts reflects the extra work required to process the consumption survey data for the econometric PMT. In addition, our second scenario takes into account the data collection costs for the household consumption survey of $41,899, which is needed to calibrate the econometric PMT. Hence we calculate a total cost of $47,660

[17] We focus on direct implementation cost and do consider neither opportunity cost of the survey respondents nor the communities' focus group participants.

for the econometric PMT, while the cost of the asset index remains unchanged. In the third scenario, for both statistical methods, we add the cost of collecting the census data of $36,053, amounting to total costs of $83,713 and $38,718 for the econometric PMT and the asset index, respectively. To these fixed targeting costs we add the hypothetical aggregate benefits paid to beneficiary households as variable costs, to obtain the total cost of a hypothetical targeted welfare program.[18]

Turning to the benefits of targeting, the most frequent approach in the existing literature on targeting accuracy has been to consider alternative sets of beneficiary households, one for each targeting method, and calculate, typically in numerical simulations, hypothetical changes in a specific poverty measure resulting from increasing each beneficiary household's per capita income (or consumption) by alternative transfer amounts (Ravallion, 2009; Alatas et al., 2012; Stoeffler et al., 2016). Adopting this approach here, which involves a poverty line that is fixed across communities, is little meaningful because beneficiary quotas in our community targeting exercise are not synchronized with community poverty rates or, in other words, our research design does not include a preceding geographical targeting step. Instead we define the social benefit of a hypothetical transfer program by two criteria. First, to identify 'socially valuable' beneficiaries, we rely on the stated purpose of the community targeting exercise, to identify the 20 percent poorest households in each community. Second, for the amount of the social benefit, we choose a metric that has the same unit of measurement as the program cost, the aggregate transfers received by 'socially valuable' households.[19] This is equivalent to considering the change in a welfare function where the consumption of each household which belongs to the poorest quintile of a community pre-intervention enters with a weight of one and all other households with a weight of zero. Moreover, under the assumption that the poverty line equals MPC at the percentile corresponding to the beneficiary quota, we show in Appendix 2 that this definition of the social

[18] Given that the area in which the data collection for the statistical methods has been carried out is relatively small and the fact that statistical methods typically exhibit considerable economies of scale, our cost figures for the statistical methods may be viewed as upper bounds to those incurred by a nation-wide program.

[19] If 'socially valuable' is replaced by 'consumption-poor', our social benefit measure is equal to the Distributional Characteristic (Coady and Skoufias, 2004), in which the aggregate transfers received by initially consumption-poor households are the social benefit of a targeted program.

benefit is equivalent to taking a multiple of the poverty gap index as the social welfare function, at least when the benefits per beneficiary household are small.

This insight establishes the connection between our approach and the one taken, e.g., in Alatas et al. (2012), where an elaborate geographical targeting step precedes within-community targeting such that community-wise beneficiary quotas are close to community poverty rates and hence the poverty line close to MPC at the percentile corresponding to each community's beneficiary quota. Hence our measure of the benefits of targeting are comparable in a straightforward fashion to those concerning the poverty gap index in the studies cited above.[20]

To avoid numerical poverty simulations to quantify the effects of targeting on poverty, we link the targeting accuracy of a specific targeting method to the social benefit of a welfare program through the exclusion error. We think that this explicit relationship instead of less transparent simulations is a substantial advantage for the understanding of the link between targeting accuracy and poverty reduction. To fix ideas, we denote by B the average benefit per consumption-poor household in response to a targeted welfare program which transfers t dollars to each eligible household and relies on a specific targeting method. Since, in our setup, the probability that a consumption-poor household is a beneficiary equals one minus the exclusion error we have that

$$B = (1 - E)t,$$

where E denotes the exclusion error of the targeting method under consideration. The cost of such a program per eligible household will be denoted by

$$C = t + TC,$$

[20] A referee of this paper pointed out that our approach to the social benefits of targeting implies that transfers received by several households under the national consumption poverty line are not classified as 'socially valuable', especially in the rural part of our study area where the poverty rate stands at around 60 percent. While we share this concern, we think that the greatest general insights valid beyond the specific parameters and (lack of) preceding geographical targeting in our intervention are obtained from defining each method's benefits narrowly according to the ex-ante stated targeting objective.

where TC denotes the fixed targeting costs per eligible household, i.e. the total targeting costs divided by the number of beneficiary households. Consolidating the two equations, we obtain

$$B(C; E, TC) = (1 - E)(C - TC). \tag{1}$$

As pointed out above, this benefit is approximately proportional to the reduction in the community's poverty-gap index for a poverty line equal to the community's MPC at the threshold between the lowest and second-lowest quintile, where the factor of proportionality is independent of C, E and TC. Following Ravallion (2009), we are interested in which targeting method delivers the greatest benefit given a budget for the total cost per eligible household, C.

We now compare the benefits of a transfer program involving community-based targeting with the econometric PMT and the asset index, respectively. The econometric PMT is always the most cost-effective choice for programs with a sufficiently large budget because, as C tends to infinity, the limit of the benefit-to-cost ratio approaches one minus the targeting method's exclusion error. For small costs, in contrast, it is solely the fixed targeting cost TC that matters for cost-effectiveness. For all three cost scenarios regarding statistical targeting, CBT always accrues considerably less than the targeting cost of either of the two statistical methods, implying that it is the most cost-effective method for targeted anti-poverty programs with a small budget. Table 9 contains transfer-amount thresholds for pairwise comparisons of the three targeting methods. When only data processing costs accrue (column 1), community-based targeting is more cost-effective than econometric targeting for program costs of up to 0.78 and 41.70 dollars per eligible household in rural and semi-urban communities, respectively. When all data collection costs are taken into account for the two statistical methods, these figures increase to 45.91 and 1,028.40 dollars, respectively. While the asset index is always more cost effective than CBT in rural areas in a data-processing-only scenario, the opposite is true for the semi-urban sector, where CBT is always cheaper and more accurate than the asset index.

[Table 9 about here]

It is also interesting to compare the two statistical procedures with each other. Recall that employing principal components does not require the use of consumption data. Accordingly, in cost scenarios 2 and 3, the econometric PMT is more cost-effective only for relatively large program budgets, of about 60 and 260 dollars in rural and semi-urban areas, respectively. For the two sectors taken together we find little scope for the relatively less expensive statistical methods, which employ no consumption data for the calibration of index weights, as the upper envelope of the cost-benefit frontier is largely formed by community-based and econometric targeting.

To put these figures into perspective, the effective average benefit per eligible household in our intervention, a discount on the premium of a health-insurance policy valid for 24 months, amounts to $1.28, which implies a total cost of $2.47 per eligible household (benefit of $1.28 plus targeting cost of $1.19). We conclude that, among the targeting procedures considered here, CBT was indeed the most cost-effective method for targeting consumption-poor households - even though CBT's targeting cost of $1.19 per eligible (or consumption-poor) household equals about twice the average transfer benefit received by a consumption-poor household ($0.64, one minus CBT's average exclusion error of 0.5 times $1.28).

Given this seeming disproportion, would an untargeted subsidy have been more cost-effective? As there are no targeting costs for such a program and, by construction, twenty percent of households in each community are consumption-poor in our application, the benefits received by poor households equal one fifth of the total costs of the program. This implies that a universal program is always most cost-effective for very small program budgets - because no fixed costs accrue. For a uniform transfer, a universal program is more cost effective than CBT up to cost thresholds of $2.18 and $1.89 in the rural and semi-urban sector, respectively. Given the total cost per beneficiary household of $2.47, it appears that community-based targeting has indeed been the most cost-effective method for the small-scale intervention considered here.

7 Discussion

While targeting accuracy assessments of specific welfare programs are numerous, there is only a small number of studies comparing alternative targeting methods within the same setting (Grosh and Baker, 1995; Alatas et al., 2012; Filmer and Scott, 2012; Klasen and Lange, 2014; Sabates-Wheeler et al., 2015; Brown et al., 2018; Karlan and Thuysbaert, 2016; Stoeffler et al., 2016). Evidence is even scarcer when it comes to comparisons between statistical and decentralized targeting methods (Alatas et al., 2012; Sabates-Wheeler et al., 2015; Karlan and Thuysbaert, 2016; Stoeffler et al., 2016). In order to fill this gap we have compared a community-based targeting intervention in Burkina Faso with various common statistical targeting methods, which we have calculated from household survey data.

In the following we shall summarize our findings and make explicit how they contribute to the existing literature. First, regarding the performance of various statistical targeting methods, we confirm the common and little surprising finding that the econometric PMT is by far the most accurate method. Regarding other statistical methods, our findings are partially in accordance with Filmer and Scott (2012), who find no statistical differences when comparing the asset index with other common statistical indices that do not involve consumption data for calibration. Second, regarding CBT and econometric targeting, our targeting accuracy results are similar to those obtained in large field experiments in Indonesia, a middle-income country (Alatas et al., 2012; Yamauchi, 2010). Third, our finding of CBT's remarkable performance in semi-urban areas is novel. CBT initially emerged from so-called rapid rural appraisals and has so far predominantly been applied in rural settings (Chambers, 1994a). Coady et al. (2004a) expect the method to perform worse in urban areas, where anonymity is greater and hence the information advantage of local community members smaller. Our findings suggest that communities in rural and semi-urban areas hold systematically different poverty concepts and that survey consumption is more strongly correlated with the latter ones'.

Finally, findings from our cost-benefit analysis demonstrate the trade-off between CBT's lower program costs on the one hand and the econometric PMT's higher accuracy on the other. Even if there is much anecdotal evidence for CBT's relative cost advantage over statistical targeting methods, there are very few studies including cost data (Alatas et al., 2012; Karlan and Thuysbaert, 2016). In our context, where we consider an inexpensive decentralized expert assessment, community-based targeting is more cost-effective than any of the statistical methods. The accuracy gains of the econometric PMT outweigh CBT's cost advantage only for very large program budgets and there is little scope for less expensive statistical methods, such as the asset index or scorecards. For the budget available for our intervention, participatory targeting is clearly the method of choice. But even for more generous programs encountered in practice, CBT may dominate statistical targeting in this African context. To illustrate, first, we consider the Indonesian unconditional cash transfer program investigated by Alatas et al. (2012). Per individual in an eligible household, the program's annual cash transfer equals about 1.4 percent of Indonesia's per capita GDP in 2008. Translated to Burkina, this figure amounts to about \$70 per household and year in 2014 US dollars. If households from rural areas are to be targeted, a program of this size exceeds the threshold of \$46 below which community-based targeting is the most cost-effective method. On the other hand, if benefits are granted to households from both rural and urban communities CBT will be more cost-effective. A Burkinabé cash transfer program geared at improving schooling and access to health care for children in poor families, the Nahouri Cash Transfers Pilot Project, employed sophisticated econometric proxy-means testing (Akresh et al., 2014). Implemented between 2008 and 2010, it involved transfers of about \$160 per targeted household over the course of two years. Again, econometric targeting would be the method of choice for rural but nor for urban households.

