*Journal of African Economies,* 2020, Vol. 29, number 3, 271–305
doi: 10.1093/jafeco/ejz026
Advance Access Publication Date: 10 December 2019
Articles

OXFORD

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

**Michael Hillebrecht [1,2,3,*], Stefan Klonner [2], Noraogo A. Pacere [4] and Aurélia Souares [3]**

[1]German Society of International Cooperation (GIZ), Eschborn, Germany, [2]South Asia Institute, Department of Development Economics, Heidelberg University, Heidelberg, Germany, [3]Heidelberg Institute of Global Health, Medical Faculty, Heidelberg University, Heidelberg, Germany, [4]Centre de Recherche en Santé de Nouna, Nouna, Burkina Faso.

*Corresponding author: Michael Hillebrecht, German Society for International Cooperation (GIZ), Eschborn, Germany. Email: michael.hillebrecht@giz.de

## Abstract

Targeting of governmental welfare programmes in low-income countries commonly relies on statistical procedures involving household-level data, while smaller-scale programmes 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 programme benefits.

**Key words:** targeting, community-based targeting, welfare programs, poverty, proxy means testing.

**JEL classification:** I13, I38, O15

## 1. Introduction

Poverty reduction programmes designed to directly improve the well-being of the poor have become increasingly popular around the global South (Honorati et al., 2015). When a programme's purpose is to maximise poverty reduction under a limited budget, its success depends on the programme'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 programmes tends to rely on statistical procedures processing suitable proxies of households' means in a proxy neans test (PMT). Alternatively, targeting may be decentralised 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, Robertson et al., 2014, Schüring, 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 programme (Robertson et al., 2014). On the downside, decentralisation of political decision making may be susceptible to capture by local elites (Bardhan and Mookherjee, 2006, Conning and Kevane, 2002). In contrast, statistical targeting is the more objective and easier-to-replicate approach. With a more centralised organisational structure it has the potential to reduce 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 CBT 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 thirty-six villages and twenty-two semi-urban neighbourhoods 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 a score that 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 scalar score. The *Econometric PMT* employs a large number of indicators available in census data (Brown et al., 2018), and 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 (PCA) (Filmer and Pritchett, 2001). This resulting score 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' (MPI). Following Alkire and Santos (2010), in this approach all indicators are first transformed into binary deprivation indicators and the score 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 PMT 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%. 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 CBT only for very large transfer benefits, even in the rural areas. Hence, for the benefits usually encountered in welfare programmes in low-income countries, CBT 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 CBT and econometric targeting are available to the policy maker.

Within the vast economic literature on targeting of welfare programmes, 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 programme at a time and targeting accuracy is measured by the share of households meeting the programme'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 programme's eligibility rules.[2]

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 (Brown et al., 2018, Filmer and Scott, 2012, Grosh and Baker, 1995, Karlan and Thuysbaert, 2016, Klasen and Lange, 2014, Stoeffler et al., 2016). Within this comparative targeting accuracy literature one may distinguish

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 decentralisation (Banerjee et al., 2014, Galasso and Ravallion, 2005) and communities' poverty perceptions (Alatas et al., 2012, Kebede, 2009, Van Campenhout, 2007).

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 programmes in Malawi and Kenya and Castañeda (2005) for Columbia's SISBEN. We describe this latter programme in more detail in Section 2. The review by Coady et al. (2004a) summarises studies of 122 anti-poverty programmes in 48 countries.

between a first branch that compares various alternative statistical targeting methods with each other (Brown et al., 2018, Filmer and Scott, 2012, Grosh and Baker, 1995, Klasen and Lange, 2014) and a second branch whose subject is the comparative assessment of statistical and CBT. Among the latter, the seminal study is Alatas et al. (2012). These authors separately investigate the targeting accuracy of pure CBT, as in the present paper, and a hybrid method, where CBT 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 CBT 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 CBT 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 CBT 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 in villages and initially been known as rapid rural appraisals (RRAs) (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 (Sabates-Wheeler et al., 2015, Stoeffler et al., 2016), we find a high targeting accuracy and a superior cost-effectiveness of this method in urban neighbourhoods, which parallel 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%, Alatas et al. (2012) and Yamauchi (2010) primarily study rural areas of Indonesia, a middle-income country with rural poverty rates of around 20%. Our study area comprises rural as well as semi-urban areas with poverty rates of around 60% and 20%, 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 CBT 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 programmes

