Policy Research Working Paper 11348

Multidimensional Poverty

Why Not Make Up the Missing Joint Distribution Data?

Benoit Decerf
Mery Ferrando
Balint Menyhert

WORLD BANK GROUP
Development Economics
Development Research Group
April 2026

Policy Research Working Paper 11348

Abstract

Poverty is inherently multidimensional, encompassing both monetary and non-monetary dimensions. However, these outcomes are often collected in separate surveys, leaving the joint distribution partially unobserved. To improve social poverty comparisons, this paper proposes a new, simple method to address this data constraint: assume a fixed value for the missing part of the joint distribution. This approach allows the integration of outcomes collected from different surveys, unlike the mainstream method currently in use. Drawing on household surveys from six developing countries where both dimensions are observed, the paper shows that the method systematically outperforms traditional single-survey measures and "mash-up" measures. Monte Carlo simulations further confirm the robustness of the results across a wide range of data-generating scenarios. The findings highlight the value of the proposed method for monitoring multidimensional poverty and suggest that it may also benefit other social indicators facing similar data limitations.

This paper is a product of the Development Research Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at benoit.decerf@unamur.be. A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click **here** for direct access.

Multidimensional Poverty: Why not Make Up the Missing Joint Distribution Data?*

Benoit Decerf,† Mery Ferrando,‡ and Balint Menyhert§

Authorized for distribution by Dean Jolliffe, Research Manager, Development Research Group, World Bank Group

JEL: I32, I38, C81, O15.
Keywords: Multidimensional Poverty Measurement; Social Comparisons; Data Constraints; Imputation Methods.

*Acknowledgments: We are grateful to Ricardo Nogales, with whom we have had extensive discussions throughout this project. We thank seminar audiences for their helpful comments at the World Bank, particularly Paola Ballon, Jed Friedman, and Berk Ozler, and at the University of Antwerp, particularly Koen Decancq, and Bea Cantillon. This project benefited from an RSB grant (P508097) from the World Bank Research Group in 2024 as well as from the GSCP1 initiative of the Poverty and Equity GP of the World Bank in 2023. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and should not be attributed in any manner to the World Bank, to its affiliated organizations, or to members of its Board of Executive Directors or the countries they represent. All errors remain ours.

†University of Namur. Email: benoit.decerf@unamur.be.
‡Tilburg University, Department of Economics. Email: m.ferrando@tilburguniversity.edu.
§Bank of Estonia. Email: balint.menyhert@eestipank.ee

1 Introduction

Poverty, like well-being, is widely recognized as a multidimensional phenomenon encompassing both monetary and non-monetary dimensions (Stiglitz et al., 2009). An individual's poverty status may thus depend on whether or not she cumulates deprivations across several dimensions. However, a common data limitation is that outcomes in different dimensions are often collected through separate surveys. For example, at the global level, outcomes in key non-monetary dimensions such as health are typically not gathered in the same surveys that measure monetary outcomes.[1] As a result, while we can observe partial distributions for each dimension separately, the joint distribution of individual-level outcomes is usually missing. This lack of information on dependence across dimensions can seriously distort poverty comparisons. Intuitively, two societies might display identical marginal rates of monetary and non-monetary deprivation but differ substantially in multidimensional poverty depending on whether these deprivations fall on the same individuals or on different groups (Decancq, 2023). The methods used to address this data constraint are therefore crucial, as they determine the extent to which poverty measures allow for appropriate social comparisons. This is particularly important because poverty measures are used by international institutions and national governments to monitor progress and allocate resources across regions.