We shall close this paper with three remarks concerning decentralization of targeting that reach beyond the somewhat narrow domain of targeting accuracy. First, our cost-benefit analysis shows that econometric targeting is very costly when census data is not readily available. Given that a general census is typically not carried out more often than every ten years, targeting based on census data will become less accurate the more outdated the underlying data. Community targeting exercises, on the other hand, may be repeated on a revolving basis at a moderate cost and in this way keep track of poverty transitions of households over time. This argument further suggests that revolving community-based targeting might be particularly suited for quickly-evolving environments. In our study area, for example, the community-based targeting exercise has been carried out three times between 2007 and 2011.

Second, the participative procedure of community-based targeting likely produces additional benefits. Since the inception of participatory appraisals, local control over the targeting process has been viewed as a desirable attribute of CBT, powerful enough to increase ownership and awareness, and foster institutional change (Chambers, 1994b). This view is supported by empirical evidence, which shows remarkably high approval rates by communities for decentralized targeting methods (Alatas et al., 2012; Robertson et al., 2014; Schüring, 2014). Savadogo (2017) confirms this for our community-based targeting intervention, where more than 85 percent of households approve of the targeting method.

Finally, it may be called into question whether consumption should be the sole targeting objective. Instead, there might be considerable value added to the targeting process when communities' concepts of poverty are taken into account. Recent empirical evidence on communities' poverty perceptions shows that communities consider more dimensions than only consumption (Alatas et al., 2012) and that their poverty concept is multidimensional (Van Campenhout, 2007). Furthermore, Kebede (2009) shows that poverty perceptions reflect local circumstances and Alderman (2002) finds that community assessments put more weight on chronic poverty. Considering the wealth criteria defined by the communities in our targeting exercise, it is striking that especially rural communities define most of the criteria in terms of capabilities such as "has insufficient food", "has nothing" or "is not able to solve problems by himself" (Savadogo et al., 2015). This fits well into Amartya Sen's capability approach (Sen, 1988) and supports the view that communities consider consumption as just one of several means. In this perspective, community-based targeting appears to be well-suited for translating deprivations in the space of capabilities into targeting outcomes.

References

Adams, A. M., T. G. Evans, R. Mohammed, and J. Farnsworth (1997). Socioeconomic stratification by wealth ranking: Is it valid? *World Development 25*(7), 1165–1172.

Ahmed, A. U. and H. E. Bouis (2002). Weighing what's practical: Proxy means tests for targeting food subsidies in Egypt. *Food Policy 27*(5), 519–540.

Akresh, R., D. De Walque, and H. Kazianga (2014). Alternative cash transfer delivery mechanisms: Impacts on routine preventative health clinic visits in Burkina Faso. In *African Successes, Volume II: Human Capital*, NBER Chapters, pp. 113–135. National Bureau of Economic Research, Inc.

Alatas, V., A. Banerjee, and R. Hanna (2012). Targeting the poor: Evidence from a field experiment in Indonesia. *American Economic Review 102*(4), 1206–1240.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, R. Purnamasari, and M. Wai-Poi (2013a). Does elite capture matter? Local elites and targeted welfare programs in Indonesia. Working Paper Series 13-008, Harvard University, John F. Kennedy School of Government.

Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, R. Purnamasari, and M. Wai-Poi (2013b). Ordeal mechanisms in targeting: Theory and evidence from a field experiment in Indonesia. Working Paper 19127, National Bureau of Economic Research.

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.

Alesina, A., R. Baqir, and W. Easterly (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics 114*(4), 1243–1284.

Alkire, S., J. Foster, S. Seth, M. E. Santos, J. M. Roche, and P. Ballon (2015). *Multidimensional poverty measurement and analysis.* New York, USA: Oxford University Press.

Alkire, S. and M. E. Santos (2010). Acute multidimensional poverty: A new index for developing countries. Human development report office background paper 2010/11, United Nations Development Programme.

Alkire, S. and M. E. Santos (2014). Measuring acute poverty in the developing world: Robustness and scope of the multidimensional poverty index. *World Development 59*(7), 251–274.

Alkire, S. and S. Seth (2013). Selecting a targeting method to identify BPL households in India. *Social Indicators Research 112*(2), 417–446.

Azevedo, V. and M. Robles (2013). Multidimensional targeting: Identifying beneficiaries of conditional cash transfer programs. *Social Indicators Research 112*(2), 447–475.

Banerjee, A., E. Duflo, R. Chattopadhyay, and J. Shapiro (2007). Targeting efficiency: How well can we identify the poor? Working Paper 3, Centre for Micro Finance.

Banerjee, A., M. R. Hanna, and J. Kyle (2014). Information is power: Identification cards and food subsidy programs in Indonesia. NBER Working Paper 20923, National Bureau of Economic Research.

Bardhan, P. and D. Mookherjee (2006). Pro-poor targeting and accountability of local governments in West Bengal. *Journal of Development Economics 79*(2), 303–327.

Beegle, K., J. De Weerdt, J. Friedman, J. Gibson, et al. (2012). Methods of household consumption measurement through surveys: Experimental results from Tanzania. *Journal of Development Economics 98*(1), 3–18.

Brown, C., M. Ravallion, and D. Van De Walle (2018). A poor means test? Econometric targeting in Africa. *Journal of Development Economics 134*, 109–124.

Caizhen, L. (2010). Who is poor in China? A comparison of alternative approaches to poverty assessment in rural Yunnan. *The Journal of Peasant Studies 37*(2), 407–428.

Castañeda, T. (2005). Targeting social spending to the poor with proxy-means testing: Colombia's SISBEN system. Social Protection Discussion Paper Series 529, The World Bank, Washington, D.C.

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

Chambers, R. (1994b). The origins and practice of participatory rural appraisal. *World Development 22*(7), 953–969.

Coady, D., M. Grosh, and J. Hoddinott (2004a). Targeting outcomes redux. *The World Bank Research Observer 19*(1), 61–85.

Coady, D., M. E. Grosh, and J. Hoddinott (2004b). *Targeting of transfers in developing countries: Review of lessons and experience.* Washington, DC: The World Bank.

Coady, D. and E. Skoufias (2004). On the targeting and redistributive efficiencies of alternative transfer instruments. *Review of Income and Wealth 50*(1), 11–27.

Conning, J. and M. Kevane (2002). Community-based targeting mechanisms for social safety nets: A critical review. *World Development 30*(3), 375–394.

De Allegri, M. (2006). *Why do people choose to enrol or not to enrol in community health insurance? The case of the Nouna Health District, Burkina Faso.* Ph. D. thesis, Ruprecht-Karls-Universität Heidelberg, Heidelberg.

Deaton, A. and V. Kozel (2005). Data and dogma: the great Indian poverty debate. *The World Bank Research Observer 20*(2), 177–199.

Deaton, A. and S. Zaidi (2002). *Guidelines for constructing consumption aggregates for welfare analysis*, Volume 135. Washington, DC: World Bank Publications.

Del Ninno, C. and e. Mills, Bradford (2015). *Safety nets in Africa: Effective mechanisms to reach the poor and most vulnerable.* Washington, D.C.: Africa Development Forum Series.

Devereux, S. and K. Sharp (2006). Trends in poverty and destitution in Wollo, Ethiopia. *The Journal of Development Studies 42*(4), 592–610.

Filmer, D. and L. H. Pritchett (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. and K. Scott (2012). Assessing asset indices. *Demography 49*(1), 359–392.

Fink, G., P. J. Robyn, A. Sié, and R. Sauerborn (2013). Does health insurance improve health: Evidence from a randomized community-based insurance rollout in rural Burkina Faso. *Journal of Health Economics 32*(6), 1043–1056.

Foster, J., J. Greer, and E. Thorbecke (1984). A class of decomposable poverty measures. *Econometrica 52*(3), 761–766.

Galasso, E. and M. Ravallion (2005). Decentralized targeting of an antipoverty program. *Journal of Public Economics 89*(4), 705–727.

Grosh, M. and J. L. Baker (1995). Proxy means tests for targeting social programs. Living Standards Measurement Study Working Paper 118, The World Bank, Washington DC.

Handa, S., C. Huang, N. Hypher, C. Teixeira, F. V. Soares, and B. Davis (2012). Targeting effectiveness of social cash transfer programmes in three African countries. *Journal of Development Effectiveness 4*(1), 78–108.

Hargreaves, J. R., L. A. Morison, J. S. Gear, M. B. Makhubele, J. D. Porter, J. Busza, C. Watts, J. C. Kim, and P. M. Pronyk (2007). Hearing the voices of the poor: Assigning poverty lines on the basis of local perceptions of poverty. A quantitative analysis of qualitative data from participatory wealth ranking in rural South Africa. *World Development 35*(2), 212–229.

Honorati, M., U. Gentilini, and R. G. Yemtsov (2015). *The state of social safety nets 2015.* Washington, DC: World Bank Group.

Howe, L. D., J. R. Hargreaves, S. Gabrysch, and S. R. Huttly (2009). Is the wealth index a proxy for consumption expenditure? A systematic review. *Journal of Epidemiology and Community Health 63*(11), 871–877.

Innovations for Poverty Action (2014). 2014 global report on poverty measurement with the progress out of poverty index. Report, Grameen Foundation, Washington, DC. `http://www.progressoutofpoverty.org/global_report` (visited on 7/13/2017).

Jolliffe, D. (2001). Measuring absolute and relative poverty: the sensitivity of estimated household consumption to survey design. *Journal of Economic and Social Measurement 27*(1, 2), 1–23.

Karlan, D. and B. Thuysbaert (2016). Targeting ultra-poor households in Honduras and Peru. *The World Bank Economic Review 36*(3), 1–38.

Kebede, B. (2009). Community wealth ranking and household surveys: An integrative approach. *The Journal of Development Studies 45*(10), 1731–1746.

Klasen, S. and S. Lange (2014). Accuracy and poverty impacts of proxy means-tested transfers: An empirical assessment for Bolivia. Discussion Paper 164, Courant Research Centre, Göttingen.

Lietz, H., M. Lingani, A. Sié, R. Sauerborn, A. Souares, and Y. Tozan (2015). Measuring population health: Costs of alternative survey approaches in the Nouna Health and Demographic Surveillance System in rural Burkina Faso. *Global Health Action 8*(1), 28330.

