



Regular article

Perceived need and measured well-being: How well do subjective rankings capture relative poverty?☆

Immacolata Ranucci ^a, Kibrom A. Abay ^b, Luca Tiberti ^{a,c}

^a Department of Economics and Management, University of Florence, Via delle Pandette 32, 50127 Firenze (FI), Italy

^b International Food Policy Research Institute, 1201 Eye St., NW, Washington, DC 20005-3915, USA

^c Partnership for Economic Policy, Duvvill Campus, Kasarani, Nairobi, 30772-00100, Kenya

ARTICLE INFO

Dataset link: <https://doi.org/10.5281/zenodo.2768864>

JEL classification:

I32
D63
D83
O12

Keywords:

Subjective poverty rankings
Household welfare
Community-based targeting
Ethiopia

ABSTRACT

Subjective well-being rankings are widely used to target social protection programs, yet their ability to capture relative welfare and wealth remains debated. This study compares three commonly used subjective ranking methods: self-, peer-, and elite-based poverty rankings, against consumption- and wealth-based measures using a large household survey in Ethiopia, where about 20 households per village were ranked from most to least “needy” by themselves, their peers, and community leaders. We assess concordance between subjective and conventional welfare rankings and explore sources of divergence. Subjective rankings align more with relative wealth than consumption and with total rather than per capita welfare, suggesting they overlook household composition. Among the three subjective ranking methods, elite-based rankings best capture conventional measures, although overall concordance between subjective and welfare-based rankings remains weak. Subjective rankings also better reflect relative deprivation among households exposed to covariate shocks. A composite index combining all three modestly improves concordance with standard metrics. We find suggestive evidence that information asymmetry, favoritism, and welfare dynamics partly explain discrepancies between subjective rankings and conventional welfare measures. These findings offer important insights to improve targeting in data-scarce settings.

1. Introduction

There is a growing consensus that people care not only about their absolute deprivation, but also about their relative income and social position within their communities and reference groups (Ravallion and Lokshin, 2010; Alkire and Foster, 2011; Ravallion and Chen, 2011, 2019; Ravallion, 2020; Bontan and Perez-Truglia, 2022; Hvidberg et al., 2023). Conventional survey-based welfare measures typically fail to capture non-monetary dimensions of poverty and deprivation, such as social exclusion and vulnerability (Decerf, 2023; Ligon and

Schechter, 2003). To address these limitations, these measures are usually combined with (or complemented by) subjective and self-assessed measures (Ravallion and Lokshin, 2002, 2010; Beegle et al., 2012; Ravallion et al., 2016; Ravallion, 2020). For this purpose, subjective poverty rankings (SPR) based on (i) self-ranking, (ii) peers’ ranking, and (iii) community leaders’ (elites’) ranking are frequently used for screening and targeting welfare and social protection programs due to their ability to leverage local information that is otherwise unavailable to program implementers (Rai, 2002; Alderman, 2002; Coady et al., 2004; Alatas et al., 2016b; Trachtman et al., 2025).¹ In many cases,

☆ We thank Carly Trachtman, Caitlin Brown, two anonymous reviewers, and the editor (Karen Macours) for their feedback on an earlier version of this paper. We also thank participants at the Welfare & Policy Conference (held in Bordeaux), the ECINEQ conference (held in Washington, DC), the SITES Young Economists Workshop (held in Prato), and the SITES X Annual Conference (held in Rome) for their insightful comments. We also thank Abdelkrim Araar for providing us with the Stata codes to run additional statistical tests (on the difference between Spearman rank correlation measures). We thank Gillian Hollerich for proofreading the manuscript. This work benefited from funding by the CGIAR Science Program on Policy Innovations. We would like to thank all funders who supported this research through their contributions to the CGIAR Trust Fund. The Community Based Experiment (CBT) used in this study was registered at the American Economics Association’s RCT registry (AEARCTR-0012677). Details on the CBT experiment can be found in Abay et al. (2026). The study received IRB approval from the IFPRI IRB (#00007490).

* Corresponding author.

E-mail addresses: immacolata.ranucci@unifi.it (I. Ranucci), k.abay@cgiar.org (K.A. Abay), luca.tiberti@unifi.it (L. Tiberti).

¹ Subjective rankings are usually positively correlated with survey-based poverty indicators, such as assets, income, or expenditure (Alatas et al., 2016a; Alix-Garcia et al., 2021; Hussam et al., 2022; Trachtman et al., 2025) and should ideally contribute to reducing the exclusion error, thanks to the more comprehensive knowledge of households’ conditions held by the household itself, peers, or community leaders (Alderman, 2002).

these methods are widely used to complement additional targeting methods such as geographic targeting or proxy means tests,² although subjective rankings are sometimes used as substitutes when conventional survey-based measures of welfare are not available. Community members and local leaders can help to determine who benefits from credit schemes (Vera-Cossio, 2022) and agricultural subsidy programs (Basurto et al., 2020). Beyond their use in screening and targeting of welfare and subsidy programs, subjective and relative poverty rankings contain important information that can shape respondents' life satisfaction and happiness (Ferrer-i Carbonell, 2005; Luttmer, 2005),³ preferences for redistribution (Cruces et al., 2013; Karadja et al., 2017; Hvidberg et al., 2023), and support for political leaders and institutions (Kosec and Mo, 2024).

Building on these trends, policy makers and program implementers often rely on local information to target poverty reduction programs. In practice, this is operationalized by eliciting relative rankings of potential beneficiaries—through self-assessment, peer evaluations, or community leaders' judgments.⁴ Understanding how well, and under what conditions, these subjective rankings capture standard welfare indicators is therefore crucial. However, the existing literature remains inconclusive about the extent to which subjective measures capture survey-based measures of welfare and wealth (Ravallion and Chen, 2011; Ravallion, 2012; Ravallion et al., 2016; Hvidberg et al., 2023; Dupas et al., 2024; Aiken et al., 2025). Similarly, which subjective rankings are more effective in capturing conventional welfare and wealth measures under various socioeconomic conditions and circumstances remains an important empirical question, especially in contexts characterized by data scarcity and shocks such as conflict or drought. Studies characterizing self-ranking and peer-ranking are especially limited (Cruces et al., 2013; Alatas et al., 2016b; Hvidberg et al., 2023; Dupas et al., 2024), and it remains poorly understood whether these rankings capture long-term or transient welfare and economic conditions (Trachtman et al., 2025), and how best these subjective measures respond to covariate and idiosyncratic shocks (Hvidberg et al., 2023). For example, identifying whether and which of these subjective ranking methods can serve in data-scarce environments or amid dynamically evolving shocks such as conflict and drought can inform effective targeting and response to these shocks. Finally, whether the different subjective rankings complement or substitute each other remains underexplored. If these ranking methods provide complementary information, aggregating these ranking methods can improve their performance. On the other hand, if these ranking methods are substitutes for each other, administering the least costly method can conserve resources, especially given the important differences in the cost of administering these methods. To the best of our knowledge, there is no comprehensive systematic comparative assessment of self, peers', and elites' rankings aimed at assessing their potential to capture conventional welfare and wealth outcomes, and the factors associated with

their discordance with these conventional measures of economic well-being. This type of comprehensive comparison is important because administering these different ranking methods involves varying levels of effort and cost.

To address these gaps, we administered these subjective ranking methods in a large household survey in Ethiopia, where subjective rankings of approximately 20 households per village were elicited (from neediest to least needy, meaning from poorest to richest) through self, peers', and elites' assessments. Following previous practices, we asked respondents to rank their own household as well as their peers, while also embedding a similar exercise in a comprehensive community survey targeting community leaders (elites), who assumed various responsibilities in each village. In contrast to most previous work, which either focuses on one type of ranking or lacks benchmark comparisons, we directly compare the concordance among these three methods and conventional expenditure- and asset-based measures. Beyond comparing the relative performance (in terms of capturing conventional welfare measures) of the three subjective rankings, we offer alternative ways to aggregate these three rankings into a single composite measure that proves to modestly reduce the discordance with welfare and wealth metrics.

We document a series of empirical patterns that are worth noting. First, although the overall concordance between the three subjective ranking methods and conventional welfare and wealth measures appears not to be strong, these subjective rankings better reflect wealth-based rankings than consumption-based rankings, and they are more concordant with total household welfare than per capita welfare. Second, among the three subjective measures, elite-based ranking performs better than the rest in terms of capturing conventional welfare and wealth. Third, the performance of the three ranking measures varies across contexts (rural and urban areas, as well as across highlands and pastoralist communities) and in the presence and absence of covariate shocks. Interestingly, these subjective ranking methods better capture relative deprivation among households exposed to covariate shocks such as conflict and drought. We also find that these subjective ranking methods perform relatively poorly in capturing relative welfare and wealth in pastoral communities, where ranking and targeting practices face major challenges arising from social and cultural norms (Sabates-Wheeler et al., 2013; Lind et al., 2022). We also discuss potential drivers of discordance between subjective and benchmark rankings, including information asymmetry, favoritism (in peers' and elites' rankings), and the dynamic evolution of households' economic conditions. Finally, we demonstrate that aggregating multiple subjective rankings results in a composite measure that, only modestly, better reflects conventional metrics than individual rankings.

Our findings offer important and practical lessons on the potential and pitfalls of alternative measures of subjective and relative well-being. For example, the evidence that these subjective rankings are better correlated with conventional relative measures during shocks (conflict and drought) is encouraging and confirms their potential and rationale for deploying them to identify those affected by covariate shocks. This is particularly important in contexts where detailed data on households' welfare are not available, as well as in contexts where conventional data-driven approaches such as proxy means tests perform poorly, such as in Africa (Brown et al., 2018). Compared to those approaches that require detailed consumption and asset measures, these subjective methods can be implemented cost-effectively (Hillebrecht et al., 2023, 2020).⁵ The better performance of elite-based ranking compared to the other methods implies that some types of information may not be freely available to all community members. The lower implementation cost of elite-based ranking supports its practical relevance.

² For example, in Ethiopia, the national Productive Safety Net Program combines geographic targeting and community-based targeting, which involves ranking of households through community leaders and community members (Gilligan et al., 2009; Berhane et al., 2014). Similarly, Alatas et al. (2016b) show that self-targeting enhances the overall efficiency of social protection targeting in Indonesia.

³ For example, Ferrer-i Carbonell (2005) note that "the income of the reference group is about as important as the own income for individual happiness". Another strand of literature argues that positional concerns may be more pronounced in high-income countries and households while such concerns get weaker in poor societies (Kingdon and Knight, 2007; Akay and Martinsson, 2011; Akay et al., 2012).

⁴ Despite their similarity in generating rankings of households and individuals, decision-making processes in peer or self-ranking and community-based targeting (CBT) vary. For example, in many cases, decision-making in CBT builds on collective consensus among committee members, while self- and peer-ranking involves individual decisions and exercises.

⁵ However, Aiken et al. (2025) find that both proxy means tests and phone-based targeting are more cost-effective than community-based targeting in Bangladesh. But these types of unconventional sources of data are not widely available to support targeting.

On the other hand, the generally weak concordance between subjective measures and conventional welfare and wealth-based ranking suggests that targeting based on these subjective measures would be prone to typical inclusion and exclusion errors.

This study makes several distinct contributions to the literature on subjective welfare assessments. First, we contribute to the empirical literature comparing alternative subjective ranking methods by providing a relatively more comprehensive comparison of self-, peer-, and elite-based rankings, benchmarked against conventional welfare and wealth indicators. Our results are consistent with earlier evidence documenting limited agreement between community-based rankings and survey-based welfare measures (Trachtman et al., 2025; Dupas et al., 2024; Aiken et al., 2025). While most previous studies compare survey-based welfare with a specific type of subjective ranking, we compare self-, peer-, and elite-based rankings within the same setting, helping identify which method performs better relative to its implementation cost. For example, while the peer and self-ranking methods entail visiting all households in villages to create complete rankings, the elite-based ranking can be constructed simply by visiting a small group of knowledgeable “elites”. We further show that aggregating information from multiple subjective sources modestly reduces the gap with benchmarks more than relying on any single measure, suggesting that these subjective rankings capture distinct dimensions of households’ latent deprivation.

Second, we add more nuance to the literature by documenting how the degree of concordance between subjective rankings and conventional welfare and wealth measures varies across contexts — for example, in the presence and absence of covariate shocks and across rural and urban areas — and by examining whether observed patterns are consistent with differences in information asymmetry and network structures (Alatas et al., 2016a; Trachtman et al., 2025; Bloch and Olckers, 2021). We find suggestive evidence that information asymmetry and favoritism may explain some of the discrepancies between subjective rankings and conventional welfare measures. More specifically, we find evidence of favoritism in the context of peer and elite ranking (Galasso and Ravallion, 2005; Ravallion, 2008; Alatas et al., 2012; Bloch and Olckers, 2021; Bardhan and Mookherjee, 2005; Alatas et al., 2019; Basurto et al., 2020; Conning and Kevane, 2002; Vera-Cossio, 2022), and a “race to the middle” effect in the case of self-ranking, whereby poorer individuals overstate their status while wealthier ones understate theirs (Dupas et al., 2024; Ghiglini and Goyal, 2010; Hvidberg et al., 2021). Finally, our study contributes to the debate on the veracity of relative social and economic positions of households as well as potential (mis)perceptions that may shape respondents’ well-being, preferences, and decisions (Ferrer-i Carbonell, 2005; Luttmer, 2005; Cruces et al., 2013; Karadja et al., 2017; Hvidberg et al., 2023; Kosec and Mo, 2024). As demonstrated by Karadja et al. (2017), correcting these misperceptions can generate meaningful changes in respondents’ preferences and decisions.

The remainder of this paper is structured as follows. Section 2 describes the data we used, defines the main ranking measures, and shows some descriptive statistics. Section 3 analyzes the concordance between SPR and welfare-based rankings. Section 4 analyzes the sources of discrepancies by different community- and household-level characteristics. Finally, Section 5 concludes the study.