The literature has developed three main approaches to deal with the missing joint distribution data.[2] Currently, the mainstream solution is to restrict the set of dimensions to a subset whose outcomes are all collected in a single survey. Prominent examples include monetary poverty measures as well as the main global multidimensional poverty measures, such as the Global MPI of UNDP-OPHI. These "single-survey" measures enjoy credibility because all their necessary data is observed, but they ignore the dimensions not captured in their survey. The second solution allows combining data from several surveys, but adopts an index whose definition does not depend on the unobserved dependence. Typically, this implies adopting a "mash-up" index that combines statistics from each partial distribution, an approach that is easy to compute but widely criticized for its weak conceptual foundations (Fleurbaey, 2009; Ravallion, 2011). Mash-up poverty measures, such as the Human Poverty Index (Watkins, 2006), are rarely used in practice. The third solution is to estimate the missing part of the joint distribution through survey-to-survey imputation techniques (Christiaensen et al., 2012; Dang and Lanjouw, 2023). Although appealing, their complexity makes them infrequently used.

[1] Evidence from subjective well-being research shows that both income and health are among the strongest predictors of subjective well-being (Diener et al., 2018).

[2] A fourth solution could be provided by the total consumption approach, for example, as used by Gethin (2025) to study global inequality. An individual's "total consumption" is obtained by adding a monetary valuation of their non-monetary outcomes to their monetary consumption. One simple way to address the missing data problem could be to assume that all individuals living in the same area share the same non-monetary outcomes. However, this approach is not yet fully mature for the purpose of poverty measurement, as it still faces complex challenges, such as the difficulty of valuing non-monetary outcomes.

In this paper, we propose a new solution to the problem of missing joint distribution data that (i) allows the use of partial distribution data from separate surveys, (ii) is simple to implement, and (iii) does not require the use of conceptually weak poverty indices. In a nutshell, our proposal is to assume (i.e., "make up") a fixed value for the missing joint distribution data. We call "made-up" the measures obtained in this way. In the special case where the missing dependence can be summarized by a one-dimensional coefficient, our solution simply assumes a fixed value for the coefficient and applies it uniformly across units. Our approach can be viewed as a simplified version of survey-to-survey imputation, in which missing values are assumed rather than estimated.

For pedagogical purposes, we select a particular multidimensional headcount ratio as our benchmark poverty measure. Individuals are identified as multidimensionally poor when they are monetary poor and/or non-monetary poor. These latter two statuses can be determined from the partial distributions respectively observed in the monetary and non-monetary surveys. The multidimensional headcount ratio is a function of three statistics, namely the monetary headcount ratio, the non-monetary headcount ratio, and a coefficient of overlap between these two statuses, whose value is unobserved. The monetary (respectively, non-monetary) headcount ratio is a single-survey measure. The multidimensional headcount ratio obtained for a fixed value of the coefficient of overlap is a made-up measure. Finally, a weighted sum of the monetary headcount ratio and the non-monetary headcount ratio yields a mash-up measure. We use three alternative criteria to evaluate how close these measures' social poverty comparisons are to "ideal" comparisons.

We empirically study which of these solutions (single-survey, mash-up, and made-up measures) yields less biased social poverty comparisons. We draw on rich datasets from six countries (Bolivia, Brazil, Ecuador, Ethiopia, Ghana, and Uganda) where both monetary and non-monetary outcomes are observed for the same households (Evans et al., 2024). These data allow us to construct the "ideal" measure based on fully observed joint distributions and to compare it directly with constrained alternatives, including single-survey, mash-up, and our made-up measures. The six countries differ widely in geography, development levels, and poverty profiles, allowing us to test the approach across very different contexts.

The empirical results show that single-survey measures are outperformed by both the mash-up and the made-up measures. Surprisingly, made-up measures outperform single-survey measures regardless of the value assumed for the overlap coefficient. The benefit of integrating partial distributions thus outweighs the cost of assuming an incorrect value for the missing part of the joint distribution, even under implausible assumptions. Interestingly, assuming the value corresponding to independence yields an almost optimal performance. This has practical relevance because the independence case is simpler (see Section 5). Our results also show that made-up measures based on plausible assumptions enjoy some advantages over mash-up measures: the former typically slightly outperform the latter in most cases. Our findings are robust to alternative poverty cutoffs and to the use of a depth-sensitive poverty index.