Mayoux, L. and R. Chambers (2005). Reversing the paradigm: Quantification, participatory methods and pro-poor impact assessment. *Journal of International Development 17*(2), 271–298.

McGee, R. (2004). Constructing poverty trends in Uganda: A multidisciplinary perspective. *Development and Change 35*(3), 499–523.

Niehaus, P. and A. Atanassova (2013). Targeting with agents. *American Economic Journal: Economic Policy 5*(1), 206–238.

Odusola, A., R. Lal, R. Dhilwayo, J. Neuhaus, and I. Sabo (2017). Drivers of income inequality in Burkina Faso, Ghana and the United Republic of Tanzania: A comparative analysis. In *Income inequality trends in sub-Saharan Africa: Divergence, Determinants and Consequences*, Chapter 14, pp. 317–341. New York, USA: United Nations Development Programme.

Panda, S. (2015). Political connections and elite capture in a poverty alleviation programme in India. *Journal of Development Studies 51*(1), 50–65.

Ravallion, M. (1993). Poverty alleviation through regional targeting: A case study for Indonesia. In K. Hoff, A. Braverman, and J. E. Stiglitz (Eds.), *The Economics of Rural Organization*, Chapter 23, pp. 453–467. Oxford: Oxford University Press.

Ravallion, M. (2003). Targeted transfers in poor countries: Revisiting the trade-offs and policy options. Policy Research Working Paper 3048, The World Bank, Washington, D.C.

Ravallion, M. (2009). How relevant is targeting to the success of an antipoverty program? *The World Bank Research Observer 24*(2), 205–231.

Robano, V. and S. C. Smith (2013). Multidimensional targeting and evaluation: A general framework with an application to a poverty program in Bangladesh. IZA Discussion Papers 7593, Institute for the Study of Labor (IZA).

Robertson, L., P. Mushati, M. Skovdal, J. W. Eaton, J. C. Makoni, T. Crea, G. Mavise, L. Dumba, C. Schumacher, L. Sherr, C. Nyamukapa, and S. Gregson (2014). Involving communities in the targeting of cash transfer programs for vulnerable children: Opportunities and challenges. *World Development 54*(1), 325–337.

Sabates-Wheeler, R., A. Hurrell, and S. Devereux (2015). Targeting social transfer programmes: Comparing design and implementation errors across alternative mechanisms. *Journal of International Development 27*(8), 1521–1545.

Savadogo, G. (2017). *Targeting the poorest households for subsidizing their premium for community health insurance in Nouna, Burkina Faso.* Ph. D. thesis, Karl-Ruprecht-Universität Heidelberg, Heidelberg.

Savadogo, G., A. Souares, A. Sie, D. Parmar, G. Bibeau, and R. Sauerborn (2015). Using a community-based definition of poverty for targeting poor households for premium subsidies in the context of a community health insurance in Burkina Faso. *BMC Public Health 15*(84), 1–12.

Schreiner, M. (2015). A comparison of two simple, low-cost ways for local, pro-poor organizations to measure the poverty of their participants. *Social Indicators Research 124*(2), 537–569.

Schüring, E. (2014). Preferences for community-based targeting - field experimental evidence from Zambia. *World Development 54*(3), 360–373.

Scoones, I. (1995). Investigating difference: Applications of wealth ranking and household survey approaches among farming households in southern Zimbabwe. *Development and Change 26*(1), 67–88.

Sen, A. (1988). The concept of development. In H. Chenery and T. Srinivasan (Eds.), *Handbook in Economics 9*, Volume 1 of *Handbook of Development Economics*, Chapter 1, pp. 9–26. Elsevier.

Shaffer, P. (1998). Gender, poverty and deprivation: Evidence from the Republic of Guinea. *World Development 26*(12), 2119–2135.

Skoufias, E., B. Davis, and S. De La Vega (2001). Targeting the poor in Mexico: An evaluation of the selection of households into PROGRESA. *World Development 29*(10), 1769–1784.

Souares, A., G. Savadogo, H. Dong, D. Parmar, A. Sie, and R. Sauerborn (2010). Using community wealth ranking to identify the poor for subsidies: A case study of community-based health insurance in Nouna, Burkina Faso. *Health and Social Care in the Community 18*(4), 363–368.

Stoeffler, Q., B. Mills, and C. del Ninno (2016). Reaching the poor: Cash transfer program targeting in Cameroon. *World Development 83*(1), 244–263.

Takasaki, Y., B. L. Barham, and O. T. Coomes (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.

Temu, A. E. and J. M. Due (2000). Participatory appraisal approaches versus sample survey data collection: A case of smallholder farmer's well-being ranking in Njombe District, Tanzania. *Journal of African Economies 9*(1), 44–62.

Thomas, B. K., R. Muradian, G. De Groot, and A. De Ruijter (2009). Multidimensional poverty and identification of poor households: A case from Kerala, India. *Journal of Human Development and Capabilities 10*(2), 237–257.

UN-DESA (2015). World urbanization prospects: The 2014 revision. Report ST/ESA/SER.A/366, United Nations, Department of Economic and Social Affairs, Population Division.

Van Campenhout, B. F. (2007). Locally adapted poverty indicators derived from participatory wealth rankings: A case of four villages in rural Tanzania. *Journal of African Economies 16*(3), 406–438.

Wagstaff, A. and N. Watanabe (2003). What difference does the choice of SES make in health inequality measurement? *Health Economics 12*(10), 885–890.

World Bank (1993). Indonesia: Public expenditures, prices and the poor. Report 11293-IND, The World Bank, Indonesia Resident Mission, Jakarta.

Yamauchi, C. (2010). Community-based targeting and initial local conditions: Evidence from Indonesias IDT program. *Economic Development and Cultural Change 59*(1), 95–147.

Table 1: Community wealth rankings and community-based targeting: procedural details

	Country	Study Population (Villages/ Households)	Focus Group Discussions	Wealth Brackets	Complete Wealth Ranking	Targeting	Number of Informants
			(1)	(2)	(3)	(4)	(5)
Latin America							
Takasaki et al. (2000)	Peru	8/300	*n.r.*	*n.r.*	YES	NO	3 – 4
Karlan and Thuysbaert (2016)	Honduras	40/1,060	YES	YES	NO	YES	*n.r.*
Karlan and Thuysbaert (2016)	Peru	40/1,007	YES	YES	NO	YES	*n.r.*
Asia							
Adams et al. (1997)	Bangladesh	55/1,637	*n.r.*	YES	NO	NO	5
Banerjee et al. (2007)	India	5/213	*n.r.*	YES	NO	YES	*n.r.*
Caizhen (2010)	China	1/473	*n.r.*	YES	YES	YES	ALL
Alatas et al. (2012)	Indonesia	640/5,753	YES	NO	YES	YES	ALL
Sub-Saharan Africa							
Scoones (1995)	Zimbabwe	1/21	YES	YES	NO	NO	ALL
Shaffer (1998)	Guinea	1/8	*n.r.*	YES	YES	NO	8
Temu and Due (2000)	Tanzania	12/300	*n.r.*	YES	YES	YES	6
Hargreaves et al. (2007)	South Africa	8/9,671	YES	YES	YES	YES	*n.r.*
Van Campenhout (2007)	Tanzania	4/877	*n.r.*	YES	YES	NO	1
Kebede (2009)	various	37/1,300	*n.r.*	*n.r.*	*n.r.*	NO	*n.r.*
Souares et al. (2010)	Burkina Faso	57/910	YES	YES	YES	YES	3
Handa et al. (2012)	Malawi	7/9,840	NO	NO	YES	YES	5
Robertson et al. (2014)	Zimbabwe	30/12,000	YES	YES	YES	YES	5
Sabates-Wheeler et al. (2015)	Kenya	48/5,108	*n.r.*	*n.r.*	YES	YES	*n.r.*
Stoeffler et al. (2016)	Cameroon	15/4,300	YES	YES	NO	YES	*n.r.*

Notes: *n.r.*: Not reported, ALL: Whole community

Table 2: Community-based targeting and community characteristics

	Pooled	Rural	Semi-urban
Community-based targeting			
Ranked households per community	106	114	93
	(72)	(87)	(39)
Targeted households per community	21	23	18
	(14)	(18)	(7)
Targeted households per community (share)	0.20	0.20	0.20
	(0.01)	(0.01)	(0.01)
Targeted by all 3 informants	0.08	0.08	0.08
	(0.04)	(0.04)	(0.02)
Targeted by exactly 2 informants	0.11	0.10	0.12
	(0.31)	(0.30)	(0.32)
Targeted by exactly 1 informant	0.21	0.21	0.20
	(0.41)	(0.41)	(0.40)
Mean rank correl., CBT key informants	0.65	0.65	0.65
	(0.12)	(0.14)	(0.10)
Any key informant with education	0.66	0.59	0.75
	(0.48)	(0.50)	(0.44)
Community characteristics			
Gini Index Consumption	0.35	0.31	0.41
	(0.12)	(0.10)	(0.12)
Gini Index Assets (PC)	0.42	0.42	0.43
	(0.04)	(0.05)	(0.03)
ELF Index for ethnicity	0.46	0.33	0.67
	(0.28)	(0.25)	(0.18)
Number of communities	58	36	22
Number of households	5708	3655	2053

Notes: Standard deviations in parentheses. All sample means are calculated at the community level. A community is a village or urban sub-sector. *ELF* is the ethno-linguistic fractionalization index and measures the probability that two randomly drawn individuals belong to different ethnic or religious groups, respectively. *Gini* is the Gini index for an asset index obtained trough principal-component analysis from a total of 25 assets.