2. Data, ranking methods, and descriptive statistics

2.1. Context and data sources

This study uses a unique and large survey in Ethiopia that covers most of the country’s regions. The subjective poverty ranking (SPR) elicitation exercises were only embedded in a follow-up survey (conducted in 2023) of baseline data collected in 2019. The 2019 household survey focuses on regions and districts included in USAID’s Feed the Future (FtF) Ethiopia Zone of Influence (ZOI). The FtF-ZOI in Ethiopia

represents the target geography for assessing the impact of the broader FtF program of the United States’ (US) government. The FtF program in Ethiopia is the US government’s flagship food security and agricultural development initiative, which aims to increase agricultural productivity, expand markets and income for smallholder farmers, and ultimately build resilience of smallholder farmers. The FtF program in Ethiopia was implemented through several initiatives and partnerships with the government of Ethiopia, international development partners, and the private sector.⁶ The FtF-ZOI in Ethiopia covered 132 *woredas* (districts) in six regions of Ethiopia: Afar, Amhara, Oromia, Somali, Tigray, and Southern Nations, Nationalities and Peoples (SNNP).⁷ The 2019 survey was designed to represent the population in the FtF-ZOI area. The survey adopted a multi-stage cluster sampling design. From each of the 132 *woredas*, two enumeration areas (EAs) were selected using probability proportional to size. EAs correspond to villages within a *kebele*, the lowest administrative unit in Ethiopia, which typically comprises 150–200 households. Then, from each of the 264 EAs, 20 households were randomly selected for interviews.

The 2019 survey was then used as a baseline for an incentivized community-based targeting (CBT) experiment and follow-up surveys in 2023, both at the household and community level. Of the 264 EAs included in the 2019 baseline survey, only 180 EAs were identified as accessible for the 2023 survey.⁸ Both the 2019 and 2023 household-level surveys included detailed food and non-food expenditure modules that were used to construct a within-village ranking of households according to their consumption levels. Total expenditure is obtained as the sum of food and non-food expenditures. Temporally deflated expenditures from 2019 and 2023 are expressed in 2016 prices. As an additional survey-based measure of relative poverty, we also use an asset index based on ownership of assets.

The CBT was implemented as a randomized experiment (Abay et al., 2026) during the community survey, 1–2 months before the household survey. Following targeting practices in Ethiopia, including for the Productive Safety Net Program (PSNP), six community leaders (“elites”) were selected from each community, representing diverse roles: (i) a *kebele* leader or member of the *kebele* leadership, (ii) an elder man/woman, (iii) a religious leader, (iv) a women’s representative, (v) a teacher, development agent, or health extension worker, and (vi) a youth representative. Most of these elites (community leaders) had previous targeting experience, and most had previously participated in targeting for the PSNP.⁹ As described in Abay et al. (2026), villages were randomly assigned to four treatment groups, each with varying levels of budget and discretion. The budget was either hypothetical or actually provided in cash, with a relaxed (20,000 Ethiopian Birr, about US\$360 at the time of the survey) or constrained (10,000 Ethiopian Birr) amount. The level of discretion depended on the required criteria to assess neediness, either rule-based, meaning the community representatives were provided with a predetermined list of criteria (including food security and assets), or discretionary, meaning they could freely choose their criteria and eventually disclose them. The four randomly

⁶ The FtF program in Ethiopia aims to sustainably reduce poverty, hunger, and malnutrition by: (i) reorienting agrifood system and supply chains, (ii) promoting climate-smart production and small-scale irrigation, (iii) advancing nutrition-led interventions, and (iv) facilitating enterprise growth and financial inclusion through public–private partnerships.

⁷ Following recent referendums, the SNNP region was dissolved and split into four regions: Sidama Region, Southwest Ethiopia Region, South Ethiopia Regional State, and Central Ethiopia Regional State. This split was triggered by protests and local demand, which ultimately led to successive referendums occurring between 2020 and 2023.

⁸ Remaining EAs were inaccessible because of insecurity and ongoing conflicts in Amhara and Oromia regions.

⁹ For example, more than half of the community leaders and elites included in this process had prior experience in the targeting of the PSNP (Abay et al., 2026).

assigned groups, illustrated in Figure A1, were organized as follows: (i) a group that conducted a rule-based ranking for hypothetical social assistance amounting to 20,000 Ethiopian Birr, (ii) a group asked to rank households using similar criteria for an actual transfer of 20,000 Ethiopian Birr, (iii) a group tasked with ranking households using similar criteria for an actual transfer of half the amount, and (iv) a final group that ranked households discretionarily and was given 20,000 Ethiopian Birr to be allocated to needy households. In each group, community members could decide how many households to allocate the funds to and how to divide the amount among them.

Ranking and network data. During the follow-up survey in 2023, households were asked to complete a peer-ranking exercise: respondents ranked the 20 randomly selected households within their village, including themselves, from the poorest (i.e., most in need of social assistance) to the richest (i.e., least in need). Each respondent within the village received the same list of households to rank. Prior to the ranking exercise, respondents reported their relationship with each listed household. A filter question assessed whether the respondent knew (i.e., had heard of) the household; if so, a follow-up question captured the degree of familiarity using six categories, ranging from “not familiar at all” to “family members”. Respondents were instructed to rank all households, including those they reported not knowing. In addition, the availability of household GPS coordinates allows us to compute pairwise distances between households. The household questionnaire module used for the ranking exercise is reproduced in Figure A2.

Parallel to the peer and self-ranking exercise, we also administered a CBT process using community leaders, whom we refer to as local “elites”. The CBT exercise was conducted 1–2 months before the household survey and hence before the peer and self-ranking exercise. Community leaders jointly reported, for each household, the closest relationship held by any member of the group, using the same categories as in the peer survey. They then provided a single ranking of all households in the village based on their need for social assistance, from the neediest (poorest) to the least needy (richest). Finally, they allocated the total budget assigned to the community among the “eligible” and most needy households in the list. The community survey also recorded the GPS location of a central village facility, in most cases the center where the committee met to discuss and deliberate on the targeting decisions. We use the distance between each household and this point as a proxy for proximity to local elites, under the assumption that community representatives primarily operate from the village center.

Three features of the elite-based ranking are worth noting.¹⁰ First, unlike the peer ranking, the elites were exposed to either a hypothetical or an incentivized ranking and targeting process. Whether incentivized ranking outperforms hypothetical ranking is an empirical question we explore in Section 4.6. Second, although the actual transfers are not large enough to change poverty rankings, the fact that some households received some actual transfers may mean that the survey-based welfare measures for these households could change.¹¹ To understand the impact of the different treatments and households’ access to cash transfers, we disaggregate the elite-based rankings across these treatment groups

¹⁰ Abay et al. (2026) provide more details on the design and implementation of the elite-based targeting. Using the elite-based targeting experiment and associated exogenous variations in budget constraint and level of constraint given to community leaders, Abay et al. (2026) examine the impact of budget constraints on targeting outcomes, including inclusion and exclusion error. Except in our use of the elite-based ranking, the two papers address completely different research questions. In effect, this paper focuses on comparing self-, peer-, and elite-based ranking, while the other paper (Abay et al., 2026) focuses on studying the implication of changing the design of CBT processes.

¹¹ As shown in Abay et al. (2026), about 80% of households received some transfer, implying that most households have received the transfer, although the size of the transfer is not large enough.

and compare them with conventional benchmark rankings. These comparisons and tests are reported in Section 4.6. Third, the elite-based ranking was administered about 1–2 months before the peer and self-ranking. We are not aware of major events and shocks that might have happened during the 1–2 month difference in survey timing.

Conflict data. Households’ geolocalization is used to merge survey data with conflict data from the Armed Conflict Location and Event Data (ACLED) Project (Raleigh et al., 2010). We measure exposure to conflict as the occurrence of at least one battle event within a three-year period preceding the 2023 survey, mainly to capture exposure to the armed conflict that started in 2020. We consider battles that happened within 5 km of every urban household and within 10 km of rural households. With such a definition, 35% of our sample was exposed to conflict prior to the survey (Table A1).¹²

Weather data. Data on precipitation anomalies from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS), provided with 0.05 degrees resolution, are geographically merged with other variables. The Standard Precipitation Index (SPI) is the most used indicator for the identification of meteorological droughts. It measures precipitation anomalies by constructing an indicator of deviation of precipitation in a given time period relative to historical average rainfall in a given location. The distribution of the SPI is particularly skewed to the left in rural areas (see Figure A4). We define households affected by droughts as those living in areas where the SPI was below -0.5 (corresponding to 48% of our sample as shown in Table A1), a distribution-driven threshold that allows us to divide the sample into two groups of comparable sizes and to capture a negative deviation within the urban sample.

2.2. Welfare- and wealth-based benchmark rankings (BR)

To assess the relative accuracy of SPR coming from self, peers, and elites, we compare them with benchmark rankings (BR) based on conventional measures of welfare and wealth. Specifically, we construct expenditure-based rankings using household daily food, non-food, and total expenditures, both in aggregate and per capita. In addition, we consider three asset-based wealth measures. The first is a general asset score, based on the first principal component of ownership of all assets recorded in the survey, along with improved housing conditions. Second, since subjective rankings may be influenced by more visible wealth indicators, we also compute a *visible assets* score, using a subset of visibly observable assets and housing characteristics. The first component for both assets and visible assets is calculated separately for rural and urban areas. Then, household rankings are calculated based on both assets and visible asset scores within each village. Third, we consider livestock in tropical livestock units (TLU) as an additional indicator of wealth, and we rank households within the village according to the quantity of livestock owned. Benchmark rankings are obtained using a ranking function that assigns ordinal values from 1 to the total number of ranked households within the village:

$$BR_{jv} = \text{Rank}(W_{jv}) \quad (1)$$

where W_{jv} denotes one of the survey-based welfare or wealth indicators for household j in village v .

2.3. Subjective poverty rankings (SPR)

Each respondent i is asked to rank all participating households in the village, including itself, according to their needs for potential social assistance, from the neediest (first) to the least needy (last). The majority of the participating households provided a full and exactly

¹² Figure A3 illustrates the share of urban and rural households exposed to battles by the number of battles that occurred in the three-year window.

progressive ranking of other households in the village. Whenever a respondent did not assign a rank to household j (or did not provide any rankings at all), that value is treated as missing. This is because it may reflect the observer's lack of knowledge about household j 's condition, and we lack any anchoring information to impute how the observer might have ranked that household. Missing ranks account for about 0.1% of all possible pairwise ranks.

Let M_v be the total number of households being ranked in the village v , and N_v be the number of ranking respondents in the village v , with $N_v \leq M_v$. Let $S_v = \{1, 2, \dots, M_v\}$ be the set of possible values for j in village v so that $i \in \{1, 2, \dots, N_v\} \subseteq S$. We denote the rank assigned to household j in village v , as given by respondent i , as g_{ijv} , which represents the relative position of household j as ranked by respondent i . We also define J_i as the number of households in village v for which respondent i provided a ranking and for which expenditure and asset data are available, such that $J_i \leq M_v$.

Tied and missing ranks in pairwise rankings from peers. In 7% of all pairwise rankings, respondents i assigned the same rank position to multiple households, that is, $g_{ijv} = g_{ikv}$ with $j \neq k$. In this case, the ranks are considered *tied*. By contrast, only 0.1% of all pairwise rankings were missing, meaning that the respondent did not provide a rank position for household j . We examine whether missing and tied ranks are systematically associated with specific characteristics of the observers or the ranked households. The results indicate that tied ranks are more likely to occur when observers are uneducated, less familiar with the household being ranked, or when the household appears in a higher position in the list (i.e., above 15). However, only the latter characteristic is also significantly associated with *missing* ranks (Table A2). Importantly, enumerators may play a key role in supporting respondents in providing complete and strictly ordered rankings.

Trimming rankings presenting a high number of missing or tied positions. We exclude rankings from observers who assigned the same rank too frequently to different households. Specifically, we calculate for each ranking household i a rank-progressivity measure α_{iv} , which corresponds to the share of unique rank positions assigned by household i . We define J_i as the set of households actually ranked by household i . The set of rank positions that respondent i in village v assigned to one and only one household j is: $U_{iv} = \{r \in S_v : \exists! j \in J_i \mid g_{ijv} = r\}$. Therefore, the rank progressivity measure is calculated as: $\alpha_{iv} = n(U_{iv})/J_i$, where $n(U_{iv})$ represents the cardinality of the set U_{iv} . We drop all rankings falling in the first percentile of the progressivity measure (α_{iv}^{q1}), corresponding to 0.2. In other words, we drop the rankings assigned by observers who assigned non-missing, unique positions to fewer than 20% of listed households. This trimming is quite conservative, as it preserves most of the rankings even if imprecise or containing some missing values. For the remaining rankings, the average rank-progressivity measure is 0.97, meaning that, on average, observers assigned a unique rank to 97% of the listed households. Therefore, the self-rank and peers' rank used in the rest of the paper are based on such selection of pairwise rankings. For consistency of notation, we define the post-trimming pairwise ranking as p_{ijv} where $p_{ijv} = g_{ijv}$ if $\alpha_{iv} > \alpha_{iv}^{q1}$. We later discuss the implications of different choices regarding the trimming method.

Self-ranking. The self-rank (SR) of household j in village v is defined as:

$$SR_{jv} = p_{ijv} \quad \forall i = j.$$

Peers' ranking. The mean position of household j in village v is given by:

$$\bar{p}_{jv} = \frac{1}{K_{jv}} \sum_{i=1}^{K_{jv}} p_{ijv} \quad \forall i \neq j.$$

where K_{jv} is the number of peers providing the rank for household j , excluding j itself. The peers' rank is then defined by applying a ranking function to $\bar{p}_{jv}$:

$$PR_{jv} = \text{Rank}(\bar{p}_{jv}) \quad (2)$$

where $\text{Rank}(\cdot)$ is a ranking function that assigns discrete ordinal positions from 1 to J_i to all j households ranked by peers. The ranking function is the same as that applied in Eq. (1). In the rest of the paper, the expression *peers' ranking* will be used to indicate PR_{jv} , while *pairwise ranking* will be used as the equivalent of p_{ijv} .