3 Another related article is Sabates-Wheeler et al. (2015), who compare the targeting accuracy of pure CBT 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.

with limited budgets in severely impoverished settings, where comparatively inexpensive CBTs are especially popular. Moreover, our exploration of CBT in semi-urban areas is of great practical importance since urbanisation rates are rapidly rising in Western Africa, from 9% in 1950 to 45% 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 CBT, 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 CBT. 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 programmes in Mexico (Progresa/Oportunidades), Colombia (Familias en Acción) and Chile (PASIS and SUF). National food-subsidisation programmes such as those in Indonesia and Egypt use statistical targeting as well (Ahmed and Bouis, 2002, Coady et al., 2004a). Statistical methods are also popular among small-scale poverty reduction programmes, 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 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 weights used for scoring. We shall discuss four eminent statistical methods, which we categorise along these lines, and contrast them with CBT.

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).

### 2.1 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 (Alatas et al., 2012, Filmer and Pritchett, 2001, Klasen and Lange, 2014). Weights are obtained from a regression of per capita consumption on the proxy-means variables with an additional sample 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 programme's purpose is to reduce consumption poverty, this method is motivated by the fact that the resulting score is the best linear predictor of household consumption given the information available in a census. Most of the recent comparative targeting accuracy literature studies this statistical targeting method, and the large-scale cash transfer programmes in Mexico (Progresa) and Indonesia (BLT) are two prominent applications.

### 2.2 Asset index

For the asset index weights are obtained from the joint distribution of the proxy-means variables themselves. Specifically, 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 asset index is frequently used by researchers 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. In this literature the index has also 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 20 years. In this system, eligibility for various social programmes relies on a wealth index with 13 proxy-means variables and PCA-based weights (Castañeda, 2005).

### 2.3 The PSI

In comparison to the two just-discussed 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 realisations 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 the Grameen Foundation and, lately, the non-governmental organisation Innovations for Poverty Action (IPA), where it is called Progress out of Poverty

Index (PPI). The index has also been used for targeting of anti-poverty interventions and is increasingly used in contexts other than microfinance, such as health and education (Alkire et al., 2015, Schreiner, 2015). According to IPA's 2014 report, the PSI is being used by more than 200 organisations for anti-poverty programmes around the global South. Among them are the Bangladesh Rural Advancement Committee, the Grameen Bank, the Ford Foundation and the International Finance Corporation (Innovations for Poverty Action, 2014). Customised scorecards for 46 countries are available as of 2016. The selection of indicators is based on 'statistics and judgement', 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.

### 2.4 The Multidimensional Poverty Index

In this approach all indicators are first transformed into binary deprivation indicators and the score 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 judgements 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 PMT 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 programmes (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 Poverty Line scorecard for India (Alkire and Seth, 2013, Azevedo and Robles, 2013, Robano and Smith, 2013, Thomas et al., 2009). 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.

### 2.4 Community-based targeting

In CBT 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 RRA. RRAs have become popular in the late 1970s because of a growing discontent with statistical poverty assessments and, in particular, their high costs (Chambers, 1994a). 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 asset creation programmes for ultra-poor households funded by the Consultative Group to Assist the Poor. Karlan and Thuysbaert (2016) analyse two such programmes in Honduras and Peru. Banerjee et al. (2007) investigate CBTs within the context of a similar asset creation programme in rural India. In their cross-country analysis of targeted anti-poverty interventions, Coady et al. (2004a) state that CBT is similarly often used as proxy-means testing, equally popular on all continents and especially wide-spread in very poor countries.

To the best of our knowledge, there is no structured summary of the procedural details of community wealth rankings and CBT 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 the fact 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 summarised 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 programme 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% 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% 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% 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% 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

7 The national definition of 'urban locality' is a 'locality with 10,000 inhabitants or more and with sufficient socio-economic and administrative infrastructures' (UN-DESA, 2006).