To further assess the generality of these findings beyond the six countries for which we have data, we conduct a series of Monte Carlo simulations that systematically vary key data-generating parameters, namely the variance of the monetary and non-monetary poverty measures, the variance of the coefficient of overlap, the correlation between the monetary and non-monetary poverty measures, and the mean absolute gap between these two measures. Across all scenarios, the simulations confirm that made-up measures more closely approximate the ideal multidimensional benchmark than either single-survey or mash-up alternatives. This outperformance is most pronounced when poverty rates vary widely across regions, but the dependence between dimensions remains relatively stable. Yet, we still find benefits in the opposite scenarios. Even under adverse conditions, such as when the correlation between dimensions is very high or when one dimension dominates, the made-up measures perform at least as well as single-survey measures for most assumed values of the overlap coefficient. Overall, the simulation exercise demonstrates that assuming a fixed overlap coefficient provides a robust and low-cost strategy for improving social poverty comparisons across a wide range of empirically relevant contexts.

Our findings provide the following insights about the solutions to split-survey data constraints. The dominant single-survey measures are suboptimal for social poverty comparisons. Integrating more dimensions from separate surveys is valuable, even when this integration is done imperfectly. As a result, the lack of conceptual foundations for mash-up poverty measures is not a sufficient reason to discard them in favor of single-survey measures. (Our results still suggest using made-up measures rather than mash-up measures.) Similarly, while one can always question whether the context in which survey-to-survey imputation models are trained is sufficiently similar to the contexts where they are applied (Mathiassen and Wold, 2021; Mathiassen, 2013; Newhouse et al., 2014), our results again suggest that such external validity concerns are not a sufficient reason to disqualify these imputations. Indeed, to the extent that their estimation of the missing joint distribution data is better than assuming a fixed value, arguably a low bar, this approach will outperform made-up measures and thus the other alternatives.

While our application focuses on poverty, the method we develop is more general. It is adaptable to a wide range of multidimensional settings, where data on different dimensions are rarely available in a single source, provided that the measure aggregates dimensions at the individual level before aggregating across individuals. This opens avenues for future work in multidimensional welfare analysis and highlights the framework's value as a general strategy for combining fragmented data into robust multidimensional indicators.

Our paper contributes to two main strands of literature. The first is the literature on multidimensional poverty measurement itself. Since Sen's early contribution (1976), a large axiomatic tradition has examined how to identify and aggregate poverty across dimensions

(Tsui, 2002; Bourguignon and Chakravarty, 2003; Duclos et al., 2006; Chakravarty et al., 2008; Alkire and Foster, 2011; Decancq et al., 2019; Piacquadio, 2026). Much of this literature, however, assumes that the joint distribution of outcomes is observed in a single dataset. In practice, this is rarely the case. Our paper extends this literature by asking how poverty can be assessed when data on relevant dimensions come from separate surveys, how alternative solutions can be contrasted, and by proposing a simple, feasible approach that allows using poverty indices with solid conceptual foundations.

Second, we contribute to an important line of work within the broader poverty measurement literature that has focused on synthetic "mash-up" indices. These are typically constructed as weighted averages of several statistics, each summarizing the partial distribution of a specific dimension. Well-known examples include the Human Development Index (UNDP, 1990) and the Human Poverty Index (UNDP, 1997). These measures are often criticized for their lack of clear conceptual foundations (Fleurbaey, 2009; Ravallion, 2011). Our paper relates to this discussion by showing that it is possible to retain the simplicity and communicability of synthetic indices while anchoring them in the structure of a well-grounded poverty index. In other words, we show that it is possible to build synthetic indices that are tractable yet grounded in theory.