Robustness of ranking construction. We assess the sensitivity of our results to alternative procedures for constructing both peers' and self rankings. First, we compare the baseline mean-based aggregation of pairwise rankings with several alternatives, including the median, the normalized Borda count, the stitching method from Dupas et al. (2024), and the HodgeRank algorithm adopted by Bloch and Olckers (2021). Second, we consider different selection rules for pairwise comparisons, including alternative trimming thresholds and restrictions based on respondents' knowledge of other households. Across all specifications, the resulting rankings are highly consistent. Correlations across aggregation methods are uniformly high (Table A3), and the relationship between subjective rankings and benchmark measures of consumption and wealth remains largely unchanged. Similarly, alternative selection rules yield very similar results (Tables A4 and A5), except for an expected increase in the correlation between peers' rankings and benchmark rankings when ranks assigned to unknown households are excluded from the construction of the peers' ranking. The peers' ranking obtained after excluding ranks to unknown households is, however, highly correlated with the one obtained on the full set of pairwise ranks, between 82 and 89% depending on the method used (see Table A6). Taken together, these findings indicate that our results, given the nature of the full pairwise rankings we have, are not sensitive to the specific aggregation method or selection rule employed, supporting the use of the baseline specifications for ease of interpretation. A more detailed discussion is provided in Appendix B.1.

Elites' ranking. The third ranking is provided by community representatives. We denote this ranking as ER_{jv} . Unlike the peer ranking, no aggregation is required here, since the six community representatives are asked to come to an agreement and provide a unique ranking of the households in their village.

Sample selection. The 2023 survey covered 2906 households. The sample used in the analysis is selected through a five-step process, as illustrated in Table A7. First, we retain only households with complete data on both consumption and wealth measures for 2023. To reduce the influence of extreme values on the expenditure-based ranking, we further exclude households in the bottom and top 1% of the distribution of food, non-food, and total expenditure. Then, we remove households for which any of the subjective rankings (self, peer, or elite) is not available. The final sample used in the analysis is therefore composed of 2625 households.

To ensure comparability, we reapply the within-village ranking procedure to benchmark, peer, and elite rankings, restricting it to the final sample of households retained for analysis. Table A1 reports summary statistics for the selected sample. After the sample selection process and the drop of rankings according to the rank progressivity parameter α , the average number of peers providing a ranking for household j is 16.4, leading to an average peers' rank (PR) of 8.7. Because in some villages the number of households listed was less than 20 (minimum 14), and because we reapply the ranking function on the final selected sample, the average elites' rank (ER) is also 8.7. Most households live in rural areas (88%), and 86% of them are involved in crop farming or livestock business (mixed farming households).

Cost of administering welfare-based and subjective rankings. The welfare data (from consumption expenditure) and subjective ranking were collected together with additional modules, and we do not have disaggregated data collection costs for each ranking method nor for expenditure. However, we have three pieces of information from our survey implementation records. First, the consumption module in our survey was comprehensive and took about 45–60 min, about half of the time needed to complete the survey.¹³ In terms of duration, the elite and peer rankings took significantly less time than the consumption module. Thus, collecting conventional welfare- and asset-based measures remains the most expensive undertaking. Second, if we assume that recruiting and surveying elites and respondents entails a similar cost, the total number of elites recruited (six from each village) in our sample is about a third of the total number of households, which implies that the cost of administering the elite-based ranking can be about a third of the self- or peer ranking. Third, the total cost of administering the elite-based ranking (and targeting experiment) was less than half of the survey cost for the household survey (which includes the consumption expenditure module and self- and peer ranking).

3. Do subjective rankings capture relative poverty?

We motivate the subjective rankings (self-, peer-, and elite-based) as proxies and approximations of households' unobserved latent deprivation or associated need for social assistance (Alkire and Foster, 2011; Lee et al., 2025). We hypothesize that the different subjective rankings capture both common and distinct dimensions of households' underlying latent need for social assistance or associated deprivation. Thus, we do not assume that subjective rankings are merely noisy measurements of a single welfare ordering. Rather, different rankers may observe or emphasize different dimensions of deprivation, leading to both shared and ranking-specific information. To assess whether the information captured by these subjective rankings predicts relative poverty as measured by conventional welfare and wealth measures, we quantify concordance and correlations using alternative measures. The concordance between subjective ranking and benchmark rankings is assessed through the estimation of Spearman's rank correlation between the subjective and the wealth- and welfare-based rankings. The Spearman's correlation coefficient (ρ) is calculated as:

$$\rho = 1 - \frac{6 \sum_{j=1}^n d_j^2}{n(n^2 - 1)},$$

where n is the number of ranked observations. The squared difference (d^2) between the positions assigned to unit j in two ranking vectors implies that larger discrepancies are penalized more heavily than smaller ones.¹⁴ Two identical rankings will have $\rho = 1$ and two opposite rankings will have $\rho = -1$. To assess whether differences in predictive accuracy across SPR are statistically meaningful, we test the null hypothesis that two Spearman correlations with the same benchmark are equal using a cluster-robust reflected bootstrap procedure. For each pair of rankings, we compute the observed difference in Spearman correlations and its distribution across 500 bootstrap resamples drawn with replacement at the village level. To impose the null of no difference, we recenter the bootstrap distribution around zero and define the p -value as the proportion of resampled differences whose absolute value exceeds the observed one. This approach is assumption-free with respect to the marginal distributions of the ranks and fully accounts for within-village correlation in the data. As a robustness test, we also estimate linear rank-on-rank correlations between the different subjective and benchmark rankings. Through these rank-on-rank correlations, we estimate the slope and hence relationships between BR and

Table 1

Spearman's correlations between subjective rankings.

	Self rank	Peers' rank	Elites' rank
Self rank	1.000		
Peers' rank	0.335	1.000	
Elites' rank	0.250	0.451	1.000
Observations	2625	2625	2625

Authors' calculation.

SPR. Thus, we also test whether the slopes are different from one and across different methods of subjective rankings.

Before turning to the correlation between subjective rankings and measures of expenditure and wealth, we first examine the rank correlation across our benchmark measures (see Table A8). The Spearman correlation between food and non-food consumption is 42% at the household level and 33% when considering per capita expenditure. At the same time, asset-based rankings correlate only weakly with per capita expenditure-based rankings. The correlation between assets and household-level non-food expenditure reaches 47% but is lower when considering food expenditure or per capita measures. Table 1 shows that correlations across subjective rankings are also relatively modest. The lowest correlation is observed between self and elites' rankings (only 25%), while peers' and elites' rankings exhibit a stronger, yet still moderate, correlation of 45%.

Table 2 presents the correlation between subjective rankings and benchmark rankings, namely per capita expenditure, total expenditure, and assets. Four important patterns are worth noting. First, the correlations between subjective rankings and rankings based on conventional welfare and wealth measures are relatively weak, ranging from 7% (between self rank and per capita expenditure) to 36% (between elites' rank and asset-based rank).¹⁵ Second, elite-based ranking generally outperforms both self and peers' rankings in capturing conventional welfare and wealth indicators. Peers' rank typically lies between the two. Third, all subjective rankings are better at capturing differences in household-level expenditure than in per capita expenditure. For example, the rank correlation between elite-based and per capita-based ranking amounts to only 20%, while the corresponding rank correlation with total expenditure stands at 31% with total household expenditure. The same comparison yields 6% versus 17% for self rank and 14% versus 24% for peers' rank.¹⁶ Fourth, subjective ranks are more strongly correlated with measures of wealth than with measures of consumption. The correlation with asset-based rankings is 25%, 33%, and 36%, respectively. This finding is consistent with (Trachtman et al., 2025), who document that subjective rankings are more closely aligned with asset-based measures than with consumption, and that targeting decisions are primarily driven by observable asset ownership. All of the above patterns are confirmed through our linear rank-on-rank correlations, which are reported in Table A10. In all cases, we can easily reject the null hypotheses that subjective rankings fully explain welfare-based rankings. Similarly, we can reject the equality of coefficients associated with each subjective ranking, and the elite-based ranking performs better than the other subjective rankings.

¹⁵ Figure A5 shows the correlations between subjective rankings and alternative benchmark measures, such as food and non-food expenditure, visible assets, and livestock. The plots include 95% confidence intervals computed using 500 bootstrap repetitions, clustered at the village level. For comparison, Spearman correlation coefficients for the full set of subjective and benchmark rankings are reported in Table A9.

¹⁶ We also examine whether subjective rankings implicitly reflect equivalized economies of scale by applying the OECD adult equivalence scale to food, non-food, and total consumption measures. However, correlations between subjective rankings and equivalized expenditure-based rankings are even lower than those based on per capita expenditure.

¹³ The average duration of the survey was about two hours.

¹⁴ In the calculation of Spearman's coefficient, tied observations receive the same average rank position.

Table 2
Spearman's correlations between subjective and benchmark rankings.

	Benchmark Ranking (BR) based on:		
	Total expenditure		Assets
	Per capita	Household level	
Self rank	0.065	0.175	0.254
Peers' rank	0.145	0.238	0.331
Elites' rank	0.201	0.314	0.358
Aggregate rank	0.194	0.324	0.417
<i>Test for coefficient differences:</i>			
Peers' – Self	0.080***	0.063**	0.078**
Elites' – Peers'	0.056*	0.076**	0.027
Elites' – Self	0.136***	0.140***	0.104***
Aggregate – Self	0.129***	0.150***	0.163***
Aggregate – Peers'	0.049**	0.086***	0.085***
Aggregate – Elites'	–0.007	0.010	0.059**
Observations	2625	2625	2625

Authors' calculation based on final sample (N = 2625). The bottom panel shows bootstrap tests for the difference between two dependent Spearman correlations (H0: $\rho(SPR_1, BR^W) = \rho(SPR_2, BR^W)$). P-values are obtained via a cluster-robust reflected bootstrap (500 replications, clusters defined at the village level). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The aggregate rank used in this table is the one based on the interaction between self, peers', and elites' rank.

3.1. Ranking aggregation

The low correlation between self, peers', and elites' rankings suggests that these measures capture different dimensions of household well-being and associated deprivation.¹⁷ These differences may reflect both the distinct information sets available to each observer and the different aspects of welfare these subjective rankings capture. We therefore investigate whether an aggregate ranking, combining information from self, peers', and elites' rankings, aligns more closely with welfare- and wealth-based rankings than any of the individual measures alone.

We propose and compare several aggregation methods. First, an interaction-based method combines the three rankings through their joint interaction (i.e., the product of the rankings). Second, principal component-based methods derive weights from the variance structure of the rankings — either using only the first component or all three components weighted by their explained variance — and use them to construct weighted aggregates. Third, a normalized Borda approach rescales rankings within villages and averages positions. Fourth, pairwise comparison methods are used: the stitching procedure proposed by Dupas et al. (2024) aggregates win-loss comparisons across raters, while the HodgeRank algorithm (used in Bloch and Olckers, 2021) recovers a globally consistent ranking from pairwise comparisons using a least-squares framework. Finally, a reliability parameter-based method assigns data-driven weights to each ranking based on their empirical proximity to benchmark measures, estimating these weights in calibration samples and applying them to validation samples across repeated bootstrap splits to construct a weighted aggregate. We provide more details on these aggregation methods in the Methodological Appendix B.2.

Concordance of aggregate rankings. Fig. 1 presents the estimated Spearman correlation coefficients between aggregated subjective rankings and benchmark rankings, with 95% confidence intervals. We note two observations from the patterns in Fig. 1. First, the alternative aggregation methods perform comparably, suggesting that computationally easier methods — such as the interaction-based aggregation — can generate results comparable to those of more computationally intensive approaches when all contributing rankings are available. Second, comparing the univariate subjective rankings with the aggregate rankings,

we observe modest improvements. Notably, the largest improvements relative to the best-performing individual subjective ranking (the elites' ranking) are observed for the benchmark that exhibits the strongest correlation with subjective assessments, namely the asset-based ranking. Table 2 shows bootstrap tests for differences between correlations: the interaction-based aggregation produces a significant improvement over both self and peer rankings for both benchmarks, and over elites' rankings for the asset-based benchmark only. The rank-on-rank regressions reported in Table A10 show consistent patterns: we can reject the null that the slopes associated with the elite-based and aggregate rankings are statistically indistinguishable for the asset ranks but not for expenditure-based rankings. Quantitatively, the interaction-based aggregation improves correlation with the asset-based benchmark by approximately 16.5%–18% relative to the elites' ranking, depending on whether the estimate is based on Spearman rank correlation or regression coefficients.

It is worth noting that weighting subjective rankings using the reliability parameter does not yield stronger correlations with benchmark welfare and wealth measures compared to alternative aggregation methods. This finding carries important policy implications. Comparable results can be achieved using non-anchored aggregation methods, which may be simpler to implement in practice.

4. Concordance and discrepancy between subjective rankings and welfare measures across various settings

Overall, the weak correlations between subjective rankings and welfare or wealth rankings shown in Fig. 1 indicate substantial discordance between conventional measures of welfare and wealth and subjective ranking methods. This discordance may arise from multiple factors. First, subjective assessments may capture different aspects and dimensions of poverty than those reflected in conventional metrics of welfare and wealth (Alatas et al., 2012; Basurto et al., 2020). In our case, the subjective rankings aim to identify households' perceived need for social assistance. Local communities and individuals may define “neediness” in broader or alternative ways compared to the economic indicators typically used in surveys. These perceptions may incorporate social vulnerability, illness, or marginalization — dimensions not always visible through consumption, asset, or income data. The fact that subjective rankings correspond to household wealth rather than consumption strengthens this. Second, it is important to acknowledge that the relatively low correlation between benchmark rankings themselves may reflect measurement error, especially in consumption data (Gibson et al., 2015; Abate et al., 2022). In this context, part of the discordance between subjective and benchmark rankings could arise from limitations in accurately capturing households' actual expenditure or asset holdings through survey instruments. As such, the observed discrepancies should not be taken as evidence against the relevance of subjective rankings. Instead, they highlight that our benchmark indicators represent conventional proxies for welfare and wealth rather than an objective truth. Third, the divergence between subjective assessments and benchmark measures may also result from unintentional misclassification driven by factors orthogonal to the characteristics of either the observing or observed households (misreporting). For instance, ranking 20 peers, and positioning oneself among them, can impose a significant cognitive burden, introducing random noise into the rankings. While such errors may not produce systematic bias, they can reduce the overall concordance between ranking systems. If discrepancies are mostly driven by misreporting during the ranking process, perhaps because of the cognitive burden and limited capacity to process a large set of information, such errors are unlikely to be correlated with respondents' perceptions of relative economic condition and subjective well-being. Fourth, subjective assessments may reflect beliefs, intentions, or misperceptions on the part of the respondents. If discrepancies mostly arise from these misperceptions, they are inherent to the observer's characteristics or point of view (Abay et al., 2021;

¹⁷ This is not surprising given that many households in rural economies face deprivations in multiple dimensions (Alkire and Foster, 2011; Lee et al., 2025).