**Table 1:** Community Wealth Rankings and CBT: Procedural Detailstab:lit˙cwr

| | 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 survey data by a local government-funded health research centre, the 'Centre de Recherche en Santé de Nouna', and Heidelberg University's Institute of Public Health has been ongoing in forty-one of the department's fifty-nine villages as well as in Nouna Town since 1993 (De Allegri, 2006).[8]

With the objective of developing a nationwide 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, 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 50% discount on the premium, essentially a voucher, was offered to the poorest quintile of households in each village and urban neighbourhood. 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 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 thirty-six villages of the study area as well as in twenty-two neighbourhoods of Nouna Town.[9] In what follows, we shall refer to both villages and semi-urban neighbourhoods 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 were 'has insufficient food' and 'has nothing' (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 algorithm: first, households ranked among the $m$ poorest by all three informants are automatically eligible (about 40% of beneficiary households; see Table 2). Second, all

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

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

households ranked among the $m$ poorest by exactly two informants are included if the resulting number of beneficiaries does not exceed $m$, and, in a consultation among the key informants, the remaining beneficiary households are chosen from the set of households ranked among the $m$ poorest by exactly one informant previously. Otherwise, only a subset of the households ranked among the $m$ poorest by exactly two informants is chosen in a consultation among the key informants. This occurred in only four of the fifty-eight communities. In thirteen communities, the number of households ranked among the $m$ poorest by at least two informants equaled $m$ precisely 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% of households in each community were targeted. In rural and urban communities alike, there is a high 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% of beneficiary households (8% relative to 20%). Educational attainments of CBT informants reflect education and literacy levels among household heads in Nouna, where literacy is about 20% 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 PCA. 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 (Odusola et al., 2017). For a measure of ethnic heterogeneity, we calculate an index commonly known as ethno-linguistic fractionalisation, 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 this measure, urban neighbourhoods are twice as diverse as the villages.

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.

**Table 2:** CBT and Community Characteristicstab:summ˙cwr

| | Pooled | Rural | Semi-urban |
|---|---|---|---|
| *CBT* | | | |
| 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 fractionalisation 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 PCA from a total of twenty-five assets.

For the merged data set 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 10% are 55 years 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% of household heads reporting an illness during the month preceding the interview. Livestock possession is widespread in rural and semi-urban

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% more frequent than in the Nouna department.

communities alike. Bicycles are the dominant means of transportation and close to a quarter of households owns a motorbike.

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 years and older. It records market purchases with two recall periods, 1 and 6 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% of the value of food consumption and 8% and 23% 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%.[12]

Descriptive statistics for the resulting MPC measure are set out in Table 3. Per capita consumption is not only considerably higher but also much more dispersed in the urban neighbourhoods. The consumption levels underlying the figures in Table 3 very likely understate actual consumption for two reasons. First, a short recall period of 1 week rather than 1 month as in our survey for high frequency consumption items has been found to result in consumption figures that are 25% and 13% 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

11 In the 2014 Burkina LSMS, the recalls for the two categories are 1 week and 3 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 1,800 (1,400) for adult and 900 (700) for adolescent rural (urban) inhabitants. These thresholds very likely understate actual levels of home-produced consumption by 30% to 40% 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.

**Table 3:** Household Survey Summary Statisticstab:summ˙nhhs

| | Pooled | Rural | Semi-urban |
|---|---|---|---|
| *CBT* | | | |
| 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) years. Ethnic minority is a dummy variable equal to one if the household head belongs to the community's smallest ethnic group.

consumption by around 30% relative to a detailed questionnaire, which is used, for example, in LSMSs (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% 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 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]

### 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

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

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

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.

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.

**Table 4:** Targeting Methodstab:targ˙sets

| 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 PMT | 44 | None | OLS |
| PCA | | | | |
| $T^{PCA}$ | Asset index | 44 | None | PCA |
| Scorecard | | | | |
| $T^{PSI}$ | PSI | 9 | Ordered categorical | hybrid |
| Counting of deprivations | | | | |
| $T^{MPI}$ | MPI | 9 | Binary | Global MPI |
| *CBT* | | | | |
| $T^{CBT}$ | 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; 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.

negative) occurs when any of the consumption-poorest households is not targeted by the targeting method under consideration. Conversely, non-poor households that 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%, 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 $\delta$ 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.

**Table 5:** Targeting Errors (in Percent)tab:mte

| | 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) |
| *PSI* | 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) |
| *MPI* | 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) |
| *CBT* | 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 that are erroneously classified as either poor or non-poor. The exclusion error is defined as the proportion of households that are consumption poor but not targeted by the targeting method under consideration. The inclusion error is defined as the proportion of households that are consumption non-poor but targeted by the targeting method under consideration.