The remainder of the paper is organized as follows. Section 2 develops the theoretical framework, introducing the ideal poverty measure considered, the proposed multi-survey measures, and the criteria used to evaluate them. Section 3 turns to the empirical analysis: it covers data and parameter choices, descriptive statistics, the main results comparing single- and multi-survey measures, and robustness checks to alternative poverty cutoffs and depth-sensitive poverty indices. Section 4 reports simulation exercises designed to assess the generality of our findings under alternative data-generating scenarios. Section 5 concludes.

2 Framework

Let $N = \{1, \dots, n\}$ denote the set of individuals in a given population. For each $i \in N$, let $x_i = (x_{i1}, \dots, x_{i\ell}) \in X$ denote the vector of poverty-relevant outcomes for individual i. Assume, for simplicity, that the first k variables in vector x capture monetary outcomes, while the remaining $\ell - k$ variables capture non-monetary outcomes, i.e.,

$$x_i = (\underbrace{x_{i1}, \dots, x_{ik}}_{monetary}, \underbrace{x_{ik+1}, \dots, x_{i\ell}}_{non-monetary}).$$

Let $x_i^M = (x_{i1}, \dots, x_{ik})$ and $x_i^{NM} = (x_{ik+1}, \dots, x_{i\ell})$ denote, respectively, the vectors of monetary and non-monetary outcomes, so that $x_i = (x_i^M, x_i^{NM})$. Let $x = (x_i)_{i \in N}$ denote the distribution of outcomes in the population.

Let P denote a subgroup-decomposable poverty measure defined as $P(x) = \frac{1}{n} \sum_{i \in N} p(x_i)$, where $p : X \to [0,1]$ is the contribution function, which aggregates an individual's outcomes into a poverty score. The main poverty index used in practice is the headcount ratio, defined as

$$H(x) = \frac{1}{n} \sum_{i \in N} \mathbb{1}(x_i)$$

where the poverty identification function $\mathbb{1} : X \to \{0,1\}$ is such that $\mathbb{1}(x_i) = 1$ when i is identified as poor and $\mathbb{1}(x_i) = 0$ otherwise.

2.1 Data Constraint and Alternative Solutions

We consider the data constraint that arises when no survey captures *all* the poverty-relevant outcomes. At the global level, for instance, monetary and non-monetary outcomes are typically collected through separate surveys. LSMS surveys collect monetary outcomes while DHS or MICS surveys collect non-monetary outcomes. For a given society, these surveys respectively provide $x^M = (x_i^M)_{i \in N'}$ and $x^{NM} = (x_i^{NM})_{i \in N''}$ for two different populations, i.e., $N' \neq N''$.

This data constraint implies that we lack information on how monetary and non-monetary outcomes are jointly distributed in the society. Thus, we do not know to what extent individuals with low monetary outcomes are also more (or less) likely to experience low non-monetary outcomes compared to individuals with high monetary outcomes. Crucially, such data do not allow computing poverty measure P. However, they do allow computing *"constrained" poverty measures* $\tilde{P}$, which do not use more (and may use less) data than the information available in $(x_i^M)_{i \in N'}$ and $(x_i^{NM})_{i \in N''}$.

The mainstream solution to this data constraint is to consider "*single-survey*" poverty measures, which focus on one type of poverty-relevant outcomes. At the global level, the World Bank monitors monetary poverty measures $P_M(x^M)$ based on LSMS-type surveys, while UNDP-OPHI monitors non-monetary poverty measures $P_{NM}(x^{NM})$ based primarily on DHS or MICS surveys. For instance, the monetary headcount ratio is defined as $H_M(x^M) = \frac{1}{n'} \sum_{i \in N'} \mathbb{1}^M(x_i^M)$, where the monetary poverty identification function checks whether the individual's monetary welfare is below a monetary poverty line, namely

$$\mathbb{1}^M(x_i^M) = \begin{cases} 1 & \text{if } m(x_i^M) < z \\ 0 & \text{otherwise} \end{cases}$$