Table 3: Household survey summary statistics

	Pooled	Rural	Semi-urban
Community-based targeting			
Number of ranked households	11.3	11.8	10.3
	(7.9)	(9.2)	(5.2)
Number of targeted households	2.5	2.5	2.4
	(2.1)	(2.3)	(1.6)
Share of targeted households	0.22	0.21	0.23
	(0.41)	(0.41)	(0.42)
Consumption			
Monthly per capita consumption (CFA)	4,808	3,026	8,152
	(8,287)	(2,209)	(13,104)
Market consumption (CFA)	4,119	2,240	7,645
	(8,319)	(2,006)	(13,151)
Home-produced consumption (CFA)	689	785	508
	(523)	(529)	(460)
Food share (percent)	60.72	67.01	48.90
	(23.47)	(21.45)	(22.56)
Demographics			
Household size	8.48	8.90	7.69
	(5.60)	(5.93)	(4.85)
Female headed household	0.12	0.09	0.19
	(0.33)	(0.29)	(0.39)
Share of adolescent household members	0.40	0.43	0.33
	(0.19)	(0.19)	(0.18)
Share of elderly household members	0.10	0.09	0.13
	(0.19)	(0.18)	(0.21)
Household head literate (incidence)	0.36	0.33	0.40
	(0.48)	(0.47)	(0.49)
HH head occup. non-agric. (incidence)	0.28	0.16	0.50
	(0.45)	(0.37)	(0.50)
Ethnic minority (incidence)	0.17	0.25	0.02
	(0.38)	(0.43)	(0.15)
HH head sick during last month (incidence)	0.21	0.21	0.22
	(0.41)	(0.41)	(0.42)
Asset possession (incidences)			
Bullock	0.47	0.52	0.39
	(0.50)	(0.50)	(0.49)
Goat or sheep	0.76	0.85	0.59
	(0.43)	(0.36)	(0.49)
Motorbike	0.23	0.17	0.33
	(0.42)	(0.38)	(0.47)
Bicycle	0.91	0.89	0.93
	(0.29)	(0.31)	(0.25)
Number of communities	58	36	22
Number of households	653	426	227

Notes: Standard deviations in parentheses. Adolescent (elderly) household members are all household members below (above) the age of 16 (55). Ethnic minority is a dummy variable equal to one if the household head belongs to the community's smallest ethnic group.

Table 4: Targeting methods

Set	Description	Number of indicators	Transformation of indicators	Weighting of indicators
Benchmark				
T^{CON}	Monthly household per capita consumption (eqv. scales)	1	None	n.a.
Statistical targeting				
Linear regression				
T^{ECON}	Econometric Proxy-means Test	44	None	OLS
Principal component analysis				
T^{PCA}	Asset Index	44	None	PCA
Scorecard				
T^{PSI}	Poverty Scorecard Index (PSI)	9	Ordered categorical	hybrid
Counting of deprivations				
T^{MPI}	Multidimensional Poverty Index	9	Binary	Global MPI
Community-based targeting				
T^{CBT}	Households identified by three local informants	n.a.	n.a.	n.a.

Notes: n.a.: not applicable; OLS: ordinary least squares regression of consumption, transformed into a dummy variable, on the set of all indicators; PCA: principal component analysis; hybrid: implicit weights for the different indicators as chosen by the creators of the Burkina poverty scorecard; hybrid: implicit weights for the indicators obtained by predicting consumption from the scorecard indicators using a national household survey (Schreiner, 2015); equal: equal weight is given to each transformed indicator; Global MPI: Weighting scheme follows that of the global MPI.

Table 5: Targeting errors (in percent)

	Pooled			Rural			Semi-urban		
	(1) Mean targeting error	(2) Mean exclusion error	(3) Mean inclusion error	(4) Mean targeting error	(5) Mean exclusion error	(6) Mean inclusion error	(7) Mean targeting error	(8) Mean exclusion error	(9) Mean inclusion error
Econometric PMT	23.3	53.1	14.9	23.5	54.9	14.9	22.9	50.0	14.9
	(1.7)	(4.2)	(1.6)	(2.1)	(5.2)	(1.9)	(2.8)	(7.0)	(2.7)
Asset Index	27.0	61.5	17.3	28.2	65.9	17.9	24.7	53.8	16.0
	(1.7)	(4.1)	(1.7)	(2.2)	(5.0)	(2.1)	(2.9)	(7.0)	(2.8)
Poverty Scorecard Index	28.5	65.0	18.2	29.1	68.1	18.5	27.3	59.6	17.7
	(1.8)	(4.0)	(1.7)	(2.2)	(4.9)	(2.1)	(3.0)	(6.9)	(2.9)
Multidimensional Poverty Index	30.0	68.5	19.2	28.6	67.0	18.2	32.6	71.2	21.1
	(1.8)	(3.9)	(1.7)	(2.2)	(5.0)	(2.1)	(3.1)	(6.3)	(3.1)
Community-based targeting	29.4	67.1	18.8	32.4	75.8	20.6	23.8	51.9	15.4
	(1.8)	(3.9)	(1.7)	(2.3)	(4.5)	(2.2)	(2.8)	(7.0)	(2.7)
Random targeting error	34.3	78.0	22.0	33.2	79.0	21.0	35.4	77.0	23.0
Number of households	653	143	510	426	91	335	227	52	175

Notes: Standard errors in parentheses. Mean targeting error is defined as the proportion of households which are erroneously classified as either poor or non-poor. The exclusion error is defined as the proportion of households which are consumption poor but not targeted by the targeting method under consideration. The inclusion error is defined as the proportion of households which are consumption non-poor but targeted by the targeting method under consideration.

Table 6: Targeting errors by consumption expenditure classes

	Rural				Semi-urban			
	(1) Extremely poor	(2) Moder. poor	(3) Around median	(4) Affluent	(5) Extremely poor	(6) Moder. poor	(7) Around median	(8) Affluent
	Exclusion error		Inclusion error		Exclusion error		Inclusion error	
Econometric PMT	50.0	58.2	16.7	12.9	38.1	58.1	16.1	13.4
	(8.5)	(6.7)	(2.8)	(2.7)	(10.9)	(9.0)	(3.8)	(3.8)
Asset Index	69.4	63.6	16.1	20.0	28.6	71.0	22.6	8.5
	(7.8)	(6.5)	(2.7)	(3.2)	(10.1)	(8.3)	(4.4)	(3.1)
Poverty Scorecard Index	63.9	70.9	20.6	16.1	76.2	48.4	23.7	11.0
	(8.1)	(6.2)	(3.0)	(3.0)	(9.5)	(9.1)	(4.4)	(3.5)
Multidimensional Poverty Index	66.7	67.3	18.3	18.1	52.4	83.9	26.9	14.6
	(8.0)	(6.4)	(2.9)	(3.1)	(11.2)	(6.7)	(4.6)	(3.9)
Community-based targeting	80.6	72.7	19.4	21.9	33.3	64.5	22.6	7.3
	(6.7)	(6.1)	(3.0)	(3.3)	(10.5)	(8.7)	(4.4)	(2.9)
Random targeting error	79.0	79.0	21.0	21.0	77.0	77.0	23.0	23.0
Number of households	36	55	180	155	21	31	93	82

Notes: All figures are in percent. Standard errors are in parentheses. The expenditure classes are defined such that for each community the shares of "extremely poor" and "moderately poor" households are roughly equal and sum up to the sample targeting share of the community-based targeting exercise. Analogously, for each community the sample shares of "around median" and "affluent" households sum up to one minus the community's sample targeting share of the CBT.

Table 7: Targeting errors and community characteristics

Dependent variable:	**Targeting error**							
	PMT	Asset Index	PSI	MPI	Community-based targeting			
		Full sample			Full sample		Restricted sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rural Community (Dummy)	0.006	0.035	0.018	-0.039	0.088**	0.106**	0.109*	0.099*
	(0.037)	(0.037)	(0.038)	(0.030)	(0.036)	(0.053)	(0.057)	(0.057)
Number of households (hundreds)						-0.009	-0.045	-0.000
						(0.183)	(0.192)	(0.202)
Gini Index Consumption						0.143	0.070	0.047
						(0.208)	(0.211)	(0.212)
Gini Index Assets (PC)						0.351	0.650	0.662
						(0.500)	(0.705)	(0.665)
ELF-Index for ethnicity						0.006	0.023	-0.014
						(0.081)	(0.088)	(0.091)
Any key informant with education								0.036
								(0.035)
Mean rank corr., CBT key informants								0.166
								(0.118)
Observations	653	653	653	653	653	653	581	581
Communities	58	58	58	58	58	58	47	47

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Table reports marginal effects of probit estimations. Standard errors clustered at the level of a community in parentheses. PCA consumption assets include number of rooms, roof quality, wall quality, floor quality, toilet mode, sewage mode, water source, bicycle, motorbike, TV, fridge, and kitchen.

Table 8: Determinants of CBT selection

	Dependent variable: Targeted by community					
	(1)	(2)	(3)	(4)	(5)	(6)
	Urban	Rural	T-test for equal coeff.	Urban	Rural	T-test for equal coeff.
Log(Monthly per capita consumption, CFA)	-0.095***	0.025	[0.004]	-0.026	0.053***	[0.011]
	(0.035)	(0.019)		(0.022)	(0.019)	
Consumption Asset Index (PC)				-0.069***	-0.078***	[0.998]
				(0.014)	(0.021)	
Agricultural Asset Index (PC)				-0.055***	-0.041***	[0.317]
				(0.014)	(0.012)	
Household head literate				-0.082	-0.075*	[0.793]
				(0.050)	(0.039)	
Female headed household				0.100**	0.074	[0.627]
				(0.048)	(0.060)	
Log(Household size)				-0.001	-0.031	[0.631]
				(0.043)	(0.038)	
Share of elderly household members				0.024	0.337***	[0.088]
				(0.127)	(0.109)	
Share of adolescent household members				-0.119	-0.006	[0.476]
				(0.130)	(0.106)	
Household belongs to ethnic minority				-0.236	-0.075*	[0.284]
				(0.153)	(0.043)	
Household head sick during last month				-0.049	0.070*	[0.083]
				(0.053)	(0.041)	
Households	227	426		227	426	

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Table reports marginal effects of probit estimations. Robust standard errors in parentheses. P-values for t-tests of equal coefficients in rural and urban communities in brackets. PCA consumption assets include number of rooms, roof quality, wall quality, floor quality, toilet mode, sewage mode, water source, bicycle, motorbike, TV, fridge, and kitchen. PCA agricultural assets include chicken, pig, goat, sheep, horse, donkey, bullock, cart, and plow.

Table 9: Cost-benefit analysis

Method A delivers higher benefits than method B for program costs smaller than ...		Cost scenarios for statistical targeting		
		(1)	(2)	(3)
A	B	Data processing and no data collection	Data processing, consumption data collection	Data processing, full data collection
Rural				
CBT	*Econometric PMT*	0.78	25.03	45.91
CBT	*Asset index*	0	0	43.24
Asset index	*Econometric PMT*	3.47	68.41	50.38
Semi-urban				
CBT	*Econometric PMT*	41.70	572.05	1028.40
CBT	*Asset index*	always	always	always
Asset index	*Econometric PMT*	17.49	272.19	254.16

Notes: All figures are in 2014 US dollars, not PPP adjusted. The maximum cost per eligible household for which method A is more cost-effective than method B is calculated by solving the equation $B(C; E^A, TC^A) = B(C; E^B, TC^B)$ for C, where B is the benefit per eligible households (see equation 1). In column 1 we consider only data processing costs of the census for the asset index ($1.33 per eligible household) and data processing costs of the census and the consumption survey for the econometric PMT ($2.88 per eligible household). In column 2 we consider data processing costs of the census for the asset index ($1.33 per eligible household), and data processing costs of the census and the consumption survey as well as data collection costs of the consumption survey for the econometric PMT ($23.83 per eligible household). In column 3 we consider data collection and processing costs of the census for the asset index ($19.36 per eligible household), and data collection and processing costs of the census and the consumption survey for the econometric PMT ($41.86 per eligible household). For community-based targeting we consider the CBT implementation and data processing costs of $1.19 per eligible household throughout.