## 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 2%. The econometric PMT has by far the lowest MTE with just 23.5% and 22.9% 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 PSI and the MPI deliver only marginal and statistically insignificant (at the 5% 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 5% level, however, while the econometric PMT performs significantly better (see Appendix Table A8 for significance levels in pairwise comparisons of targeting methods).

Regarding CBT, 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%, which is a statistically significant improvement at the 1% 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 significant 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 (Kebede, 2009, 2013, for Honduras and Peru). In contrast, in our setting, CBT achieves a high accuracy in the semi-urban neighbourhoods (columns 7 to 9). As the second-best performing method, it reduces the random targeting error by 33% (significant at 1%), 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% over random targeting. The improvement is greater than the 12% reported by the same authors or the 17.5% 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 neighbourhoods. This double difference of five percentage points is not statistically significant in their estimations, however.

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 6we 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 CBT 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%, the affluent and around median expenditure brackets roughly contain the first and second 39% 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]

Only the econometric PMT performs consistently well across sectors and expenditure classes. In the semi-urban neighbourhoods, CBT and the asset index share the favourable feature of identifying correctly households in the lowest and highest expenditure brackets, where they improve on random targeting by around 60%. 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 neighbourhoods 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. Towards this, we decompose CBT outcomes into an observable and an unobservable 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

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:** Targeting Errors by Consumption Expenditure Classestab:wmte

| | 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) |
| *PSI* | 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) |
| *MPI* | 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) |
| *CBT* | 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 CBT 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 Characteristicstab:com˙char

| Dependent variable: | Targeting error | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | PMT | Asset index | PSI | MPI | CBT | | | |
| | 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.

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.

Consistent with the 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

As a second potential 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 favouritism 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).

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 5% (or one percentage point) in urban neighbourhoods, while the same probability decreases by just a little more than 1% for rural households.

**Table 8:** Determinants of CBT Selectiontab:det˙cbt

| | 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 Analysistab:cba

| Method *A* delivers higher benefits than method *B* for programme 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 CBT we consider the CBT implementation and data processing costs of $1.19 per eligible household throughout.

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% (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 favouritism 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 years and older has an 80% (or seventeen 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 10% 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 the 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. Towards 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 twenty-two urban and thirty-six rural communities, we obtain MTEs of 17.9% and 19.0%, respectively, with standard errors of 2.5% and 1.9%. 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 PMT and partly also the asset index over CBT 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 programme'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 programmes with relatively small budgets, while the opposite holds for relatively generous transfer programmes. This is precisely what Alatas et al. (2012) find in the 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, 2003; ; Klasen and Lange, 2014).

We use cost information from the 2009 CBT 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 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 programme.[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 (Alatas et al., 2012, Ravallion, 2009, 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 synchronised 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 programme by two criteria. First, to identify 'socially valuable' beneficiaries, we rely on the stated purpose of the community targeting exercise, to identify the 20% 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 programme 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 that 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 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]

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 nationwide programme.

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 programme.

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%. 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.

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 programme 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 programme that 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 programme per eligible household will be denoted by

$$C = t + TC,$$

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 programme involving CBT with the econometric PMT and the asset index,. The econometric PMT is always the most cost-effective choice for programmes 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 programmes 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), CBT is more cost-effective than econometric targeting for programme costs of up to \$0.78 and \$41.70 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, 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.

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 programme budgets, of about $60 and $260 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 CBT 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 programme and, by construction, 20% 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 programme. This implies that a universal programme is always most cost-effective for very small programme budgets because no fixed costs accrue. For a uniform transfer, a universal programme 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 CBT has indeed been the most cost-effective method for the subsidy intervention fielded in Nouna.

## 7. Discussion

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

In the following we shall summarise 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. We have shown 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 programme 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 decentralised expert assessment, CBT 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 programme budgets, and there is little scope for less expensive statistical methods, such as the asset index or scorecards. For the budget spent in the subsidy campaign considered here, participatory targeting is clearly the method of choice. But even for more generous programmes encountered in practice, CBT may dominate statistical targeting in this African context. To illustrate, first, we consider the Indonesian unconditional cash transfer programme investigated by Alatas et al. (2012). Per individual in an eligible household, the programme's annual cash transfer equals about 1.4% 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 programme of this size exceeds the threshold of $46 below which CBT 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 programme geared at improving schooling and access to health care for children in poor families, the Nahouri Cash Transfers Pilot Project, employed sophisticated econometric PMT (Akresh et al., 2014). Implemented between 2008 and 2010, it involved transfers of about $160 per targeted household over the course of 2 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 decentralisation 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 are not readily available. Given that a general census is typically not carried out more often than every 10 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 CBT might be particularly suited for quickly evolving environments. In our study area, for example, the CBT exercise has been carried out three times between 2007 and 2011.