where $z > 0$ is the monetary poverty line and $m(x_i^M)$ is the monetary welfare (income or consumption) of individual i computed from her monetary outcomes. Following Alkire and Foster (2011), the non-monetary headcount ratio is defined as $H_{NM}(x^{NM}) = \frac{1}{n''} \sum_{i \in N''} \mathbb{1}^{NM}(x_i^{NM})$, where the non-monetary poverty identification function checks whether the individual's weighted sum of deprivations exceeds a cutoff (Seth and Santos, 2019), namely

$$\mathbb{1}^{NM}(x_i^{NM}) = \begin{cases} 1 & \text{if } \sum_{j=k+1}^{\ell} w_j x_{ij}^{NM} \geq k \\ 0 & \text{otherwise} \end{cases}$$

where $k \in [0, 1]$ is the non-monetary poverty cutoff, $x_{ij}^{NM} = 1$ if i is deprived in variable j and $x_{ij}^{NM} = 0$ otherwise (i.e., non-monetary variables are binary), $w_j \in [0, 1]$ is the weight assigned to being deprived in variable j, and $\sum_{j=k+1}^{\ell} w_j = 1$.

Poverty measures serve two types of purposes: interpersonal comparisons and social comparisons. First, they allow *interpersonal poverty comparisons*. Measure P indicates when individual i is poorer than individual i', i.e., when $p(x_i) > p(x_{i'})$. This is, for example, the case when i is identified as poor while i' is not. Interpersonal comparisons are used in practice to target social protection benefits, for example, by targeting benefits for individuals identified as poor.[3] Second, they allow *social poverty comparisons*. Measure P indicates when a society whose outcome distribution is x is poorer than another society whose outcome distribution is y, i.e., when $P(x) > P(y)$. Social comparisons are necessary to *monitor progress* against poverty, in which case y is the outcome distribution of the same society but at a later time than x. Monitoring progress is the cornerstone of policy impact evaluation. Social comparisons are also necessary for across-space comparisons, in which case y is the outcome distribution of a different society than x at the same time period. Across-space comparisons are necessary for the *allocation of poverty alleviation efforts* across countries. For instance, should the World Bank prioritize this country or the other? Similarly, a national government may decide to allocate part of its budget across its regions based on needs. Each region could, for example, receive a fixed amount per poor individual and then decide on how to spend it, taking into account the region's main poverty issues.

We provide evidence that, for the purpose of making *social poverty comparisons*, the mainstream solution of considering single-survey poverty measures is suboptimal. Social comparisons should instead be based on "*multi-survey*" poverty measures, which account for both types of poverty-relevant outcomes: monetary and non-monetary. In general, a multi-survey poverty measure $\hat{P}$ uses the outcomes collected in the two separate surveys, so that

$$\hat{P}\left(M\left((x_i^M)_{i\in N'}, (x_i^{NM})_{i\in N''}\right)\right)$$

for some mapping M. Our goal is to assess whether multi-survey measures perform better than single-survey measures. More precisely, we examine whether social comparisons based

[3]It is important that the government be able to provide an objective reason why one household can claim benefits while another cannot. This justification should not ultimately rely on outcomes observed for a third household. For this reason, we do not necessarily recommend our solution for the purpose of interpersonal comparisons.

on $\hat{P}$ more closely approximate those based on P than do social comparisons based on P_M or P_{NM}. Showing this requires combining answers to three different conceptual questions, namely

$Q1:$ What is the "ideal" poverty measure P in the absence of data constraints?

$Q2:$ Which multi-survey poverty measure(s) $\hat{P}$ to consider (for performance comparison with single-survey measures P_M and P_{NM})?

$Q3:$ What criteria are used to capture how *close* social comparisons with a constrained measure $\tilde{P}$ are to those with the ideal measure P?

We require that our answers to these questions are chosen conservatively or that our empirical results are robust to the particular choices made. We next present, in turn, the answers we use for these three questions.