Correlation with benchmark rankings based on:

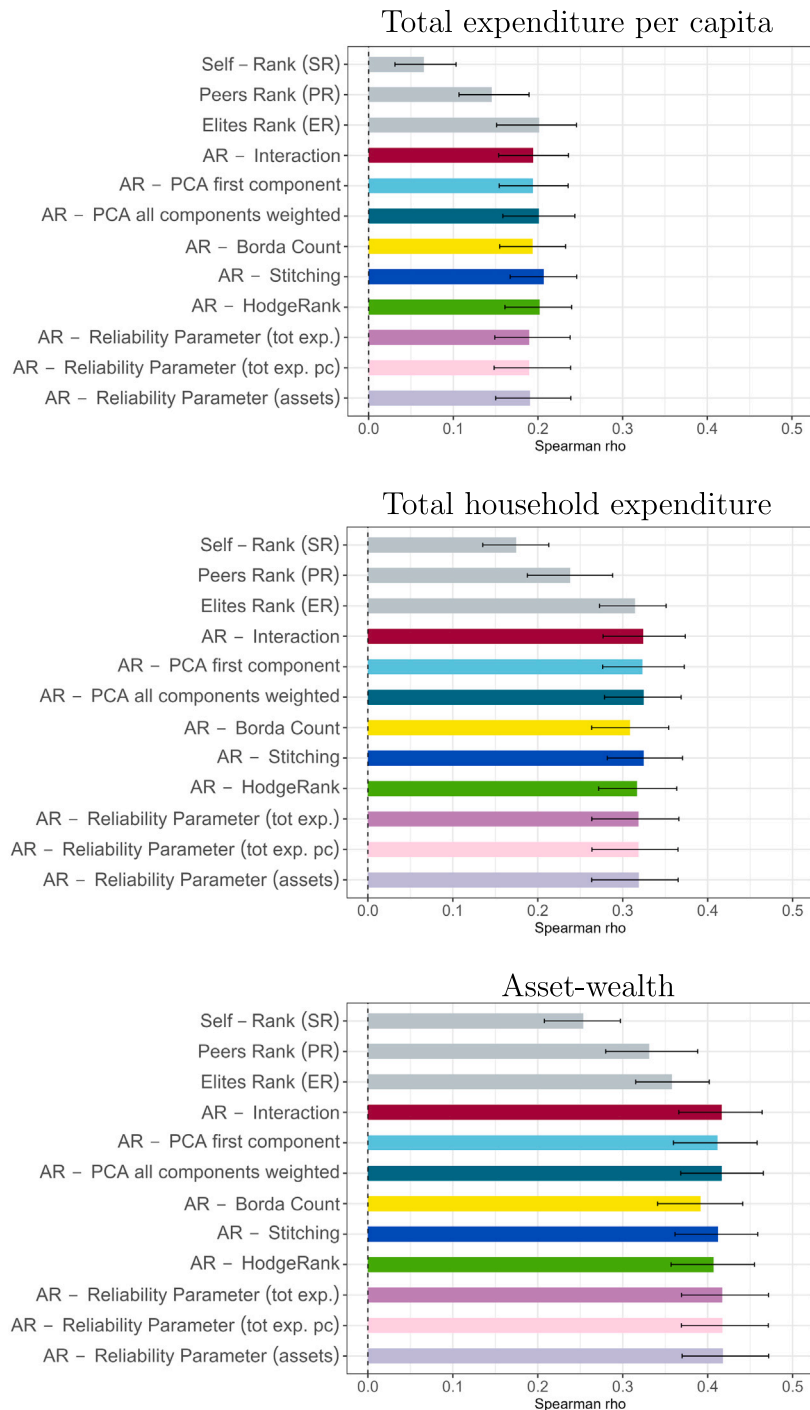


Fig. 1. Concordance of subjective poverty rankings with benchmark rankings.

Authors' elaboration. The graphs report Spearman ρ correlations between SPR, including various aggregation methods, and benchmark rankings based on expenditure and assets. The length of the horizontal bar corresponds to ρ , while the whiskers represent 95% confidence intervals computed using 500 bootstrap repetitions. AR indicates Aggregate ranks, computed with different methods as indicated in the labels. Number of observations = 2625.

Wossen et al., 2022), implying that these errors can be updated through additional information (Karadja et al., 2017). Even when individuals attempt to rank others (or themselves) based on conventional notions of welfare or wealth, their views may be shaped by limited information, social dynamics, or personal biases, ultimately leading to systematic discrepancies. Finally, the concordance and discordance between conventional measures of welfare and subjective rankings may vary across

contexts, geographies, and households. In the next subsections, we formally and parametrically examine these heterogeneities.

4.1. Heterogeneity in concordance and ranking discrepancies

Understanding and attributing the discordance between subjective and benchmark rankings is crucial for interpreting these differences

Table 3
Sociodemographic heterogeneity.

	Benchmark Rankings (BR) based on:					Assets				
	Household-level total expenditure									
	$\rho^{H=1}$	$\rho^{H=0}$	$\Delta\rho$	$\hat{\gamma}_{H \times SPR}$	$\hat{\beta}_H$	$\rho^{H=1}$	$\rho^{H=0}$	$\Delta\rho$	$\hat{\gamma}_{H \times SPR}$	$\hat{\beta}_H$
Urban ($H = 1$, $N_{H=1} = 315$) vs Rural ($H = 0$, $N_{H=0} = 2310$)										
<i>Self rank</i>	0.19	0.18	0.01	-0.04	-0.06	0.18	0.27	-0.09	-0.11**	0.19
<i>Peers' rank</i>	0.21	0.23	-0.03	-0.01	-0.60*	0.27	0.33	-0.06	-0.06	-0.42
<i>Elites' rank</i>	0.38	0.30	0.08	0.08	-0.97***	0.44	0.34	0.10*	0.12**	-0.97***
<i>Aggregate rank</i>	0.38	0.31	0.07	0.07	-0.81***	0.42	0.40	0.02	0.03	-0.56*
Pastoralist ($H = 1$, $N_{H=1} = 295$) vs Mix-farming household ($H = 0$, $N_{H=0} = 2077$)										
<i>Self rank</i>	0.05	0.20	-0.15**	-0.15***	0.42	0.02	0.30	-0.28***	-0.25***	0.83***
<i>Peers' rank</i>	0.08	0.24	-0.16*	-0.13	-0.45	0.21	0.34	-0.13*	-0.13**	-0.37
<i>Elites' rank</i>	0.20	0.30	-0.10	-0.07	-0.53*	0.19	0.36	-0.17**	-0.14**	-0.27
<i>Aggregate rank</i>	0.20	0.32	-0.12	-0.11	-0.48*	0.25	0.42	-0.17***	-0.19***	-0.18
Small households ($H = 1$ if less than 5 members, $N_{H=1} = 1072$) vs larger households ($H = 0$, $N_{H=0} = 1553$)										
<i>Self rank</i>	0.16	0.13	0.03	-0.05	-0.44**	0.24	0.23	0.01	-0.06*	-0.05
<i>Peers' rank</i>	0.19	0.24	-0.05	-0.06	0.13	0.28	0.35	-0.07*	-0.08**	0.34**
<i>Elites' rank</i>	0.22	0.33	-0.11***	-0.10**	0.24*	0.30	0.37	-0.07*	-0.09**	0.30*
<i>Aggregate rank</i>	0.26	0.32	-0.06	-0.08*	0.17	0.36	0.42	-0.06	-0.08**	0.26*

Authors' calculation. H is a binary subgroup indicator defined separately in each panel header. ρ^H is the Spearman's correlation coefficient between SPR and BR, computed for the two groups defined by H . The two-sided sign-test p -value for $\Delta\rho = \rho^{H=1} - \rho^{H=0}$ is obtained with cluster bootstrap at the village level (1000 repetitions with replacement). $\hat{\gamma}_{H \times SPR}$ is the estimated coefficient from the regression specification $BR = \gamma_0 + \gamma_1 SPR + \gamma_2 H + \gamma_{H \times SPR} H \times SPR + \epsilon$, with standard errors clustered at the village level. $\hat{\beta}_H$ is the estimated coefficient from the regression specification $|SPR - BR| = \beta_0 + \beta_H H + e$, with standard errors clustered at the village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

as well as appraising the potential and value of subjective ranking methods. We study heterogeneity in the alignment of SPR with BR using three complementary metrics. First, we assess concordance and hence associated heterogeneity across dichotomous groups (H) by computing subgroup-specific Spearman rank correlations ($\rho^{H=1}$ and $\rho^{H=0}$) between SPR and BR and their difference ($\Delta\rho$) estimated through bootstrap repetitions. The difference between subgroup-specific Spearman rank correlations ($\Delta\rho$) captures the heterogeneities in concordance across groups. Second, we use a regression-based framework with interaction terms, which capture differences in the strength of SPR-BR relationships across groups. We specifically estimate the following specification:

$$BR_j = \gamma_0 + \gamma_1 SPR_j + \gamma_2 H_j + \gamma_{H \times SPR} H_j \times SPR_j + \epsilon_j. \quad (3)$$

While the Spearman coefficient provides a non-parametric measure of ordinal agreement within each group, the interaction coefficient $\gamma_{H \times SPR}$ offers a formal test of whether the SPR-BR relationship differs across groups. Third, we compute *ranking discrepancies*, measured by absolute differences between SPR and BR, which capture the magnitude of divergence between the two rankings. We then test whether these absolute differences are statistically distinguishable across dichotomous groups (H) using the following empirical specification:

$$|SPR_j - BR_j| = \beta_0 + \beta_H H_j + \epsilon_j. \quad (4)$$

where the coefficient β_H captures the parametric difference in ranking discrepancies across groups.

To facilitate comparison with subgroup-based Spearman correlations, both regressions (Eqs. (3) and (4)) adopt a parsimonious specification. Tables 3–5 and 7 presented in this section (and corresponding tables in the Appendix) include: (i) the Spearman's rank correlations between subjective rankings and benchmark measures of total expenditure and wealth across the different heterogeneity factors and their difference ($\Delta\rho$), (ii) the regression-based coefficient of the interaction term between SPR and the heterogeneity factor on BR ($\hat{\gamma}_{H \times SPR}$), and (iii) the regression-based coefficient of the heterogeneity factor on the absolute differences between SPR and BR ($\hat{\beta}_H$).

The following discussion is organized by groups of factors that can explain these differences: (i) contextual and household characteristics; (ii) economic conditions and associated dynamics; (iii) exposure to shocks; (iv) social networks and spatial proximity; and (v) CBT participation and experience.

4.2. Contextual and household characteristics

Previous studies highlight the importance of context in shaping the performance of subjective rankings (Hillebrecht et al., 2020; Trahtman et al., 2025; Dupas et al., 2024). In this subsection, we focus on three contextual and sociodemographic factors: rural-urban differences, economic-based local norms, and household demography.

Rural-urban differences. The first panel of Table 3 reports concordance and discrepancy between SPR and BR in urban and rural areas. Overall, subjective rankings perform somewhat differently across the two settings, although the pattern is not uniform across ranking methods. For *self-rank*, concordance with total expenditure-based ranking is nearly identical in urban and rural areas ($\rho = 0.19$ vs. 0.18), and neither the interaction coefficient nor the discrepancy regression is statistically significant. With asset-based ranking, however, concordance is lower in urban areas ($\Delta\rho = -0.09$), and this is confirmed by a negative and statistically significant interaction coefficient, indicating weaker alignment of self-assessments with asset-based welfare in urban settings. For *peers' rank*, concordance is slightly lower in urban areas for both welfare measures, although differences are not statistically significant. At the same time, the discrepancy regression is negative and statistically significant only for total expenditure-based rankings, indicating that peer rankings are closer to expenditure-based rankings in urban areas in terms of absolute distance, while no clear difference emerges for asset-based rankings. For *elites' rank*, concordance is higher in urban areas, especially with asset-based ranking, and this is mirrored by a positive and significant interaction term. Moreover, the regression characterizing the discrepancy between the SPR and BR measures shows a negative and statistically significant coefficient associated with the “urban” dummy, indicating that elite rankings align more closely with benchmark measures in urban areas both in ordinal terms and in absolute distance. A similar pattern emerges for the *aggregate rank*, for which discrepancies are significantly smaller in urban areas and concordance is slightly higher.

Overall, these results point to a nuanced pattern. While self and peer rankings perform similarly across urban and rural areas, elites' and aggregate rankings perform better (in capturing conventional benchmark measures) in urban areas. These findings add some nuance to the mixed evidence in the literature. On the one hand, Dupas et al. (2024) and Aiken et al. (2025) document stronger performance of CBT in rural contexts, attributing it to denser social networks and better

Table 4
Exposure to shocks.