Second, the participative procedure of CBT likely generates additional value. 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 decentralised targeting methods (Alatas et al., 2012, Robertson et al., 2014, Schüring, 2014). Savadogo (2017) confirms this for our CBT intervention, where more than 85% of households have approved 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, CBT appears to be well suited for translating deprivations in the space of capabilities into targeting outcomes.

## Supplementary material

Supplementary material is available at *Journal of African Economies* online.

## References

Adams A. M., Evans T. G., Mohammed R., Farnsworth J. (1997) ' Socioeconomic stratification by wealth ranking: Is it valid?' *World Development*, 25 (7): 1165–72.

Ahmed A. U., Bouis H. E. (2002) ' Weighing what's practical: Proxy means tests for targeting food subsidies in Egypt', *Food Policy*, 27 (5): 519–40.

Akresh R., De Walque D., Kazianga H. (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*. National Bureau of Economic Research, Inc, University of Chicago Press, Chicago, pp. 113–35.

Alatas V., Banerjee A., Hanna R. (2012) ' Targeting the poor: Evidence from a field experiment in Indonesia', *American Economic Review*, 102 (4): 1206–40.

Alatas V., Banerjee A., Hanna R., Olken B. A., Purnamasari R., Wai-Poi M. (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., Banerjee A., Hanna R., Olken B. A., Purnamasari R., Wai-Poi M. (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., Baqir R., Easterly W. (1999) ' Public goods and ethnic divisions', *The Quarterly Journal of Economics*, 114 (4): 1243–84.

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

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

Alkire S., Santos M. E. (2014) ' Measuring acute poverty in the developing world: Robustness and scope of the multidimensional poverty index', *World Development*, 59 (7): 251–74.

Alkire S., Seth S. (2013) ' Selecting a targeting method to identify BPL households in India', *Social Indicators Research*, 112 (2): 417–46.

Azevedo V., Robles M. (2013) ' Multidimensional targeting: Identifying beneficiaries of conditional cash transfer programs', *Social Indicators Research*, 112 (2): 447–75.

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

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

Bardhan P., Mookherjee D. (2006) ' Pro-poor targeting and accountability of local governments in West Bengal', *Journal of Development Economics*, 79 (2): 303–27.

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

Brown C., Ravallion M., Van De Walle D. (2018) ' A poor means test? Econometric targeting in Africa', *Journal of Development Economics*, 134: 109–24.

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): 407–28.

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

Chambers R. (1994a) ' Participatory rural appraisal (PRA): Analysis of experience', *World Development*, 22(9): 1253–68.

Chambers R. (1994b) ' The origins and practice of participatory rural appraisal', *World Development*, 22 (7): 953–69.

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

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

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

Conning J., Kevane M. (2002) ' Community-based targeting mechanisms for social safety nets: A critical review', *World Development*, 30 (3): 375–94.

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', in *Ph.D. thesis*. Heidelberg: Ruprecht-Karls-Universität Heidelberg.

Deaton A., Kozel V. (2005) ' Data and dogma: The great Indian poverty debate', *The World Bank Research Observer*, 20 (2): 177–99.

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

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

Devereux S., Sharp K. (2006) ' Trends in poverty and destitution in Wollo, Ethiopia', *The Journal of Development Studies*, 42 (5): 592–610.

Filmer D., Pritchett L. H. (2001) ' Estimating wealth effects without expenditure data—Or tears: An application to educational enrollments in states of India', *Demography*, 38 (1): 115–32.

Filmer D., Scott K. (2012) ' Assessing asset indices', *Demography*, 49 (1): 359–92.

Fink G., Robyn P. J., Sié A., Sauerborn R. (2013) ' Does health insurance improve health: Evidence from a randomized community-based insurance rollout in rural Burkina Faso', *Journal of Health Economics*, 32 (6): 1043–56.