Online Appendix

Table of Contents:

List of Figures

List of Tables

1) Construction of sample target sets

In our main analysis we construct and compare five different sets of target households to a reference set of consumption-poorest households. First, the set of households actually targeted by the communities. We denote the corresponding set of targeted households in community c in our sample of 653 households in total by T_c^{CBT}. The remaining four statistical targeting sets are constructed from the household survey data as follows. Let n_c^{CBT} denote the number of sample households targeted by CBT in community c. To construct T_c^m, the hypothetical targeting set for statistical method m in community c, we first sort all sample households in c by the score of method m and select the n_c^{CBT} households with the lowest scores.[21] The aggregate sample target set is the union $T^m = \cup_{c=1}^{C} T_c^m$, where C denotes the number of communities. Following the recent comparative targeting accuracy literature (Alatas et al., 2012; Filmer and Scott, 2012; Karlan and Thuysbaert, 2016; Klasen and Lange, 2014; Brown et al., 2018; Stoeffler et al., 2016), we take consumption as the benchmark and construct the reference set of consumption-poorest households analogously: we sort all sample households in community c by per capita consumption and select the n_c^{CBT} households with the lowest values to obtain T_c^{CON}. The aggregate reference set is the union $T^{CON} = \cup_{c=1}^{C} T_c^{CON}$.

[21] To illustrate, consider a village with 100 households from which 20 have been targeted by the community-based method. Suppose that, from this village, the household survey includes 10 households with 3 sample households targeted and 7 not targeted. While the population targeting set contains 20 percent of all households, all six of our sample target sets contain precisely 3 sample households from that village.

a. Econometric PMT

For the econometric PMT we define a dummy variable $Elig_{ci}^{CON}$ equal to one if household i in community c is targeted according to our benchmark. We then estimate the linear regression model

$$Elig_{ci}^{CON} = \alpha_c + \beta x_{ci} + u_{ci},$$

where x_{ci} is a vector of proxy variables, β is the corresponding coefficient vector, α_c are community fixed effects, and u is a stochastic error term.[22] We conduct this regression analysis separately for rural and semi-urban households (see Table A3 in the Online Appendix for the regression output). To avoid overfitting, we estimate β with data from the 2008 Nouna Household Survey. In a second step, a score is calculated for each household as $\widehat{Elig}_{ci}^{CON} = \widehat{\beta} x_{ci}$, where x_{ci} comes from the 2009 household survey data. Third, for each community, households are sorted with respect to their scores and the n_c^{CBT} lowest ranked households are assigned to T_c^{ECON}.

b. Asset Index

We conduct a principal component analysis to obtain weights that derive from the joint distribution of the proxy-means variables themselves. Our asset index employs the same set of indicators as the econometric PMT. To focus on within-community differences across households, we first subtract community means from each indicator. Following Filmer and Pritchett (2001), we take the first principal component as score and the corresponding factor loadings as weights (see Appendix Table A4 for details).

c. Poverty Scorecard Index

For the Poverty Scorecard Index (PSI) we use the 2011 Burkina Faso poverty scorecard from IPA's website[23] to calculate indicator scores and to construct the hypothetical target set T^{PSI}. Of its

[22] An alternative regression specification has logarithmic MPC as dependent variable (Alatas et al., 2012). The accuracy of our econometric PMT deteriorates only marginally with this alternative dependent variable.

[23] See `http://www.progressoutofpoverty.org/country/burkina-faso`.

ten indicators we omit one which is not covered in our household survey, "Does the household own a bed or mattress?". As this indicator accounts for not more than three percent of the full score, we are confident that this omission will not threaten its overall performance. Table A5 of the Online Appendix provides details on this scorecard.

d. The Multidimensional Poverty Index

We build on the Global Multidimensional Poverty Index (MPI) by Alkire and Santos (2010). For reasons of data availability, we make two modifications regarding the health dimension. We omit the nutrition indicator and substitute the child mortality indicator by the incidence of a recent severe health shock, which is recorded in our sample survey with a recall of one month.[24] We adjust the weights for this dimension accordingly (see Online Appendix Table A6 for details).

2) Monte-Carlo study of the mean targeting error estimator

A potential challenge to our estimation of mean targeting errors at the level of the community is the relatively small sample size per community of about eleven households. While there is little variation in the share of households targeted by our CBT intervention across communities (according to Table 2 the average share equals 0.20 with a standard deviation of merely 0.01), this figure varies greatly in our sample. According to Table 3, the average share equals 0.22 with a standard deviation of 0.41. As outlined in the first section of this Online Appendix, by construction, the reference set as well as the statistical target sets for a given community each have an identical number of households equal to the number of sample households targeted by CBT in the respective community. While our procedure delivers consistent estimates of all methods' targeting errors when the asymptotics are such that the share of sampled households approaches

[24]With an indicator weight of about 16 percent, the omission of nutrition in the MPI could have a greater effect on targeting performance than the omission of the type of bed on the Poverty Scorecard and we shall be cautious in considering this limitation when interpreting the results.

one, there is a concern of small sample bias.

To address this issue, in this section, we conduct a Monte-Carlo simulation of our method with a data-generating process that employs a multivariate normal distribution, which we calibrate with alternative true MTEs for the different targeting methods. In a nutshell, the findings from this exercise are as follows. While there is a downward bias regarding the levels of MTEs of around ten percent across all targeting methods, biases are very small regarding differences in MTEs between targeting methods, the main object of interest of our study, especially when the true MTEs are similar. In comparisons of CBT with statistical targeting methods, these biases always slightly favor CBT but do not exceed one fifth of a standard error in the empirically relevant range of MTE values. When the true MTEs differ across two methods, the bias of our procedure does not favor any one method systematically and the bias' magnitude continues to be relatively small. Bias corrections based on these simulations would not qualitatively affect any of our main findings.

In this Monte-Carlo simulation, for simplicity, we will speak of comparisons of community-based targeting and statistical targeting using the asset index, but the insights are applicable to all pairwise comparisons of CBT and one given statistical method. For the simulation, we consider one community and three outcomes at the level of households, consumption, C, asset index, A, and a hypothetical index that underlies community-based targeting, B. We assume that these three random variables are jointly normally distributed with mean zero and correlation matrix Σ. The three parameters to be calibrated are the three pairwise correlations between the three methods. When procedure D is the reference method (in our case consumption), E the method under examination (in our case either CBT or the statistical targeting based on the asset index), and the share of eligible households under either procedure is q, the *population mean targeting* *error* for this data-generating process is

$$\begin{aligned}
MTE_D^E(c,\rho_{DE}) &= \Pr(E<c\cap B>c)+\Pr(E>c\cap B<c)\\
&= 1-\Pr(E<c\cap B<c)-\Pr(E>c\cap B>c)\\
&= 1-\Pr(E<c\cap B<c)-\Pr(E<-c\cap B<-c)\\
&= 1-\Phi^{Biv}(c,c,\rho_{DE})-\Phi^{Biv}(-c,-c,\rho_{DE}),\\
\text{where } c &= \Phi^{-1}(q).
\end{aligned}$$

Here $\Phi(\ )$ denotes the standard normal cumulative distribution function and $\Phi^{Biv}(\ ,\ ,\rho)$ the cumulative distribution function of a bivariate normal distribution with unit variances and correlation coefficient ρ.

To simulate sampling, we make n draws of 1×3 vectors from this distribution, where n corresponds to the number of households in a community. This gives a $n\times 3$ matrix, S say, from which we obtain a matrix of targeting outcomes with each of the three methods, O say, by defining

$$o_{ik}=\begin{cases}1 & \text{if } s_{ik}\le s_{qn:n,k}\\ 0 & \text{otherwise,}\end{cases}$$

where $s_{qn:n,k}$ is the qn'th smallest element in the k'th column of S. We thus obtain the *community mean targeting error* for procedure E, which equals either A or B, as

$$\widetilde{MTE}_C^E=\frac{1}{n}\sum_{i=1}^{n}|o_{ik}-o_{i1}|,$$

where k equals 2 for the asset index and 3 for community based targeting, respectively. To calculate the mean targeting error in a small sample from this community according to the method outlined in section 4.1, we first construct the $n\times 3$ matrix S', whose first two columns are equal to those in S while its third column equals O's third column. We then randomly draw m rows from S' without replacement and arrange them in the $m \times 3$ matrix $\widehat{S}$. The number of CBT targeted households in the sample is $n^{CBT} = \sum_{i=1}^{m} \widehat{s}_{i3}$ and we obtain the $m \times 3$ sample target outcome matrix $\widehat{O}$ by defining

$$\widehat{o}_{ik} = \begin{cases} 1 & \widehat{s}_{ik} \leq s_{n^{CBT}:m,k} \\ 0 & \text{otherwise} \end{cases}$$

for $k = 1,\ 2$, and $\widehat{o}_{ik} = \widehat{s}_{i3}$ for $k = 3$. In words, for both consumption and the asset index, the n^{CBT} poorest households are targeted where n^{CBT} equals the number of CBT targeted households in the sample. For one community, we obtain the *sample mean targeting error* for procedure D, which equals either A or B, as

$$\widehat{MTE}_{C}^{D} = \frac{1}{m} \sum_{i=1}^{m} |\widehat{o}_{ik} - \widehat{o}_{i1}|.$$

We repeat this procedure many times to obtain the expected values of $\widetilde{MTE}_{C}^{A}$, $\widetilde{MTE}_{C}^{B}$, $\widehat{MTE}_{C}^{A}$ and $\widehat{MTE}_{C}^{B}$. Given the share of targeted households, q, and the pairwise correlations between $C,\ A$ and B, the objects of interest are $\widetilde{MTE}_{C}^{A}$ and $\widetilde{MTE}_{C}^{B}$, the expected mean targeting errors at the level of the community. In our empirical exercise, on the other hand, we estimate $\widehat{MTE}_{C}^{A}$ and $\widehat{MTE}_{C}^{B}$. Hence, in our simulations we assess two kinds of biases, first bias concerning the level of sample MTEs for asset-index and community-based targeting, and second the bias of the difference between these two sample MTEs. We calibrate the simulations by first solving the following system of four equations

$$\begin{aligned} \Phi(c) &= q \\ MTE_{C}^{B}(c, \rho_{CB}) &= MTE_{C}^{B} \\ MTE_{C}^{A}(c, \rho_{CA}) &= MTE_{C}^{A} \\ MTE_{A}^{B}(c, \rho_{AB}) &= MTE_{A}^{B}, \end{aligned}$$

where, throughout, we set $q = 0.22$, the average share of targeted households in our sample, $n = 106$, the average number of households per community in our data, $m = 11$, the average number of households in our sample per community, and $MTE_A^B = 0.17$, which we have obtained from an estimation of the mean targeting error of CBT when the asset index is the benchmark (not shown in the paper). We consider alternative values of MTE_C^B and MTE_C^A between 0.10 and 0.35. Notice that a mean targeting error of 34.3% results from random targeting (Table 5, column 1). We run 1,000,000 repetitions.