	Benchmark Rankings (BR) based on:					Assets				
	Household-level total expenditure									
	$\rho^{H=1}$	$\rho^{H=0}$	$\Delta\rho$	$\hat{\gamma}_{H \times SPR}$	$\hat{\beta}_H$	$\rho^{H=1}$	$\rho^{H=0}$	$\Delta\rho$	$\hat{\gamma}_{H \times SPR}$	$\hat{\beta}_H$
Drought ($H = 1, N_{H=1} = 1247$) vs no drought ($H = 0, N_{H=0} = 1362$)										
<i>Self rank</i>	0.23	0.13	0.09***	0.11***	-0.16	0.32	0.20	0.11***	0.14***	-0.25
<i>Peers' rank</i>	0.33	0.15	0.19***	0.19***	-0.40**	0.40	0.27	0.12**	0.13**	-0.19
<i>Elites' rank</i>	0.35	0.28	0.07	0.06	-0.00	0.35	0.37	-0.02	-0.02	0.31
<i>Aggregate rank</i>	0.40	0.25	0.14***	0.14***	-0.23	0.45	0.38	0.07	0.07	-0.00
Conflict ($H = 1, N_{H=1} = 927$) vs no conflict ($H = 0, N_{H=0} = 1698$)										
<i>Self rank</i>	0.22	0.15	0.07	0.07	-0.36*	0.27	0.25	0.02	0.04	-0.21
<i>Peers' rank</i>	0.38	0.16	0.23***	0.22***	-0.75***	0.43	0.28	0.15***	0.15**	-0.49**
<i>Elites' rank</i>	0.37	0.28	0.09**	0.09**	-0.27	0.36	0.36	0.01	-0.00	0.03
<i>Aggregate rank</i>	0.42	0.27	0.15***	0.15***	-0.53***	0.45	0.39	0.06	0.05	-0.16
Self-reported shock ($H = 1, N_{H=1} = 1303$) vs none ($H = 0, N_{H=0} = 1322$)										
<i>Self rank</i>	0.21	0.13	0.08**	0.10***	-0.27	0.29	0.22	0.07	0.10***	-0.29
<i>Peers' rank</i>	0.28	0.19	0.09**	0.08*	-0.21	0.33	0.32	0.01	-0.02	0.09
<i>Elites' rank</i>	0.28	0.34	-0.06	-0.07	0.29*	0.35	0.36	-0.01	-0.02	0.06
<i>Aggregate rank</i>	0.32	0.32	0.00	-0.00	0.06	0.40	0.42	-0.02	-0.03	0.08
Idiosyncratic shock ($H = 1, N_{H=1} = 898$) vs none ($H = 0, N_{H=0} = 1727$)										
<i>Self rank</i>	0.23	0.15	0.08*	0.11***	-0.27	0.29	0.23	0.06	0.08**	-0.19
<i>Peers' rank</i>	0.27	0.21	0.06	0.06	-0.04	0.32	0.33	-0.01	-0.03	0.15
<i>Elites' rank</i>	0.27	0.33	-0.06	-0.07	0.39**	0.32	0.37	-0.05	-0.07*	0.24
<i>Aggregate rank</i>	0.32	0.32	0.00	0.00	0.11	0.38	0.42	-0.04	-0.06	0.17

Authors' calculation. H is a binary subgroup indicator defined separately in each panel header. ρ^H is the Spearman's correlation coefficient between SPR and BR, computed for the two groups defined by H . The two-sided sign-test p -value for $\Delta\rho = \rho^{H=1} - \rho^{H=0}$ is obtained with cluster bootstrap at the village level (1000 repetitions with replacement). $\hat{\gamma}_{H \times SPR}$ is the estimated coefficient from the regression specification $BR = \gamma_0 + \gamma_1 SPR + \gamma_2 H + \gamma_{H \times SPR} H \times SPR + \epsilon$, with standard errors clustered at the village level. $\hat{\beta}_H$ is the estimated coefficient from the regression specification $|SPR - BR| = \beta_0 + \beta_H H + e$, with standard errors clustered at the village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5
Economic differences.

	Benchmark Rankings (BR) based on:					Assets				
	Household-level total expenditure									
	$\rho^{H=1}$	$\rho^{H=0}$	$\Delta\rho$	$\hat{\gamma}_{H \times SPR}$	$\hat{\beta}_H$	$\rho^{H=1}$	$\rho^{H=0}$	$\Delta\rho$	$\hat{\gamma}_{H \times SPR}$	$\hat{\beta}_H$
Poorer than in 2019 ($H = 1, N_{H=1} = 1484$) vs others ($H = 0, N_{H=0} = 1141$)										
<i>Self rank</i>	0.19	0.16	0.03	0.02	-1.11***	0.25	0.25	-0.00	-0.00	-0.16
<i>Peers' rank</i>	0.27	0.21	0.06	0.05	-0.34**	0.34	0.32	0.01	0.01	-0.02
<i>Elites' rank</i>	0.31	0.35	-0.04	-0.04	-0.03	0.36	0.35	0.02	0.01	-0.03
<i>Aggregate rank</i>	0.35	0.32	0.03	0.01	-0.28*	0.42	0.40	0.01	0.01	0.04
Above median village inequality ($H = 1, N_{H=1} = 1298$) vs others ($H = 0, N_{H=0} = 1327$)										
<i>Self rank</i>	0.14	0.21	-0.08*	-0.08**	0.42**	0.22	0.29	-0.07	-0.08*	0.39**
<i>Peers' rank</i>	0.20	0.28	-0.08	-0.08	0.38*	0.31	0.35	-0.04	-0.04	0.17
<i>Elites' rank</i>	0.27	0.36	-0.09**	-0.09**	0.38**	0.30	0.42	-0.12***	-0.11***	0.53***
<i>Aggregate rank</i>	0.28	0.36	-0.08*	-0.08	0.41**	0.38	0.45	-0.07	-0.07	0.31

Authors' calculation. H is a binary subgroup indicator defined separately in each panel header. ρ^H is the Spearman's correlation coefficient between SPR and BR, computed for the two groups defined by H . The two-sided sign-test p -value for $\Delta\rho = \rho^{H=1} - \rho^{H=0}$ is obtained with cluster bootstrap at the village level (1000 repetitions with replacement). $\hat{\gamma}_{H \times SPR}$ is the estimated coefficient from the regression specification $BR = \gamma_0 + \gamma_1 SPR + \gamma_2 H + \gamma_{H \times SPR} H \times SPR + \epsilon$, with standard errors clustered at the village level. $\hat{\beta}_H$ is the estimated coefficient from the regression specification $|SPR - BR| = \beta_0 + \beta_H H + e$, with standard errors clustered at the village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

information flows. However, our urban sample includes only secondary towns, which may help explain why our results differ from those found in the metropolitan area of Abidjan (Dupas et al., 2024). On the other hand, our findings also echo (Hillebrecht et al., 2020), who argue that in more urbanized settings, community perceptions of poverty may be more closely aligned with consumption-based measures. This may help explain why, in our Ethiopian sample, elite and aggregate rankings exhibit stronger alignment with benchmark measures in urban areas, despite potentially weaker social ties. One possible explanation is that, while rural communities benefit from richer social interactions and information sharing, they may place greater weight on multidimensional aspects of deprivation that are not fully captured by standard consumption- or asset-based measures.

Mixed farming households versus pastoral communities. The second panel of Table 3 compares the alignment between SPR and BR across pastoralist (households residing in predominantly pastoral areas and regions, namely Afar and Somali) and mixed-farming households (households

residing in non-pastoralist regions and involved in crop farming or livestock activities). A clear and consistent pattern emerges across all subjective ranking methods and benchmarks. Concordance is systematically lower among pastoralist households. This holds for self, peers', elites', and aggregate rankings, and is particularly pronounced for asset-based measures, where differences are large and statistically significant. The interaction coefficients confirm this pattern, as they are consistently negative and often statistically significant, indicating weaker alignment between subjective rankings and benchmark measures among pastoralist households. The discrepancy regressions point in the same direction. Absolute differences between SPR and BR are generally larger for pastoralist households, although statistical significance varies across specifications. Taken together, the two metrics provide coherent evidence that subjective rankings perform worse among pastoralist households, both in terms of ordinal concordance and in terms of the precision of household placement.

These findings are consistent with existing evidence on targeting in pastoralist settings in Ethiopia. Targeting in pastoral communities has

Table 6
Effect of incentives and budget constraint on elite-based ranking.

	Benchmark rank (BR) based on:			
	Total expenditure		Assets	
Elites' Rank (ER)	0.307*** (0.039)	0.254*** (0.055)	0.336*** (0.036)	0.383*** (0.042)
CBT: Incentivized vs Hypothetical	-0.160 (0.453)		-0.280 (0.445)	
ER × CBT: Incentivized vs Hypothetical	0.017 (0.047)		0.031 (0.045)	
CBT: Hypothetical vs Rule-based 20 K		-0.643 (0.652)		0.257 (0.544)
CBT: Constrained vs Rule-based 20 K		-1.603** (0.618)		-0.291 (0.571)
CBT: Discretionary vs Rule-based 20 K		-0.752 (0.642)		0.241 (0.652)
ER × Hypothetical		0.053 (0.067)		-0.047 (0.056)
ER × Constrained		0.145** (0.065)		-0.003 (0.059)
ER × Discretionary		0.064 (0.068)		-0.046 (0.069)
Constant	6.031*** (0.381)	6.674*** (0.528)	5.778*** (0.362)	5.521*** (0.406)
N obs.	2625	2625	2625	2625
R ²	0.102	0.105	0.129	0.129

Authors' calculation. Standard errors clustered at the village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Villages are randomly assigned to one of four treatment arms: (i) hypothetical budget (20 K ETB) with rule-based ranking (base group in column 1 and 3); (ii) incentivized (actual) budget (20 K ETB) with rule based ranking (base group in column 2 and 4); (iii) incentivized constrained budget (10 K ETB) with rule-based ranking; and (iv) incentivized budget (20 K ETB) with fully discretionary ranking. In columns 1 and 3, the incentivized group includes the latter three treatment arms.

Table 7
CBT-related factors.

	Benchmark Rankings (BR) based on:									
	Household-level total expenditure					Assets				
	$\rho^{H=1}$	$\rho^{H=0}$	$\Delta\rho$	$\hat{\gamma}_{H \times SPR}$	$\hat{\beta}_H$	$\rho^{H=1}$	$\rho^{H=0}$	$\Delta\rho$	$\hat{\gamma}_{H \times SPR}$	$\hat{\beta}_H$
Identified as eligible for transfer ($H = 1$, $N_{H=1} = 2130$) vs others ($H = 0$, $N_{H=0} = 495$)										
Self rank	0.16	0.18	-0.02	-0.04	0.13	0.24	0.28	-0.04	-0.07	-0.01
Peers' rank	0.21	0.30	-0.09*	-0.09*	0.24	0.30	0.43	-0.13**	-0.13**	0.34*
Elites' rank	0.30	0.33	-0.03	-0.11*	-0.15	0.34	0.31	0.03	-0.07	-0.06
Aggregate rank	0.30	0.39	-0.09**	-0.13**	0.13	0.38	0.49	-0.11**	-0.16***	0.27
Transfer beneficiary in incentivized groups ($H = 1$, $N_{H=1} = 1548$) vs others ($H = 0$, $N_{H=0} = 340$)										
Self rank	0.18	0.21	-0.03	-0.05	0.08	0.24	0.28	-0.04	-0.06	-0.05
Peers' rank	0.20	0.37	-0.16***	-0.17***	0.48**	0.30	0.49	-0.19***	-0.19***	0.58**
Elites' rank	0.29	0.37	-0.07	-0.18***	0.01	0.34	0.32	0.02	-0.08	0.03
Aggregate rank	0.30	0.46	-0.16***	-0.22***	0.34*	0.39	0.52	-0.13**	-0.18***	0.38*
High experience with targeting ($H = 1$, $N_{H=1} = 1816$) vs low ($H = 0$, $N_{H=0} = 809$)										
Elites' rank	0.33	0.29	0.04	0.02	-0.03	0.38	0.31	0.07	0.06	-0.21
Aggregate rank	0.33	0.30	0.03	0.02	-0.09	0.43	0.38	0.05	0.04	-0.17

Authors' calculation. H is a binary subgroup indicator defined separately in each panel header. ρ^H is the Spearman's correlation coefficient between SPR and BR, computed for the two groups defined by H . The two-sided sign-test p -value for $\Delta\rho = \rho^{H=1} - \rho^{H=0}$ is obtained with cluster bootstrap at the village level (1000 repetitions with replacement). $\hat{\gamma}_{H \times SPR}$ is the estimated coefficient from the regression specification $BR = \gamma_0 + \gamma_1 SPR + \gamma_2 H + \gamma_{H \times SPR} H \times SPR + \epsilon$, with standard errors clustered at the village level. $\hat{\beta}_H$ is the estimated coefficient from the regression specification $|SPR - BR| = \beta_0 + \beta_H H + e$, with standard errors clustered at the village level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

been particularly challenging for many reasons (Sabates-Wheeler et al., 2013; Lind et al., 2022). For instance, while the PSNP achieved progressive targeting in highland regions, targeting performance was weak in lowland and pastoral areas (Coll-Black et al., 2012; Sabates-Wheeler et al., 2013). Lind et al. (2022) emphasize that norms of sharing and redistribution within horizontal networks may undermine the ability to distinguish sharply between households, thereby complicating targeting efforts. Our results are consistent with this interpretation. In pastoralist settings, greater mobility, more dispersed social networks, and norms of resource sharing may reduce the observability and salience of relative economic differences across households. In contrast, mixed-farming households, which are typically more sedentary and embedded in denser local networks, may have access to more reliable information flows, facilitating more accurate subjective rankings.

Household composition. The third panel of Table 3 compares the alignment between SPR and BR across small and larger households. Overall, evidence based on total expenditure is mixed and does not reveal a clear systematic pattern. A clearer pattern emerges when considering asset-based rankings. Concordance is consistently lower for small households across peers', elites', and aggregate rankings, and these differences are statistically significant. The interaction coefficients confirm this pattern, indicating weaker alignment between subjective rankings and asset-based measures among smaller households. This pattern likely reflects how household size affects perceived wealth, as economies of scale and resource pooling can lead larger households to be assessed differently depending on the benchmark used. In this context, smaller households may provide less observable information on asset accumulation and long-term wealth, making them harder to assess

accurately.¹⁸ Households with a smaller size are likely to be newly formed households, which may have a short spell of residence in the community. Overall, while household size does not appear to play a major role when welfare is proxied by consumption, it becomes more relevant for asset-based measures, where smaller households are systematically more difficult to rank accurately.