2.2 Ideal Poverty Measure

Regarding the first question, we select an "ideal" poverty measure that is both pedagogical and plausible. We later test the robustness of our results using an alternative measure that does not suffer from its main limitation.

In the absence of data constraint, we select as *ideal measure* the multidimensional headcount ratio H for which the multidimensional poverty identification function determines whether the individual is either monetary poor and/or non-monetary poor, namely[4]

$$\mathbb{1}(x_i) = \begin{cases} 1 & \text{if } \mathbb{1}^M(x_i^M) = 1 \text{ and/or } \mathbb{1}^{NM}(x_i^{NM}) = 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Ideal measure H is plausible for two reasons. First, the headcount ratio is the most well-known poverty index and is widely used in applications.[5] Second, the multidimensional poverty identification is based on the union approach, so that any individual that is monetary (respectively, non-monetary) poor is automatically identified as multidimensionally poor. This is the approach used by the Multidimensional Poverty Measures of the World Bank, which combine monetary outcomes with a small subset of non-monetary indicators that are collected in most LSMS surveys.

More importantly, measure H is pedagogical for our purposes. The reason is that H only depends on the distribution in the population of the binary "variable" $\mathbb{1}(x_i)$. Therefore, by definition of the identification function in Equation (1), H only depends on the

[4] In our empirical application, poverty is identified at the household level; all individuals within a household are therefore assigned the same poverty status.

[5] We note, however, that it is common for non-monetary poverty measures to instead use the so-called adjusted headcount ratio, which not only accounts for the incidence but also for the intensity of non-monetary poverty, as it is depth-sensitive. We also show that our results are robust to the use of such a depth-sensitive measure.

joint-distribution of two binary "variables", namely $\mathbb{1}^M(x_i^M)$ and $\mathbb{1}^{NM}(x_i^{NM})$. This joint-distribution is fully characterized by the two partial distributions for variables $\mathbb{1}^M(x_i^M)$ and $\mathbb{1}^{NM}(x_i^{NM})$ together with the *one-dimensional* dependence (at the individual level) between these two variables (see below). Under the data constraint implied by separate surveys, the two partial distributions are observed, while the dependence parameter is not observed. The two partial distributions can respectively be summarized by $H_M(x^M)$ and $H_{NM}(x^{NM})$, which are natural single-survey poverty measures when the ideal indicator is H (because they rely on the same poverty index). Hence, we can show that H is a function of three conceptually clear factors, namely

$$H(x) = f\left(H_M(x^M), H_{NM}(x^{NM}), c(x)\right) \tag{2}$$

where the first two arguments summarize the observed partial distributions and the third argument $c(x) \in [-1, 1]$ summarizes the part of the joint distribution of outcomes that is missing due to the data constraint, namely, the relevant dependence between monetary and non-monetary individual-level outcomes.

We call factor $c(x)$ the *coefficient of overlap* between monetary and non-monetary poverty statuses. In general, the value for c tunes the extent to which monetary poor individuals are more likely to be non-monetary poor than their counterparts. The value $c(x) = 0$ corresponds to the case in which the two statuses are independent. In turn, $c(x) = 1$ corresponds to the case with maximally positive dependence between these two statuses, and $c(x) = -1$ to the situation with maximally negative dependence. The two separate surveys contain all the necessary data to compute the values for $H_M(x)$ and $H_{NM}(x)$, but they do not provide the necessary information to compute the value taken by $c(x)$. The coefficient c captures the individual-level overlap between the binary poverty statuses; it is not the correlation between the levels of H_M and H_{NM} across societies.

In the next paragraphs, we provide some intuition and the exact definition of the coefficient of overlap. It should be noted that we do not claim that its definition is the only valid way to capture the dependence between variables $\mathbb{1}^M(x_i^M)$ and $\mathbb{1}^{NM}(x_i^{NM})$, nor that it is the "best" way. We merely claim that its definition is fit for our purpose, which is to capture dependence through a one-dimensional variable.