The results of our simulations are set out in Table A7 in the Appendix. The first column and row contain the calibrated population mean targeting errors for asset-index and community-based targeting in percentage points, respectively. The second column and row contain the community mean targeting errors for the two procedures, respectively, corresponding to the population MTEs displayed to the left and above, respectively. For obvious reasons, the second column and second row are identical. The third column and row contain the sample mean targeting errors. The sample MTE for CBT depends only on the population MTE for CBT while the sample MTE for asset-index targeting depends on all three population MTEs, MTE_C^B, MTE_C^A and MTE_A^B. Each entry in the third row contains the CBT sample MTE for a value of the population counterpart, MTE_C^B, displayed two rows above, while an entry in the third column contains the asset-index sample MTE for a value of the population mean targeting errors MTE_C^A and MTE_c^B equal to the respective value in the first column, and $MTE_A^B = 0.17$. Unlike the second row and column, the third row and the third column are not identical. In particular, the sample MTE for asset-index targeting is up to 2.1 percentage points higher than that of CBT - when the two population MTEs equal 10 percent. The biases of the sample MTEs relative to the community MTEs are obtained as the differences between the third and second column (row) for asset-index (community-based) targeting. For both methods the bias is similar in magnitude, increasing in the population MTEs, and equal to about two and a half percentage points for the empirically relevant range around population MTEs of 25 to 30%.

Each of the remaining cells of the table contains the bias of the difference in sample mean targeting errors corresponding to the respective column and row headers, to be precise the difference between CBT and asset-index sample MTEs relative to the respective community MTEs. For example the upper-left-most shaded entry gives the bias of the difference estimator of the sample MTEs of community based and asset-index targeting when both procedures have a population MTE of 10 percent. The bias of the difference does in general not equal the difference in biases of the two methods' sample MTEs as obtained from the second and third row and column, respectively, because the asset-index sample MTE depends also on MTE_C^B. To give an example, in the simulation for population MTEs of 20 percent for both asset-index and community-based targeting, the bias of the difference in sample MTEs equals 0.6 percentage points while the difference between the individual biases is only $0.8 - (0.4) = 0.4$.

The findings of this exercise are as follows. First, regarding the level of MTEs, our estimation procedure appears to systematically underestimates the MTEs of both targeting methods. The bias for CBT increases from 2.4 to 3.8 percentage points as the population MTE increases from 10 to 34.3 percent. For the asset index, the bias has the same increasing pattern, from 0.3 to 3.8. Second, regarding MTE comparisons between the two methods, the bias in our estimation procedure slightly favors CBT when the population MTEs of the two procedures are identical; the shaded cells on the diagonal of Table A7 are all positive. In the empirically relevant range between 20 and 30 percent (shaded 3x3 block of cells) this bias is small, however, and not greater than 0.6. Third, for the same range, our estimator of the difference in MTEs tends to slightly understate the true difference between the methods in absolute value when the population MTEs of the two procedures differ. This pattern is symmetric across the two methods. To give an example, when the asset-index and CBT population MTEs equal 20 and 25 percent (and hence the community MTEs 20.0 and 24.9 percent), respectively, our difference estimator will equal 6.1 in expectation (the true difference of 4.9 plus 1.2 bias). To illustrate the symmetry, when the asset-index and CBT population MTEs equal 25 and 20 percent (and hence the community MTEs 24.9 and 20.0

percent), respectively, our difference estimator will equal -5.2 in expectation (the true difference of -4.9 plus -0.3 bias).

To summarize, our MTE estimator appears to perform reasonably well. First, we think that the downward bias regarding the levels of MTEs of individual targeting methods is not so much of an issue given that the measurement of consumption itself is typically subject to a sizable measurement error (Deaton and Zaidi, 2002), which will affect the targeting accuracy of alternative targeting methods similarly. Second, regarding differences in MTEs, the main object of interest of our study, the bias is at most a third of one standard error in the empirically relevant range of MTE values (0.6 relative to 1.7 percentage points in the pooled sample) for the null hypothesis of equality of MTEs across two distinct targeting methods. Finally, when the true community MTEs differ across two methods, the bias of our procedure does not favor any one method systematically and the magnitude is relatively small. No single statistically significant estimate would be affected were the data-generating process as modeled here and the estimates corrected for biases accordingly.

In addition to the simulations discussed thus far we have also simulated a data-generating process corresponding to our empirical findings for the econometric proxy-means test and community-based targeting. The only difference to the previous exercise is a different calibration value of MTE_A^B, equal to 0.27 instead of 0.17 previously. The resulting biases are very similar to the ones reported in Table A7. The most important finding is that, of the MTE difference between econometric and community-based targeting of 6 percentage points in the pooled data, again around one fifth would be due to bias. As before, this would not affect the conclusions of our comparisons in any way.

3) Targeting accuracy and poverty reduction

We consider a distribution of consumption y characterized by the smooth density function $f(y)$. Now consider an anti-poverty program with a uniform transfer per beneficiary household that relies on some targeting method. We assume that the inclusion probability of that method (equal to one minus the exclusion error), IP say, may vary along the realization of consumption, and accordingly write $IP(y)$. Then, for a given transfer per eligible household of amount t, which we assume to increase consumption by the same amount, the consumption distribution after implementation of the program is

$$f(y;t) = (1 - IP(y))\, f(y) + IP(y-t) f(y-t).$$

In words, the density at y equals the exclusion error at y times the original density at y plus the inclusion probability at $y - t$ times the original density at $y - t$. It is easily verified that $f(y;t)$ integrates to one.

Regarding poverty indices, we consider the poverty gap index (Foster et al., 1984). We write

$$P(t) = \int_0^p \frac{p-y}{p} f(y;t) dy,$$

where p is the poverty line. We evaluate poverty reduction resulting from the anti-poverty program under consideration with a small transfer amount in place by a first-order Taylor expansion, which equals t times the derivative of $P(t)$ with respect to t evaluated at t equal to zero. Straightforward calculations yield

$$\Delta P \approx -\frac{F(p)}{p} E[IP(Y)|Y \leq p]t.$$

Notice that $F(p)$ equals the headcount ratio in the absence of the program, which, according to our assumptions, is equal to the share of targeted households in the population (20 percent in our application). Further, $E[IP(Y)|Y \leq p]$ is just the probability limit of the average sample inclusion probability, or one minus the exclusion error, where the latter corresponds precisely to the estimates set out in columns 2 and 5 of Table 5.

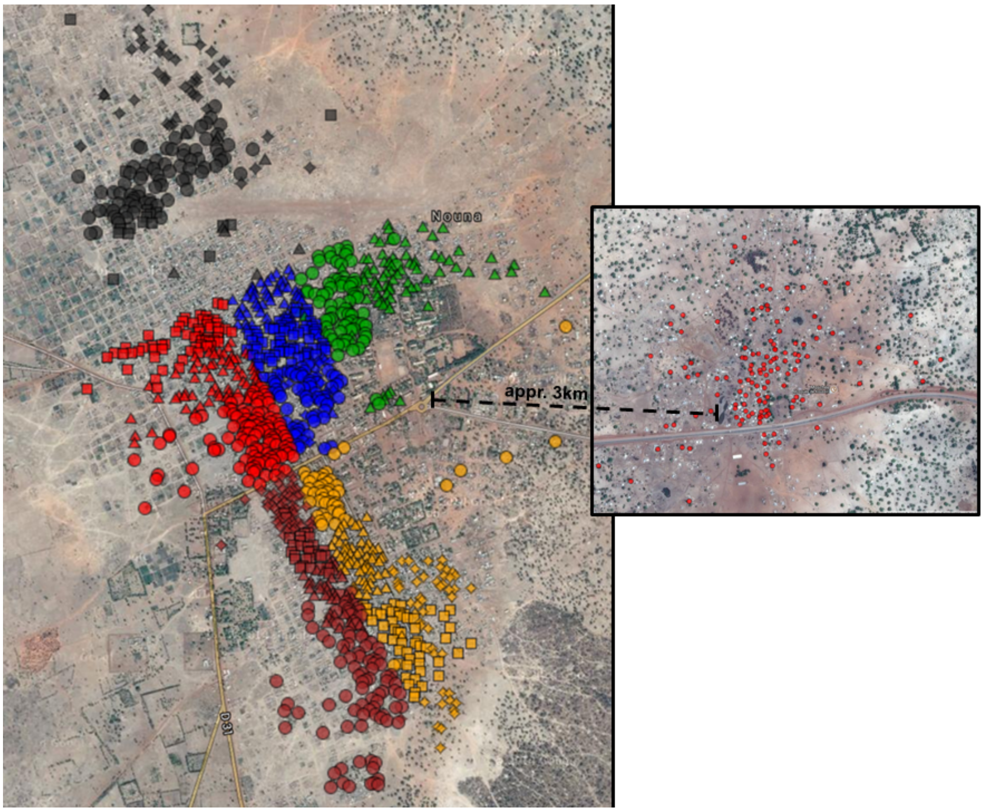

Figure A1: Nouna town and its neighborhoods

Notes: Map depicts the distribution of households across the seven sectors of Nouna town. Districts are distinguished by different colors. Within each sector there are two to four neighborhoods depicted by different symbols. District 7 is illustrated by a separate window as it is located approximately three kilometers in the West of the town's center.