4.3. Exposure to shocks

We next examine whether the concordance between SPR and BR varies with households' exposure to shocks. Shocks may affect the accuracy of subjective rankings through multiple channels: they can alter households' observable living conditions, reshape local information flows, and influence the salience of deprivation within communities. Importantly, different types of shocks may differ in their visibility to external evaluators. We begin by focusing on common covariate shocks in our setting: violent conflict and drought. About one-third of the households in our sample have been exposed to violent conflict in the three years preceding the survey, while 48% have been affected by drought. Exposure to conflict and drought varies mainly between villages, with little variation within villages (Table A12). The first and second panels of Table 4 show that concordance between SPR and BR is systematically higher among those households affected by covariate shocks such as drought and conflict, particularly for peers', elites', and aggregate rankings, when compared against total expenditure. A similar, though less uniform, pattern is observed for asset-based rankings. Overall, these results suggest that subjective ranking methods perform better at capturing the relative economic condition of households affected by covariate shocks, perhaps because these shocks facilitate the flow of information and increase the salience of deprivation within communities.

We also consider all types of self-reported shocks and explore heterogeneities between households reporting exposure to these shocks and those not reporting. These self-reported shocks include covariate shocks such as drought and idiosyncratic shocks. In contrast, when we examine heterogeneity by self-reported shocks, the improvement in concordance between SPR and BR is pronounced only for self- and peer-ranking (third panel of Table 4). Concordance between rankings based on consumption and self- and peers' rankings is statistically higher for those households reporting idiosyncratic shocks. The interaction coefficient is positive and significant for self-rank and peer-rank, confirming stronger alignment. However, no comparable pattern emerges for elites' and aggregate rankings, and the interaction terms and coefficients remain imprecise and inconsistent. Finally, when we isolate idiosyncratic events within the broader category of self-reported shocks — such as the death of a spouse or income earner, agricultural shocks (e.g., theft or lack of inputs), and discrimination-based shocks — we no longer find significant differences even for peers' and elites' rankings (fourth panel of Table 4). This suggests that idiosyncratic shocks are less observable to external evaluators and therefore less likely to be incorporated into subjective assessments of deprivation.

To conclude, these findings suggest that the informational content of shocks depends critically on their nature. Aggregate shocks, which affect large segments of the community simultaneously, increase the salience of economic differences and generate shared information that facilitates relative comparisons across households. In contrast, self-reported and idiosyncratic shocks are less observable and may be

known only within narrower social networks, limiting their usefulness for forming consistent community-wide assessments of welfare. As a result, while subjective rankings appear to better capture relative welfare in the presence of aggregate shocks, their performance does not improve systematically in response to more localized and household-specific economic disturbances. As shown in Table A13, these results can be attributed to a closer relationship between peers, which is more prevalent in cases of drought. Covariate shocks such as conflict and drought might not only affect individual households' well-being but also social dynamics and resource and information sharing within the village network. In the context of Ethiopia, covariate shocks such as conflict and drought are likely to trigger collective action and encourage intra-community support through a number of village associations such as *iddir* and *equib* (Dercon et al., 2006; Caeyers and Dercon, 2012; Abay et al., 2018).¹⁹ But, in both cases, the experience of shocks seems to make households' relative conditions more visible. This implies that subjective ranking methods can perform relatively better when communities are exposed to covariate shocks, indicating their relevance and potential to support targeting in data-scarce regions and times.

4.4. Economic conditions and dynamics

Previous studies have shown how individuals and groups may have systematic misperceptions driven by some underlying economic-related attributes. For example, Hvidberg et al. (2023) find that lower-ranked (poorer) individuals tend to overestimate their own position among others because they underestimate other people's income, while wealthier households underestimate their positions because they overestimate their peers. This is the so-called “center-bias”, also documented in Dupas et al. (2024) and Cruces et al. (2013). The presence of center-bias in our subjective rankings is clearly shown in the graphical representation of the correlations in Fig. 2, especially for richer households that systematically rank themselves lower than what the benchmark measure would require.

Related to relative people's income, we also examine whether the subjective rankings are more effective at identifying relatively poorer or richer households within the village. As shown in Table A14, subjective rankings better identify the top and bottom terciles of the relative welfare and wealth distributions than those in the middle. Building on these insights, we extend the analysis to an additional dimension of economic heterogeneity — such as intertemporal changes in welfare — that is less explored in the literature.

Dynamics in household welfare. By exploiting the panel structure of the survey data, we are able to assess the accuracy of subjective rankings relative to expenditure- and asset-based rankings collected at baseline in 2019. First, we explore the correlation between welfare outcomes across the two rounds (2019 and 2023). The intertemporal correlations between BR, reported in Table A15, appear to be weak. This is not surprising given the stochastic environment households face in the four-year gap between the two rounds. Between the two rounds, Ethiopia experienced one of the deadliest armed conflicts in its history, which has affected households and communities (Abay et al., 2023). Interestingly, the intertemporal correlation of BR is higher for households exposed to conflict with respect to those less exposed, suggesting that conflict might crystallize existing inequalities. The low correlation between welfare and wealth rankings between 2019 and 2023 suggests significant changes in the distribution of resources or living conditions over the period, plausibly driven by other types of shocks (i.e., drought) experienced by a significant share of the sample (Aragie and Thurlow, 2022). This implies that predicting welfare outcomes in such stochastic environments can be challenging. Whether

¹⁸ Moreover, additional evidence from pairwise rankings (see Table A11) suggests that household size predicts both how households are perceived and how they evaluate others: larger households tend to be assigned higher positions but also display more nuanced assessments when acting as evaluators. More broadly, demographic composition — such as the presence of children or adult males — can shape perceptions of economic status, especially when rankings reflect notions of productive capacity (Abay et al., 2026).

¹⁹ For example, in the case of Ethiopia, Nigus and Abay (2024) show that covariate shocks such as battles encourage cooperation.

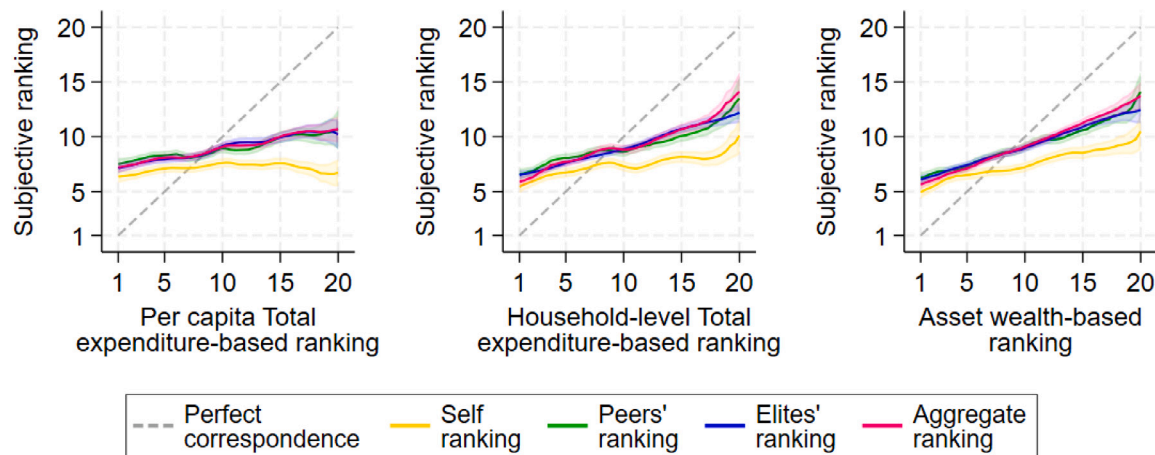


Fig. 2. Subjective vs welfare- and wealth-based ranking.

Authors' elaboration. Local polynomial smooth plots with 95% confidence intervals. The aggregate ranking used in the plot is the one based on the interaction of the three subjective rankings.

the subjective ranking measures respond and capture these dynamics is an important empirical question. We find that subjective rankings better correlate with contemporaneous (2023) rankings of welfare and wealth than with those from 2019 (see Table A16), suggesting that observers update (at least partially) their information and beliefs over time. Trachtman et al. (2025) found that subjective assessments tend to reflect longer-term wealth indicators (i.e., assets) rather than welfare metrics that are more sensitive to shocks (i.e., expenditure). While our results are consistent with this finding, we also find evidence that, despite a stronger reliance on asset indicators, observers do adapt their perception to reflect more recent conditions in both wealth and consumption. We further show in Table A17 that rankings based on past benchmark indicators predict present benchmark rankings similarly to aggregate rankings. For total expenditure, elites' rank alone also performs similarly to past expenditure-based benchmark rankings; however, this equivalence does not hold for asset-based rankings. In addition to the dynamic component, the first panel of Table 5 examines how changes in households' economic conditions are related to the alignment between SPR and BR, distinguishing households that became poorer relative to 2019 from those that did not. Overall, we do not find systematic differences in concordance across subjective and benchmark rankings. Households that became poorer are not consistently ranked differently from those that did not, across ranking methods and benchmark measures. However, households that became poorer exhibit significantly smaller absolute discrepancies between self and peers' rankings and expenditure-based rankings, indicating more precise placement (Table 5). This pattern does not extend to elite ranking or asset-based measures. Complementary evidence from pairwise rankings highlights an asymmetry across the welfare distribution: households tend to rank more accurately — relative to the expenditure-based benchmark — those peers who are wealthier than themselves (Table A18), while households that experienced a decline in welfare tend to rank themselves, and are ranked by others, more accurately (Table 5).

These findings point to two distinct but related dimensions. First, cross-sectional differences suggest that higher-status households are easier to locate in the welfare distribution. Second, intertemporal changes indicate that recent downward mobility is associated with higher ranking accuracy. Both patterns are specific to expenditure-based measures and do not extend to asset-based benchmarks. A plausible interpretation is that richer households are more easily identifiable due to more visible and persistent status signals (e.g., housing quality or social standing), and recent deteriorations in welfare are more socially visible than improvements, particularly in contexts characterized by widespread shocks, as also discussed in the previous subsection.

Within-group inequality. Economic distance can facilitate or limit understanding of the economic conditions of others in many ways. For example, Dupas et al. (2024) investigate the accuracy of peers' ranking in urban and peri-urban Côte d'Ivoire and suggest that the low correlation with traditional measures could be explained by the high income inequality among peers, which might correlate with physical distance between peers in these areas. However, large economic differences within the group of peers could have the opposite effect of making relative conditions more visible and therefore more easily incorporated into subjective rankings (Premand and Schnitzer, 2021).²⁰ We empirically test which force dominates in our context by computing within-group inequality using the village-level Gini Index based on households' per capita expenditure. In our sample, the within-village Gini Index ranges from 0.18 to 0.69, with an average of 0.33 (Table A1). The second panel of Table 5 compares the alignment between SPR and BR across households living in villages with above- versus below-median inequality. Concordance is systematically lower in high-inequality villages. This holds for self, elites', and aggregate rankings; results are similar for expenditure- and asset-based measures. Regression-based results point in the same direction. Results for asset-based ranking are statistically weaker but consistent with those for expenditure-based ranking. In sum, the three metrics provide coherent evidence that subjective rankings perform worse in more unequal contexts, consistent with the findings from Dupas et al. (2024). These findings suggest that higher inequality may complicate the accurate assessment of relative welfare, possibly because larger disparities and more heterogeneous economic conditions make it harder for community members to form consistent comparisons across households.

4.5. Social networks and spatial proximity

Social networks and spatial proximity can generate at least two important effects, which may have varying effects on the accuracy of subjective rankings. First, social networks and relationships allow observers to know their peers and hence address potential *information asymmetry*. This can create a stronger concordance between SPR and BR. Second, social relationships can trigger some level of *favoritism*

²⁰ Premand and Schnitzer (2021) show that CBT methods perform better in settings characterized by high levels of income inequality. However, their setting is less unequal than ours, with an average village-level Gini of 0.24 and values up to 0.4, and their Gini is based on proxy-means testing scores rather than observed consumption. These differences may matter if the relationship between inequality and ranking accuracy is non-linear.

and *nepotism* (Alatas et al., 2019), especially given that we elicited and motivated the ranking exercise based on households' need for potential social assistance. For example, Alatas et al. (2019) show that the families and relatives of those assuming local leadership are more likely to receive benefits in a social assistance program in Indonesia. Similarly, Vera-Cossio (2022) finds that households with direct connection to the village council were disproportionately favored in credit allocation through the community-based lending program. These tendencies in *favoritism* and *nepotism* are likely to weaken the concordance between SPR and BR. We empirically explore the relevance of these forces in our context, considering relationships between peers as well as between elites and households.

Relationship with and geographic distance between peers. Before starting the ranking exercise, the survey elicited relationships between ranking households and those being ranked, as well as between elites and households. We coded these relationships between ranking households (or elites) and those being ranked as follows (in a range from 0 to 6): 0 if completely *unknown*; 1 if *known* (heard of) *but not familiar with*; 2 if peers or elites can *recognize someone in the family but have never interacted*; 3 if peers or elites have *talked to a member of the family before infrequently*; 4 if peers or elites have *talked to a member of the family before frequently*; 5 if *close friend of the family* of peers or elites; and 6 for *family member*. The distribution of responses across these categories in our data, as reported in Table A19, is reasonably intuitive: while a small share of households are unknown to peers and elites, most households are known, and a few are closely related to peers and elites.

Fig. 3 illustrates how the relationship between observers and ranked households relates to discrepancies between welfare-based rankings and peer- and elite-based rankings. We construct such relationships for rural and urban areas separately. The top panel of Fig. 3 suggests three important empirical patterns worth noting. First, discrepancies between SPR and BR are highest for those households that are not known to observers, highlighting the role of information asymmetry in the ranking process. *Unknown* households are more likely to be ranked unfavorably for the potential social assistance, and more so in rural areas. The evidence that unknown households face more severe “punishment” both in rural and urban areas makes SPR vulnerable to the presence of information asymmetry about community members' economic conditions. In rural areas, especially, households that are related to the observer (family members) seem to be generally ranked lower than they should be according to conventional welfare and wealth measures. Second, the graph demonstrates that the discrepancy between SPR and BR is lowest in the middle of the distribution of acquaintance and relationships between peers and households, and these patterns hold for both rural and urban areas. Third, the direction of bias and associated discrepancies between peer-ranking and BR reverses for those households who are family members of the observing households. Households that are family members or closely related to the observing peers are likely to be ranked favorably for their need for social assistance, even more than what their welfare- or wealth-based measures would suggest. This points to the importance of favoritism and nepotism when social ties and relationships are strong.