Some intuition for the definition of the coefficient c can be gained from the following thought experiment. Let H_M and H_{NM} be two percentage values for the monetary and non-monetary poverty rates. The coefficient of overlap determines how to draw a number of $H_{NM} \times 100$ balls from two urns

- An "M-poor" urn with $H_M \times 100$ balls,
- An "M-rich" urn with $(1 - H_M) \times 100$ balls,

in such a way that the number of balls drawn from the M-poor urn is equal to the percentage of individuals who are *both* monetary and non-monetary poor. This implies the number of balls drawn from the M-rich urn is equal to the percentage of individuals who are non-monetary poor and *not* monetary poor. For the independence case, $c = 0$, a fraction H_{NM} of its balls should be drawn for each of the two urns. For the maximally positive dependence case $c = 1$, all the balls must be drawn from the M-poor urn, and if there are not enough balls in the M-poor urn, the remaining number of balls ($= (H_{NM} - H_M) \times 100$) must be drawn from the M-rich urn. For the maximally negative dependence case $c = -1$, all the balls must be drawn from the M-rich urn, and if there are not enough balls in the M-rich urn, the remaining number of balls ($= (H_{NM} - (1 - H_M)) \times 100$) must be drawn from the M-poor urn. Intermediate values for c are defined linearly between these two extremes.

Mathematically, for a given distribution x, the formal definition of c, which is a bit involved,[6] is given by

$$c = \begin{cases} \frac{\frac{H_M + H_{NM} - H}{H_M} - H_{NM}}{\min(1, H_{NM}/H_M) - H_{NM}} & \text{if } H_{NM} \leq \frac{H_M + H_{NM} - H}{H_M} \\ \\ \frac{\frac{H_M + H_{NM} - H}{H_M} - H_{NM}}{H_{NM} - \max(0, (H_M + H_{NM} - 1)/H_M)} & \text{else} \end{cases} \tag{3}$$

where the first line corresponds to values $c \in [0, 1]$ and the second line to values $c \in [-1, 0)$ and where we dropped the dependence to x in the notation for the sake of clarity. The reason for using this complex coefficient to capture dependence is that it preserves a common support of $[-1, 1]$ for all possible pairs of values of H_M and H_{NM},[7] which simplifies the analysis. Equation (3) provides the function f' such that

$$c(x) = f'\left(H_M(x^M), H_{NM}(x^{NM}), H(x)\right)$$

Inverting function f' provides function f in Equation (2), namely

$$H = \begin{cases} H_M + H_{NM} - H_M \times [H_{NM} + c \times (\min(1, H_{NM}/H_M) - H_{NM})] & \text{if } c \geq 0 \\ \\ H_M + H_{NM} - H_M \times [H_{NM} + c \times (H_{NM} - \max(0, (H_M + H_{NM} - 1)/H_M))] & \text{else} \end{cases} \tag{4}$$

With the coefficient of overlap formally defined, we now turn to the limitations of the ideal measure H. Clearly, H suffers from several well-known limitations (Sen, 1976). Arguably, the main limitation is that H is not sensitive to the depth of poverty. That is, two poor individuals contribute the same amount to H, even when one is much poorer

[6] Some intuition for Equation (3) is provided in Appendix A.1.

[7] In other words, for any pair $(H_M, H_{NM}) \in [0, 100\%] \times [0, 100\%]$ and any $c \in [-1, 1]$, if all three values are rational numbers, then there exists some outcome distribution x such that $H_M(x^M) = H_M$, $H_{NM}(x^{NM}) = H_{NM}$ and $c(x) = c$.

than the other. We do *not* claim that H would be the best possible poverty measure in a world without data constraints. Besides the limitation just mentioned, one could also criticize the explicit or implicit weights assumed for the different outcomes. However, the question of the best poverty measure in the absence of data constraints is not the topic of this paper. Rather, we focus on the best way to deal with the data constraint once an ideal measure has been selected. Importantly, we show in section 3.4.2 that our empirical results are robust to an alternative depth-sensitive choice for the ideal measure.