Foster J., Greer J., Thorbecke E. (1984) ' A class of decomposable poverty measures', *Econometrica*, 52 (3): 761–6.

Galasso E., Ravallion M. (2005) ' Decentralized targeting of an antipoverty program', *Journal of Public Economics*, 89 (4): 705–27.

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

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

Hargreaves J. R., Morison L. A., Gear J. S., Makhubele M. B., Porter J. D., Busza J., Watts C., Kim J. C., Pronyk P. M. (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): 212–29.

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

Howe L. D., Hargreaves J. R., Gabrysch S., Huttly S. R. (2009) ' Is the wealth index a proxy for consumption expenditure? A systematic review', *Journal of Epidemiology and Community Health*, 63 (11): 871–7.

Innovations for Poverty Action (2014) *2014 Global report on poverty measurement with the progress out of poverty index*. Washington, DC: Grameen Foundation, 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): 1–23.

Karlan D., Thuysbaert B. (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): 1731–46.

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

Lietz H., Lingani M., Sié A., Sauerborn R., Souares A., Tozan Y. (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., Chambers R. (2005) ' Reversing the paradigm: Quantification, participatory methods and pro-poor impact assessment', *Journal of International Development*, 17 (2): 271–98.

McGee R. (2004) ' Constructing Poverty Trends in Uganda: A Multidisciplinary Perspective', *Development and Change*, 35 (3): 499–523.

Niehaus P., Atanassova A. (2013) ' Targeting with agents', *American Economic Journal: Economic Policy*, 5 (1): 206–38.

Odusola A., Lal R., Dhilwayo R., Neuhaus J., Sabo I. (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*. New York, USA: United Nations Development Programme, pp. 317–41.

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*. Oxford: Oxford University Press, pp. 453–467.

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

Ravallion M. (2009) ' How relevant is targeting to the success of an antipoverty program?' *The World Bank Research Observer*, 24 (2): 205–31.

Robano V., Smith S. C. (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., Mushati P., Skovdal M., Eaton J. W., Makoni J. C., Crea T., Mavise G., Dumba L., Schumacher C., Sherr L., Nyamukapa C., Gregson S. (2014) ' Involving communities in the targeting of cash transfer programs for vulnerable children: Opportunities and challenges', *World Development*, 54 (1): 325–37.

Sabates-Wheeler R., Hurrell A., Devereux S. (2015) ' Targeting social transfer programmes: Comparing design and implementation errors across alternative mechanisms', *Journal of International Development*, 27(8): 1521–45.

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

Savadogo G., Souares A., Sie A., Parmar D., Bibeau G., Sauerborn R. (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): 537–69.

Schüring E. (2014) ' Preferences for community-based targeting—Field experimental evidence from Zambia', *World Development*, 54 (3): 360–73.

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, Handbook of Development Economics I*. Elsevier, pp. 9–26.

Shaffer P. (1998) ' Gender, poverty and deprivation: Evidence from the Republic of Guinea', *World Development*, 26 (12): 2119–35.

Skoufias E., Davis B., De La Vega S. (2001) ' Targeting the poor in Mexico: An evaluation of the selection of households into PROGRESA', *World Development*, 29 (10): 1769–84.

Souares A., Savadogo G., Dong H., Parmar D., Sie A., Sauerborn R. (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): 363–8.

Stoeffler Q., Mills B., del Ninno C. (2016) ' Reaching the poor: Cash transfer program targeting in Cameroon', *World Development*, 83 (1): 244–63.

Takasaki Y., Barham B. L., Coomes O. T. (2000) ' Rapid rural appraisal in humid tropical forests: An asset possession-based approach and validation methods for wealth assessment among forest peasant households', *World Development*, 28 (11): 1961–77.

Temu A. E., Due J. M. (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., Muradian R., De Groot G., De Ruijter A. (2009) ' Multidimensional poverty and identification of poor households: A case from Kerala, India', *Journal of Human Development and Capabilities*, 10 (2): 237–57.

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.

UN-DESA. (2006) ' World urbanization prospects: The 2005 revision', report No. 216, 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): 406–38.

Wagstaff A., Watanabe N. (2003) ' What difference does the choice of SES make in health inequality measurement?' *Health Economics*, 12 (10): 885–90.

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 Indonesia's IDT program', *Economic Development and Cultural Change*, 59(1): 95–147.