We further investigate these patterns using parametric regressions and explore whether the pairwise relationship between peers is systematically correlated with the discrepancy between pairwise ranking (p_{ijv}) and benchmark measures. Table A20 shows that, with respect to the highest-frequency category (4 – *have talked to a member of the family before frequently*), households with weaker social ties with the observer are systematically given a higher rank than their position in the welfare or wealth distribution would suggest. These patterns reverse when peers are closely related or family members among each other. These results confirm the non-parametric results.

Relationship with village elites and proximity to the village center. The relationship between households and elites is coded similarly to peers'. However, the six elite members are asked to jointly report only one degree of relationship for each listed household—i.e., the closest relationship that any of them has with a member of the household. As shown in Table A19, elite members are less likely to report a household as completely unknown (only 2.83% of the cases versus 20.6% for peers), suggesting that elite members have at least a better general knowledge of community members. At the same time, elites are closely related (family members or friends) to the ranked households as frequently as peers are.

The implications of the relationships between elites and households on community-based ranking show some similarity with those from peer-ranking. In particular, the highest level of difference between SPR and BR is observed for households that are not known to elites. On average, unknown households are “punished” by elites, in some cases even more so than by peers, as shown in the bottom panel of Fig. 3. As with the peer-ranking, the least discrepancy between elite-based ranking and BR is observed in the middle of the distribution of relationship, for cases when elites know households but they are not close enough or have no familial ties. This is intuitive: some level of familiarity with households allows elites to observe their economic condition and hence address *information asymmetry*, but very close or familial relationships make elites vulnerable to favoritism.²¹ The strong “punishment” of unknown households by elites is confirmed by regression results presented in Table A21.

Finally, as shown in Table A22, proximity to the village center is associated with improved performance for elites' and, to a lesser extent, aggregate rankings, with higher concordance and smaller discrepancies. This reinforces the interpretation that visibility and centrality within the village enhance observability and hence accuracy of rankings that rely on general community-level information. Overall, these findings suggest that social and spatial proximity among peers, as well as between households and elites, might reduce information asymmetry, while close relationships can also trigger favoritism.

Concordance after removing ranks assigned to unknown households. Excluding rankings assigned to households unknown to the observer — whether peers or elite members — slightly improves the concordance of peers' and elites' rankings with household-level expenditure- and asset-based benchmarks (Figure A6). To implement this restriction, we construct peers' rankings using only pairwise comparisons from peers who report knowing (i.e., having heard of) the household, and we exclude from elites' rankings households that elite members report not knowing at all. This reduces the sample by 80 observations, primarily due to households unknown to elite members (75 observations). We then re-rank both subjective and benchmark measures within the restricted sample to avoid downward bias in correlation coefficients arising from gaps in rank positions caused by these excluded observations. Among known households, peers' rankings capture relative asset wealth nearly as well as elites' rankings, while still performing somewhat worse in capturing expenditure, though this difference is not statistically significant. Aggregation of subjective rankings constructed from ranks provided only by peers and elites who know the household yields results similar to those obtained with aggregate ranks based on the full sample of rankings (Figure A6). In particular, aggregate rankings display comparable concordance with expenditure-based benchmarks relative to individual subjective measures, and slightly stronger concordance with asset-based benchmarks.

²¹ Another slightly unique pattern we observe is the differences in vulnerability to favoritism across rural and urban areas. Favoritism driven by close ties or familial relationships appears to be slightly more pronounced in urban areas than in rural areas.

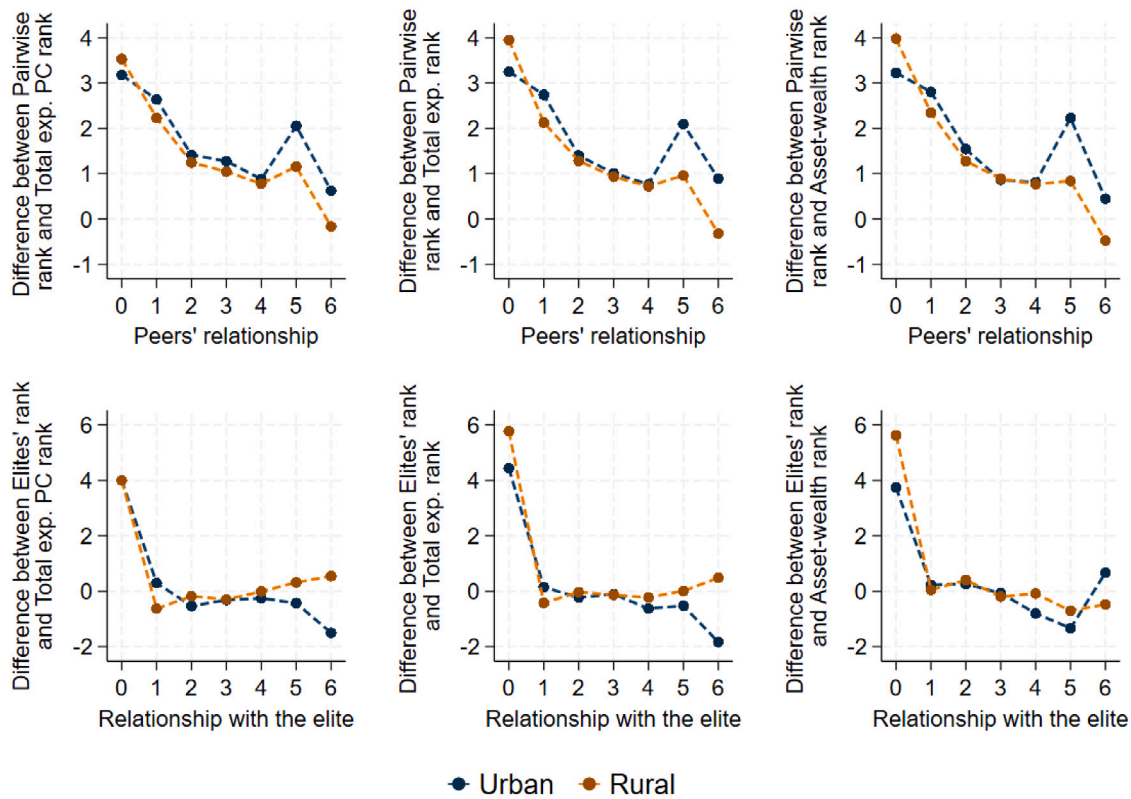


Fig. 3. Differences between subjective and benchmark rankings by relationship strength.

Authors' elaboration. Relationship: 0 = Unknown household; 1 = Not familiar at all; 2 = Know of/can recognize someone in the family but have never interacted; 3 = Have talked to a member of the family before infrequently; 4 = Have talked with a member of the family before frequently; 5 = Close friend of family member; 6 = Family member/related.

4.6. Community-based targeting participation and experience

As described in Section 2.3, the elite-based ranking was implemented under different treatment arms involving both incentivized and hypothetical targeting (ranking) as well as varying budget constraints (see Figure A1). Theoretically, these differences can generate three sources of variation on the performance of the elite-based ranking in terms of capturing welfare and wealth. First, incentivizing the targeting and ranking process can theoretically change the behavior and effectiveness of elites in identifying the poor and hence their ranking of households, compared to a hypothetical ranking without immediate consequences. Second, changes in budget constraints might affect the ranking process and outcome. Third, although the average transfer to beneficiary households remains small, and hence impacts on consumption expenditure appear to be negligible, access to the transfer through the incentivized targeting arm can, in principle, affect the welfare-based ranking of households. We empirically test these three possibilities by characterizing the relationship between the benchmark rankings and the elite-based rankings while also accounting for (and interacting with) these variations in our empirical regressions.

Hypothetical versus incentivized targeting. The first and third columns of Table 6 show that incentivized targeting (and ranking) does not generate meaningful and statistically significant improvement in concordance between elite-based ranking and benchmark rankings. The interaction effects involving elite-based ranking and the incentivized treatment arm are statistically insignificant.

Constrained versus relaxed budget. The second and fourth columns of Table 6 show weak evidence that the concordance between elites' rankings and expenditure-based measures improves under a constrained

budget, suggesting that resource scarcity may enhance targeting accuracy according to conventional metrics; however, this pattern does not extend to asset-based rankings. When the budget is constrained — so that few households receive transfers — elites might pay more attention and effort to minimize exclusion error.

Being selected to receive a transfer from the CBT experiment. Across treatment arms, elite respondents identified eligible beneficiaries by ranking households according to their perceived level of need and by allocating transfers under a predefined budget. Therefore, eligibility and transfer receipt emerged endogenously from elites' targeting decisions. The key difference between the control arm and the incentivized arms is that the former involved hypothetical allocations, while the latter involved the allocation of real transfer funds. As a result, targeting decisions in incentivized groups carried direct financial implications for beneficiary households. Importantly, households were not explicitly informed about others' beneficiary status, although such information may have been informally learned through communication among households. In the control arm, this is unlikely to have occurred since no transfers were actually implemented, meaning that unless informed by elites, households would generally not know whether they had been identified as eligible. In the first panel of Table 7, we compare the accuracy of subjective rankings in capturing benchmark indicators across households classified as eligible and non-eligible; in the second panel, we focus on incentivized groups and therefore explicitly focus on beneficiaries and non-beneficiaries. Eligible households account for roughly 80% of households across the four treatment arms (Table A1). Differences in the concordance between subjective and benchmark rankings for eligible and non-eligible households may arise for several reasons. In

the case of elites' rankings, concordance may vary with elites' relative aversion to inclusion versus exclusion errors.²² For self-rankings, differences may reflect a perception bias induced by transfer receipt in the incentivized arms: being identified as poor through beneficiary status may lead eligible respondents to rank their own households as poorer. Finally, peers may perceive eligible households as poorer than they actually are, particularly in incentivized arms. Alternatively, peers may strategically rank eligible households as less poor if they believe these households were undeservedly selected as beneficiaries. The first and second panels of Table 7 show that concordance between elite and benchmark rankings is broadly similar across eligible and non-eligible households and across beneficiaries and non-beneficiaries, except for lower discordance with expenditure rankings. No differences emerge for self-rankings. The only notable divergence concerns peers' rankings, reflected in aggregate measures: concordance is lower among beneficiaries for expenditure- and asset-based rankings.

Elites' experience with targeting. The third panel of Table 7 examines whether elite members' prior targeting experience affects the alignment between SPR and BR. Specifically, we compare villages in which at least three of six elite members had prior targeting experience with villages in which fewer than three members had such experience. We find no statistically significant evidence that higher prior targeting experience improves the concordance of either elites' rankings or aggregate rankings with expenditure- or asset-based benchmarks.

5. Conclusions

This study provides a comprehensive comparative analysis of alternative SPR methods and survey-based welfare and wealth measures. We administered three commonly used subjective ranking methods as part of a large household survey in Ethiopia: (i) self-ranking, (ii) peers' ranking, and (iii) elites' ranking. While these methods are widely applied in targeting social protection and welfare programs, a systematic comparison of their accuracy and concordance with conventional welfare outcomes remains limited in the existing literature. We evaluate the concordance as well as the potential sources of discordance between the three SPR and standard measures of consumption and wealth. We also construct and compare alternative methods of aggregating subjective rankings to improve their concordance with benchmark well-being indicators.

Our findings can be summarized as follows. First, subjective rankings generally exhibit weak concordance with conventional welfare and wealth measures, and they are only weakly correlated with each other. We also find that the three subjective ranking methods better predict household-level than per capita consumption, implying their failure to account for households' demographic size and composition properly, information that may not be easily observable. Second, among the three approaches, elite-based rankings best capture conventional measures of relative poverty, whereas self-rankings show the greatest discordance and a strong tendency to underestimate respondents' own positions. Third, and consistent with previous findings (Trachtman et al., 2025), subjective rankings are more aligned with wealth-based indicators than consumption-based measures. In our context, self, peers', and elites' rankings correlate more strongly with asset-based measures than with expenditure-based measures. This likely reflects the greater temporal stability of assets such as housing, land, and durable goods. In contrast, consumption expenditures fluctuate due to seasonal income, temporary shocks, or short-term financial variability, making them less observable

to respondents. Fourth, a composite ranking constructed by combining the three rankings into an aggregate rank only modestly improves the concordance with conventional indicators, especially asset-based rankings. Fifth, the performance of these subjective ranking methods varies across contexts (rural and urban areas, as well as across highlands and pastoralist communities) and in the presence and absence of covariate shocks. Interestingly, these subjective ranking methods better capture relative deprivation among households exposed to covariate shocks such as conflict and drought. We also find that these subjective ranking methods perform relatively poorly in capturing relative welfare and wealth in pastoral communities. Finally, we offer plausible explanations about the heterogeneities in the performance of these methods across contexts as well as major drivers of discordance between subjective and benchmark rankings, including information asymmetry, favoritism (in peers' and elites' rankings), and dynamic evolution of economic conditions of households.

Overall, our results provide important insights into the potential and limitations of SPR in capturing conventional welfare- and wealth-based ranking measures. The unique comparison of self, peer, and elite rankings, along with the analysis of the sources of discrepancy in these rankings, offers important lessons that can inform the targeting of welfare and social protection programs, especially in data-scarce contexts. For example, the generally weak concordance between subjective measures and conventional welfare-based ranking suggests that targeting based on these subjective measures would be prone to substantial inclusion and exclusion errors. However, the slightly better performance of elite-based ranking compared to the other methods justifies its relevance in actual targeting for data-scarce settings, especially given that implementing elite-based ranking appears to be the least costly. Similarly, the evidence that these subjective rankings are better correlated with conventional measures of deprivation during shocks (conflict and drought) is encouraging and confirms their potential and rationale for deploying them to identify those affected by covariate shocks. Finally, the evidence that these methods correlate better with overall consumption than per capita consumption can imply that they should only be used for household- and not individual-level targeting. The substantial deviations between subjective and conventional welfare-based rankings, as well as the heterogeneities across contexts and social connections, point to the usual criticisms related to elite capture and favoritism in elite-based targeting (Bardhan and Mookherjee, 2005; Alatas et al., 2019; Basurto et al., 2020; Vera-Cossio, 2022). Consistent with these patterns, we find suggestive evidence of favoritism toward closely related households, as well as information asymmetry associated with unknown households.