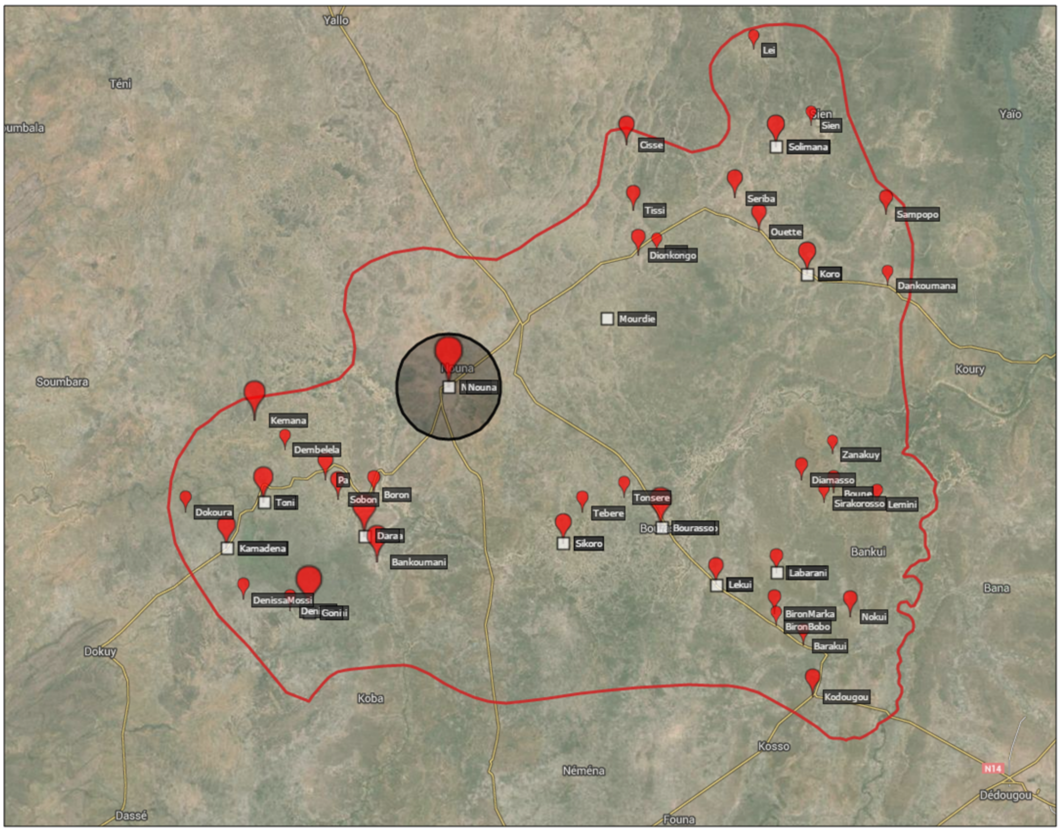

Figure A2: The survey site

Notes: Map depicts the location and population of the 41 villages. For Nouna town see Figure A1. Created with GPS Visualizer.

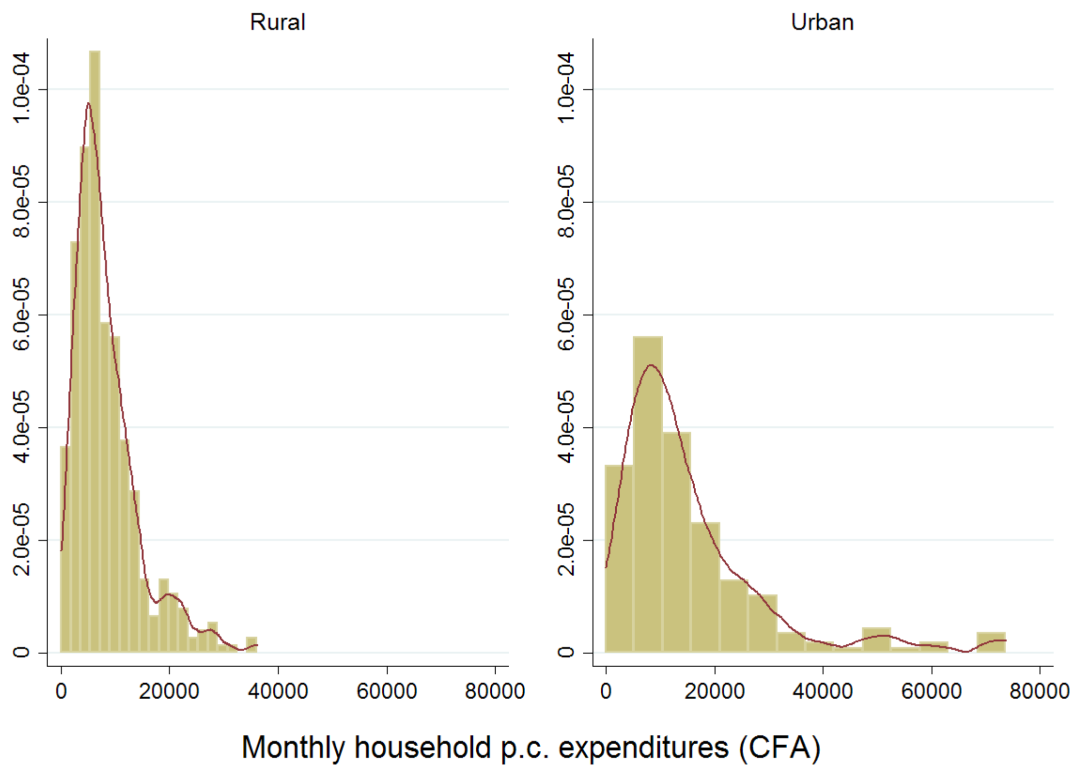

Figure A3: Histogram and kernel densities of the consumption reference variable

Table A1: Targeting errors (in percent), alternative consumption definition

	Pooled			Rural			Semi-urban		
	(1) Mean targeting error	(2) Mean exclusion error	(3) Mean inclusion error	(4) Mean targeting error	(5) Mean exclusion error	(6) Mean inclusion error	(7) Mean targeting error	(8) Mean exclusion error	(9) Mean inclusion error
Econometric PMT	25.7	58.7	16.5	26.8	62.6	17.0	23.8	51.9	15.4
	(1.7)	(4.1)	(1.6)	(2.1)	(5.1)	(2.1)	(2.8)	(7.0)	(2.7)
Asset Index	26.3	60.1	16.9	27.2	63.7	17.3	24.7	53.8	16.0
	(1.7)	(4.1)	(1.7)	(2.2)	(5.1)	(2.1)	(2.9)	(7.0)	(2.8)
Poverty Scorecard Index	27.9	63.6	17.8	28.2	65.9	17.9	27.3	59.6	17.7
	(1.8)	(4.0)	(1.7)	(2.2)	(5.0)	(2.1)	(3.0)	(6.9)	(2.9)
Multidimensional Poverty Index	29.1	66.4	18.6	28.2	65.9	17.9	30.8	67.3	20.0
	(1.8)	(4.0)	(1.7)	(2.2)	(5.0)	(2.1)	(3.1)	(6.6)	(3.0)
Community-based targeting	28.8	65.7	18.4	31.5	73.6	20.0	23.8	51.9	15.4
	(1.8)	(4.0)	(1.7)	(2.3)	(4.6)	(2.2)	(2.8)	(7.0)	(2.7)
Random targeting error	34.3	78.0	22.0	33.2	79.0	21.0	35.4	77.0	23.0
Number of households	653	143	510	426	91	335	227	52	175

Notes: Standard errors in parentheses.

Table A2: Four statistical targeting indices and their specifications

Indicators	Number of categories used	Multidim. Poverty	Poverty Scorecard	Asset Index	Econometric PMT
Total number of variables used		9	9	33	33
Demographic characteristics					
Share of HH-members at age 15 and below	2			x	x
Share of HH-members between age 16 and 60	2			x	x
Share of HH-members above 60	2			x	x
How many household members are 14-years-old or younger?	7		x		
HHHead is married	2			x	x
HHHead is not widowed	2			x	x
HHHead is male	2			x	x
Occupational choice					
HHH can read or write	2		x		
HHHead is literate	2			x	x
No one in the HH is literate	2	x			
Any HH member completed primary	2		x	x	x
Any HH member completed secondary	2			x	x
Any HH member completed tertiary	2			x	x
No HH member has completed five years of schooling	2	x			
HHH is not employed in agriculture	2			x	x
Any HH member is not employed in agriculture	2		x		
Dwelling characteristics					
HH uses running-water or good wells, any period	2	x		x	x
Drinking Water is changed at least every 2nd day	2			x	x
Water is not piped outside	2	x			
Sanitation not at the open field	2			x	x
Toilet arrangement	3		x		
number of rooms	n.A.			x	x
Roof is made of concrete, metal sheets, or tile	2			x	x
Wall is not made of ordinary mud or straw	2			x	x
Floor is made of cement	2	x		x	x
No electricity or solar panel	2	x			
Main energy source of lighting	4		x		
Cooking fuel is wood	2	x			
Asset possession at household level					
At least one cart	2			x	x
At least one plow	2			x	x
At least one bike	2			x	x
At least one mbike	2		x	x	x
At least one car	2			x	x
At least one radio	2			x	x
At least one tv	2		x	x	x
At least one fridge	2			x	x
At least one kitchen	2			x	x
HH owns no assets at all	2	x			
Livestock possession at household level					
At least one horse_donkey	2			x	x
At least one goat_sheep	2			x	x
At least one chicken	2			x	x
At least one bullock	2			x	x
At least one pig	2			x	x
Number of bulocks owned by HH head	4		x		
Other					
HH experienced at least one severe illness last month	2	x			
HHs belongs to ethnic minority group	2			x	x

Notes: The second column *Categories* specifies the number of categories of the variable. The majority of variables consists of indicator variables which only take on two values.