Finally, future research may investigate what subjective ranking methods actually capture. Similarly, examining the dynamics associated with these subjective rankings and their responsiveness to economic shocks and policy interventions merits further research. Longitudinal data on SPR are needed to provide valuable insights into the dynamics of poverty perceptions, the effectiveness of interventions, and the psychological and community-level impacts of alternative interventions. These insights are essential for effective utilization of these methods in the design of responsive social protection programs.

CRedit authorship contribution statement

Immacolata Ranucci: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Kibrom A. Abay:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization. **Luca Tiberti:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Conceptualization.

²² If elites have a greater aversion to exclusion errors, they may pay more attention to the ranking of those who are obviously perceived as needy in the village. On the other hand, if elites are more sensitive to inclusion error, they might pay more attention to ensure that undeserving households are not included or ranked as “needy”.

Declaration of competing interest

The author, Immacolata Ranucci, declares that she has no relevant or material financial interests that relate to the research described in this paper. In addition, no third party had the right to review this paper prior to its circulation.

The author, Luca Tiberti, declares that he has no relevant or material financial interests that relate to the research described in this paper. In addition, no third party had the right to review this paper prior to its circulation.

The author, Kibrom Abay declares that he has no relevant or material financial interests that relate to the research described in this paper. In addition, no third party had the right to review this paper prior to its circulation.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.jdeveco.2026.103943>.

Data availability

The dataset used in the analysis is included in the supplementary materials and is available in the replication package deposited in the following repository: <https://doi.org/10.5281/zenodo.22768864>.

References

- Abate, G.T., de Brauw, A., Gibson, J., Hirvonen, K., Wolle, A., 2022. Telescoping error in recalled food consumption: Evidence from a survey experiment in Ethiopia. *World Bank Econ. Rev.* 36 (4), 889–908.
- Abay, K.A., Berhane, G., Gilligan, D.O., Tafere, K., Taffesse, A.S., 2026. Targeting social assistance in fragile settings: An experiment in community-based targeting. *J. Dev. Econ.* 103819.
- Abay, K.A., Bevis, L.E., Barrett, C.B., 2021. Measurement error mechanisms matter: Agricultural intensification with farmer misperceptions and misreporting. *Am. J. Agric. Econ.* 103 (2), 498–522.
- Abay, K.A., Kahsay, G.A., Berhane, G., 2018. Social networks and factor markets: Panel data evidence from Ethiopia. *J. Dev. Stud.* 54 (1), 174–190.
- Abay, K.A., Tafere, K., Berhane, G., Chamberlin, J., Abay, M.H., 2023. Near-real-time welfare and livelihood impacts of an active war: Evidence from Ethiopia. *Food Policy* 119, 102526.
- Aiken, E., Ashraf, A., Blumenstock, J., Guiteras, R., Mobarak, A.M., 2025. Scalable Targeting of Social Protection: When Do Algorithms Out-Perform Surveys and Community Knowledge?.
- Akay, A., Martinsson, P., 2011. Does relative income matter for the very poor? Evidence from rural Ethiopia. *Econom. Lett.* 110 (3), 213–215.
- Akay, A., Martinsson, P., Medhin, H., 2012. Does positional concern matter in poor societies? Evidence from a survey experiment in rural Ethiopia. *World Dev.* 40 (2), 428–435.
- Alatas, V., Banerjee, A., Chandrasekhar, A.G., Hanna, R., Olken, B.A., 2016a. Network structure and the aggregation of information: Theory and evidence from Indonesia. *Am. Econ. Rev.* 106 (7), 1663–1704.
- Alatas, V., Banerjee, A., Hanna, R., Olken, B.A., Purnamasari, R., Wai-Poi, M., 2019. Does elite capture matter? Local elites and targeted welfare programs in Indonesia. In: *AEA Papers and Proceedings*. 109, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 334–339.
- Alatas, V., Banerjee, A., Hanna, R., Olken, B., Tobias, J., 2012. Targeting the poor: Evidence from a field experiment in Indonesia. *Am. Econ. Rev.* 102 (4), 1206–1240.
- Alatas, V., Purnamasari, R., Wai-Poi, M., Banerjee, A., Olken, B.A., Hanna, R., 2016b. Self-targeting: Evidence from a field experiment in Indonesia. *J. Political Econ.* 124 (2), 371–427.
- Alderman, H., 2002. Do local officials know something we don't? Decentralization of targeted transfers in Albania. *J. Public Econ.* 83 (3), 375–404.
- Alix-Garcia, J.M., Sims, K.R., Costica, L., 2021. Better to be indirect? Testing the accuracy and cost-savings of indirect surveys. *World Dev.* 142 (C), 105419.
- Alkire, S., Foster, J., 2011. Counting and multidimensional poverty measurement. *J. Public Econ.* 95 (7–8), 476–487.
- Aragie, E.A., Thurlow, J., 2022. Modeling the recovery dynamics of Ethiopia cattle population. *J. Arid. Environ.* 197, 104664.
- Bardhan, P., Mookherjee, D., 2005. Decentralizing antipoverty program delivery in developing countries. *J. Public Econ.* 89 (4), 675–704.
- Basurto, M.P., Dupas, P., Robinson, J., 2020. Decentralization and efficiency of subsidy targeting: Evidence from chiefs in rural Malawi. *J. Public Econ.* 185, 104047.
- Beegle, K., Himelein, K., Ravallion, M., 2012. Frame-of-reference bias in subjective welfare. *J. Econ. Behav. Organ.* 81 (2), 556–570.
- Berhane, G., Gilligan, D.O., Hoddinott, J., Kumar, N., Taffesse, A.S., 2014. Can social protection work in Africa? The impact of Ethiopia's Productive Safety Net Programme. *Econ. Dev. Cult. Chang.* 63 (1), 1–26.
- Bloch, F., Olckers, M., 2021. Friend-based ranking in practice. *AEA Pap. Proc.* 111, 567–571.
- Bottan, N.L., Perez-Truglia, R., 2022. Choosing your pond: location choices and relative income. *Rev. Econ. Stat.* 104 (5), 1010–1027.
- Brown, C., Ravallion, M., van de Walle, D., 2018. A poor means test? Econometric targeting in Africa. *J. Dev. Econ.* 134 (C), 109–124.
- Caeyers, B., Dercon, S., 2012. Political connections and social networks in targeted transfer programs: Evidence from rural Ethiopia. *Econ. Dev. Cult. Chang.* 60 (4), 639–675.
- Ferrer-i Carbonell, A., 2005. Income and well-being: an empirical analysis of the comparison income effect. *J. Public Econ.* 89 (5–6), 997–1019.
- Coady, D., Grosh, M.E., Hoddinott, J., 2004. Targeting of transfers in developing countries: Review of lessons and experience.
- Coll-Black, S., Gilligan, D.O., Hoddinott, J., Kumar, N., Taffesse, A.S., Wiseman, W., 2012. Targeting food security interventions in Ethiopia: The Productive Safety Net Programme. *Food Agric. Ethiop.: Prog. Policy Challenges* 280–317.
- Conning, J., Kevane, M., 2002. Community-based targeting mechanisms for social safety nets: A critical review. *World Dev.* 30 (3), 375–394.
- Cruces, G., Perez-Truglia, R., Tetaz, M., 2013. Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment. *J. Public Econ.* 98 (C), 100–112.
- Decerf, B., 2023. A preference-based theory unifying monetary and non-monetary poverty measurement. *J. Public Econ.* 222, 104898.
- Dercon, S., De Weerd, J., Bold, T., Pankhurst, A., 2006. Group-based funeral insurance in Ethiopia and Tanzania. *World Dev.* 34 (4), 685–703.
- Dupas, P., Fafchamps, M., Houeix, D., 2024. Eliciting poverty rankings from urban or rural neighbors: Methodology and empirical evidence. *Stanford Working Papers wp2068*, Stanford King Center on Global Development.
- Galasso, E., Ravallion, M., 2005. Decentralized targeting of an antipoverty program. *J. Public Econ.* 89 (4), 705–727.
- Ghigino, C., Goyal, S., 2010. Keeping up with the neighbors: Social interaction in a market economy. *J. Eur. Econ. Assoc.* 8 (1), 90–119.
- Gibson, J., Beegle, K., De Weerd, J., Friedman, J., 2015. What does variation in survey design reveal about the nature of measurement errors in household consumption? *Oxf. Bull. Econ. Stat.* 77 (3), 466–474.
- Gilligan, D.O., Hoddinott, J., Taffesse, A.S., 2009. The impact of Ethiopia's Productive Safety Net Programme and its linkages. *J. Dev. Stud.* 45 (10), 1684–1706.
- Hillebrecht, M., Klöner, S., A Pacere, N., Souares, A., 2020. Community-based versus statistical targeting of anti-poverty programs: Evidence from Burkina Faso. *J. Afr. Econ.* 29 (3), 271–305.
- Hillebrecht, M., Klöner, S., Pacere, N.A., 2023. The dynamics of poverty targeting. *J. Dev. Econ.* 161, 103033.
- Hussam, R., Rigol, N., Roth, B.N., 2022. Targeting high ability entrepreneurs using community information: Mechanism design in the field. *Am. Econ. Rev.* 112 (3), 861–898.
- Hvidberg, K., Kreiner, C.T., Stantcheva, S., 2021. Social position and fairness views. *NBER Working Paper 28099*, National Bureau of Economic Research.
- Hvidberg, K.B., Kreiner, C.T., Stantcheva, S., 2023. Social positions and fairness views on inequality. *Rev. Econ. Stud.* 90 (6), 3083–3118.
- Karadja, M., Møllerstrom, J., Seim, D., 2017. Richer (and holier) than thou? The effect of relative income improvements on demand for redistribution. *Rev. Econ. Stat.* 99 (2), 201–212.
- Kingdon, G.G., Knight, J., 2007. Community, comparisons and subjective well-being in a divided society. *J. Econ. Behav. Organ.* 64 (1), 69–90.
- Kosec, K., Mo, C.H., 2024. Does relative deprivation condition the effects of social protection programs on political support? Experimental evidence from Pakistan. *Am. J. Polit. Sci.* 68 (2), 832–849.
- Lee, S., Abay, K.A., Barrett, C.B., Hoddinott, J., 2025. Estimating multidimensional development resilience. *J. Dev. Econ.* 103583.
- Ligon, E., Schechter, L., 2003. Measuring vulnerability. *Econ. J.* 113 (486), C95–C102.
- Lind, J., Sabates-Wheeler, R., Hoddinott, J., Taffesse, A.S., 2022. Targeting social transfers in Ethiopia's agro-pastoralist and pastoralist societies. *Dev. Chang.* 53 (2), 279–307.
- Luttmer, E.F., 2005. Neighbors as negatives: Relative earnings and well-being. *Q. J. Econ.* 120 (3), 963–1002.
- Nigus, H.Y., Abay, K.A., 2024. Cooperation Among Community Leaders: The Role of Women's Leadership and Exposure to Conflict. *International Food Policy Research Institute*.
- Premand, P., Schnitzer, P., 2021. Efficiency, legitimacy, and impacts of targeting methods: Evidence from an experiment in Niger. *World Bank Econ. Rev.* 35 (4), 892–920.
- Rai, A.S., 2002. Targeting the poor using community information. *J. Dev. Econ.* 69 (1), 71–83.
- Raleigh, C., Linke, A., Hegre, H., Karlsen, J., 2010. Introducing ACLED: An Armed Conflict Location and Event Dataset: Special data feature. *J. Peace Res.* 47 (5), 651–660.

- Ravallion, M., 2008. Miss-targeted or miss-measured? *Econom. Lett.* 100 (1), 9–12.
- Ravallion, M., 2012. Poor, or just feeling poor? On using subjective data in measuring poverty. In: Clark, A.E., Senik, C. (Eds.), *Happiness and Economic Growth: Lessons from Developing Countries*. Oxford Academic, Oxford, pp. 140–178.
- Ravallion, M., 2020. On measuring global poverty. *Annu. Rev. Econ.* 12 (1), 167–188.
- Ravallion, M., Chen, S., 2011. Weakly relative poverty. *Rev. Econ. Stat.* 93 (4), 1251–1261.
- Ravallion, M., Chen, S., 2019. Global poverty measurement when relative income matters. *J. Public Econ.* 177, 104046.
- Ravallion, M., Himelein, K., Beegle, K., 2016. Can subjective questions on economic welfare be trusted? *Econ. Dev. Cult. Chang.* 64 (4), 697–726.
- Ravallion, M., Lokshin, M., 2002. Self-rated economic welfare in Russia. *Eur. Econ. Rev.* 46 (8), 1453–1473.
- Ravallion, M., Lokshin, M., 2010. Who cares about relative deprivation? *J. Econ. Behav. Organ.* 73 (2), 171–185.
- Sabates-Wheeler, R., Lind, J., Hoddinott, J., 2013. Implementing social protection in agro-pastoralist and pastoralist areas: How local distribution structures moderate PSNP outcomes in Ethiopia. *World Dev.* 50, 1–12.
- Trachtman, C., Permana, Y., Sahadewo, G., 2025. How much do our neighbors really know? The limits of community-based targeting. *J. Dev. Econ.* 103555.
- Vera-Cossio, D., 2022. Targeting credit through community members. *J. Eur. Econ. Assoc.* 20 (2), 778–821.
- Wossen, T., Abay, K.A., Abdoulaye, T., 2022. Misperceiving and misreporting input quality: Implications for input use and productivity. *J. Dev. Econ.* 157, 102869.