Table A3: Indicators and weights of the econometric PMT index

	Dependent variable: Eligible by consumption	
	Rural	Urban
Any HH-member with primary education	-0.040	0.011
Any HH-member with secondary education	-0.087	-0.021
Any HH-member with tertiary education	0.000	0.606*
Household head literate (incidence)	-0.047	-0.157
HH head occup. non-agric. (incidence)	0.015	0.191*
Ethnic minority (incidence)	0.015	-0.137
HH uses running-water or good wells, any period	0.062	-0.124
Drinking Water is changed at least every 2nd day	0.020	-0.226
Wastewater by cesspool, gutters or septic tank	0.156	-0.142
Concrete, metal sheets, or tile	-0.052	0.040
No ordinary mud or straw	-0.047	-0.093
Floor is made of cement	0.084	-0.100
Not at the open field	-0.035	-0.079
Number of rooms	0.007	0.012
At least one Cart	-0.086	0.118
At least one Plow	-0.107	-0.046
At least one Bike	-0.139	-0.201
At least one Motorbike	-0.001	0.008
At least one Car	0.060	-0.002
At least one Radio	-0.048	-0.030
At least one TV	-0.133	-0.117
At least one Fridge	0.000	0.212
At least one Kitchen	0.000	-0.166
At least one Horse or donkey	0.001	0.110
At least one Goat/Sheep	-0.171*	-0.171
At least one Chicken	-0.053	0.008
At least one Bullock	0.090	-0.131
At least one Pig	0.085	0.053
Share of HH members between 16 and 60 years	-0.508***	-0.470
Share of HH members above 60 years	-0.102	-0.030
HH head is married	-0.041	0.017
HH head is not widowed	0.256**	0.189
HH head is male	0.226	0.115
Observations	458	196
R^2	0.24	0.27
F-test (p-value)	0.00	0.00

Notes: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the household level. Regressions include community fixed effects. Regression sample includes all households surveyed in 2008.

Table A4: Indicators and weights of the asset index

	1st principal component factor loadings	
	Rural	Urban
Any HH-member with primary education	0.196	0.178
Any HH-member with secondary education	0.127	0.150
Any HH-member with tertiary education	0.120	-0.036
Household head literate (incidence)	0.000	-0.128
HH head occup. non-agric. (incidence)	-0.077	0.114
HHs belongs to ethnic minority group	-0.048	0.205
HH uses running-water or good wells, any period	-0.038	-0.115
Drinking Water is changed at least every 2nd day	0.023	0.032
Wastewater by cesspool, gutters or septic tank	-0.004	0.070
Concrete, metal sheets, or tile	0.201	0.373
No ordinary mud or straw	0.085	0.120
Floor is made of cement	0.193	-0.221
Not at the open field	0.154	-0.112
Number of rooms	0.270	-0.290
At least one Cart	0.296	0.067
At least one Plow	0.293	-0.182
Bicycle	0.195	-0.003
Motorbike	0.275	-0.148
At least one Car	0.029	-0.294
At least one Radio	0.202	0.113
At least one TV	0.221	-0.315
At least one Fridge	0.065	0.195
At least one Kitchen	0.000	0.278
At least one Horse or donkey	0.325	0.184
Goat or sheep	0.249	0.166
At least one Chicken	0.181	-0.010
Bullock	0.289	0.016
At least one Pig	0.079	-0.116
Share of HH-Members at age 16 to 60	0.043	0.203
Share of HH-Members above age 60	-0.167	0.105
HH head is married	0.109	0.171
HH head is not widowed	0.131	0.125
HH head is male	-0.093	0.097
Observations	426	229
Number of principal components	31	33

Notes: Weights are derived from a principal component analysis where all variables are first demeaned at the community level.

Table A5: The Poverty Scorecard Index

Original scorecard for Burkina Faso		*Scorecard adjusted for our study*	
Indicator	**Score**	**Indicator**	**Score**
1. How many household members are 14-years-old or younger?		1. How many household members are 15-years-old or younger?	
A. Six or more	0	A. Six or more	0
B. Five	5	B. Five	5
C. Four	6	C. Four	6
D. Three	10	D. Three	10
E. Two	13	E. Two	13
F. One	19	F. One	19
G. None	29	G. None	29
2. In what languages can the male head/spouse read and write?		2. HH head can read and/or write	
A. None, or no male head/spous	0	A. No	0
B. French only	4	B. Yes	4
C. A non-French language (regardless of French literacy)	5		
3. Has the female head/spouse completed first grade?		3. First grade completed by HH head	
A. No	0	A. No	0
B. No female head/spouse	0		
C. Yes	9	B. Yes	9
4. What is the main source of energy for lighting?		4. What is the main source of energy for lighting?	
A. Firewood, or other	0	A. Firewood, or other	0
B. Candles, kerosene, or LPG	4	B. Candles or oil lamp	4
C. Flashlight, or batteries	5	C. Flashlight	5
D. Electricity, or solar energy	8	D. Electricity, solar panel or battery	8
5. What toilet arrangement does the household have?		5. What toilet arrangement does the household have?	
A. No toilet arrangement, or other	0	A. Open field	0
B. Non-ventilated pit latrine	4	B. Latrine	4
C. Ventilated pit latrine, or flush to a septic tank	15	C. Ventilated latrine and flush toilet	15
6. Does the household own a television?		6. Does the household own a television?	
A. No	0	A. No	0
B. Yes	10	B. Yes	10
7. Does the household own a bed or a mattress?		7. Omitted	
A. No	0		
B. Yes	3		
8. Does the household own a scooter or a motorcycle?		8. Does the household own a scooter or a motorcycle?	
A. No	0	A. No	0
B. Yes	6	B. Yes	6
9. Have any household members, in their main occupation in the last seven days, worked in agriculture, animal husbandry, fishing, or forestry?		9. Is the primary occupation of the HH head in agriculture?	
A. Yes	0	A. Yes	0
B. No	8	B. No	8
10. How many head of cattle or other large animals does the household now own?		10. How many head of bullocks does the household head now own?	
A. None, or one	0	A. None, or one	0
B. Two	2	B. Two	2
C. Three to five	3	C. Three to five	3
D. Six or more	7	D. Six or more	7

Notes: This was retrieved from the following link on September 10, 2016: `http://www.progressoutofpoverty.org/country/burkina-faso`

Table A6: The Multidimensional Poverty Index

Original MPI			MPI adjusted for our study	
Indicator		**Weight**	**Indicator**	**Adjusted MPI Weight**
Education 0.333				
Years of Schooling	No household member has completed five years of schooling	0.167	No household member has completed five years of schooling	0.167
Child School Attendance	Any school-aged child is not attending school in years 1 to 8	0.167	No one is literate	0.167
Health 0.333				
Mortality	Any child has died in the family	0.167	Any severe-illness in the last month	0.333
Nutrition	Any adult or child for whom there is nutritional information is malnourished	0.167	-	
Standard of Living 0.333				
Electricity	The household has no electricity	0.056	The household has no electricity or solar panel	0.056
Sanitation	The households sanitation facility is not improved (according to the MDG guidelines), or it is improved but shared with other households	0.056	Water not piped outside	0.056
Water	The household does not have access to clean drinking water (according to the MDG guidelines) or clean water is more than 30 minutes walking from home.	0.056	Ordinary water source	0.056
Floor	The household has dirt, sand or dung floor	0.056	Floor is not cement	0.056
Cooking Fuel	The household cooks with dung, wood or charcoal.	0.056	Household cooks with wood	0.056
Assets	The household does not own more than one of: radio, TV, telephone, bike, motorbike or refrigerator, and does not own a car or truck.	0.056	HH owns no assets	0.056

Notes: The original MPI is from Alkire and Santos (2010)

Table A7: Simulations of sample mean targeting errors

				Commnity-based Targeting					
	MTE			10.0	15.0	20.0	25.0	30.0	34.3
		$\widetilde{MTE}$		10.1	15.0	20.0	24.9	29.8	34.0
			$\widehat{MTE}$	7.7	12.7	17.6	22.4	27.1	31.2
Asset index targeting	10.0	10.1	9.8	2.1	1.8	2.1	2.8	.	.
	15.0	15.0	13.8	0.9	1.0	1.4	2.0	2.9	.
	20.0	20.0	18.2	-0.4	0.1	0.6	1.2	1.9	2.7
	25.0	24.9	22.7	-2.0	-1.0	-0.3	0.3	1.0	1.6
	30.0	29.8	27.2	.	-2.4	-1.4	-0.6	0.1	0.7
	34.3	34.0	31.2	.	.	-2.3	-1.3	-0.6	0.0

Table A8: Differences between pairs of MTEs, Rural Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		-0.5	-0.9	-5.6*	3.3
MPI (#10)			-0.5	-5.2*	3.8
PCA (#33)				-4.7*	4.2*
ECON (#33)					8.9***

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the household level.

Table A9: Differences between pairs of MTEs, Urban Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		5.3	-2.6	-4.4	-3.5
MPI (#10)			-7.9*	-9.7**	-8.8*
PCA (#33)				-1.8	-0.9
ECON (#33)					0.9

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the household level.

Table A10: Differences between pairs of MTEs, Extremely poor households, Rural Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		2.8	5.6	-13.9	16.7
MPI (#10)			2.8	-16.7	13.9
PCA (#33)				-19.4*	11.1
ECON (#33)					30.6**

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors clustered at the household level.

Table A11: Differences between pairs of MTEs, Moderately poor households, Rural Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		-3.6	-7.3	-12.7	1.8
MPI (#10)			-3.6	-9.1	5.5
PCA (#33)				-5.5	9.1
ECON (#33)					14.5

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the household level.

Table A12: Differences between pairs of MTEs, Around median households, Rural Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		-2.2	-4.4	-3.9	-1.1
MPI (#10)			-2.2	-1.7	1.1
PCA (#33)				0.6	3.3
ECON (#33)					2.8

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the household level.

Table A13: Differences between pairs of MTEs, Affluent households, Rural Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		1.9	3.9	-3.2	5.8
MPI (#10)			1.9	-5.2	3.9
PCA (#33)				-7.1*	1.9
ECON (#33)					9.0*

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the household level.

Table A14: Differences between pairs of MTEs, Extremely poor households, Urban Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		-23.8	-47.6**	-38.1*	-42.9**
MPI (#10)			-23.8*	-14.3	-19.0
PCA (#33)				9.5	4.8
ECON (#33)					-4.8

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the household level.

Table A15: Differences between pairs of MTEs, Moderately poor households, Urban Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		35.5**	22.6	9.7	16.1
MPI (#10)			-12.9	-25.8*	-19.4
PCA (#33)				-12.9	-6.5
ECON (#33)					6.5

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the household level.

Table A16: Differences between pairs of MTEs, Around median households, Urban Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		3.2	-1.1	-7.5	-1.1
MPI (#10)			-4.3	-10.8*	-4.3
PCA (#33)				-6.5	-0.0
ECON (#33)					6.5

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the household level.

Table A17: Differences between pairs of MTEs, Affluent households, Urban Sector

	PSI (#10)	MPI (#10)	PCA (#33)	ECON (#33)	CBT
PSI (#10)		3.7	-2.4	2.4	-3.7
MPI (#10)			-6.1	-1.2	-7.3
PCA (#33)				4.9	-1.2
ECON (#33)					-6.1

Notes: Each cell gives the estimated difference in the mean targeting error when using the column instead of the row targeting method. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors clustered at the household level.