2.3 Multi-Survey Poverty Measures

We now turn to our second question on constrained multi-survey poverty measures. From the data constraint, they must be computable from two separate surveys, one collecting monetary outcomes and the other collecting non-monetary outcomes.

Two Simple Multi-Survey Measures

We consider two different measures. First, we compute a measure that is the simple weighted average of the two single-survey poverty measures. Assuming equal weights, as is conventional for a *mash-up measure*, we have

$$\hat{H}^{1/2}(x^M, x^{NM}) = \frac{1}{2}H_M(x^M) + \frac{1}{2}H_{NM}(x^{NM}). \tag{5}$$

The idea of a summary measure weighting several poverty measures, each capturing a different dimension, is not new. One example is the Human Poverty Index proposed by Watkins (2006). Ravallion (2011) criticizes such ad-hoc aggregation of different dimensions. Clearly, mash-up measures typically lack credible conceptual foundations for their mathematical expression (as well as for the value selected for their weight). As a result, their social comparisons are treated with caution. Also, their conceptual limitations may explain why mash-up measures are seldom adopted for official purposes, although there are well-known exceptions, such as the Human Development Index (Anand and Sen, 1994).

Second, we propose a new type of multi-survey poverty measures. This new type does not require adopting a mash-up poverty index, whose mathematical expression is explicitly independent of the missing part of the joint distribution. Rather, the idea is to use the ideal poverty index but "make up" the missing joint distribution data, that is, the dependence between monetary and non-monetary outcomes. We therefore assume how the two types of outcomes co-vary across individuals in the considered society.[8] Yet we are not necessarily interested in a full micro-level matching of individuals across surveys. This distinction is important: our objective is not to reconstruct missing individual-level outcomes, but

[8] For brevity, we use terms such as "missing part of the joint distribution," "dependence," or "overlap" to refer to this association.

to capture the aggregate (individual-level) association between dimensions that shapes multidimensional poverty at the society-level.

In the case of our ideal poverty measure H, the missing data is simply the one-dimensional coefficient of overlap, as explained in Section 2.2. Our proposed solution is to arbitrarily select a fixed value for the coefficient of overlap, which we denote by $\hat{c} \in [-1, 1]$.[9] We thus get a family of measures, which we call *made-up measures*, parametrized by the value selected for $\hat{c}$, defined as

$$\hat{H}_{\hat{c}}(x^M, x^{NM}) = f\left(H_M(x^M), H_{NM}(x^{NM}), \hat{c}\right) \tag{6}$$

where function f is defined in Equation (4). Survey-to-survey imputations also allow preserving the ideal poverty index, as we explain below. However, our approach is arguably much simpler than survey-to-survey imputations. Our approach is also readily generalizable to more complex ideal poverty measures, where the missing part of the joint distribution is not captured by a one-dimensional coefficient. For these more complex measures, one could assume a fixed dependency structure through a copula-based approach (Decancq and Jorda, 2024).

By definition, made-up measures inherit both the mathematical expression and the weights of the ideal multidimensional poverty index. Hence, if the ideal index rests on sound conceptual foundations, the associated made-up measure does as well, although some properties of the ideal index may be lost (as discussed below). This stands in sharp contrast to mash-up measures, which simply average single-survey indices without a clear normative justification. While mash-ups provide a straightforward way to combine information from multiple surveys, their ad hoc functional form limits their theoretical credibility and partly explains why they are seldom used in practice. Made-up measures, by contrast, remain anchored in the ideal index: their structure and interpretation are preserved, with the only approximation being that the missing dependence between dimensions is assumed rather than observed. As a result, they offer stronger conceptual legitimacy and may therefore be more acceptable to practitioners, despite their reliance on an assumed rather than observed dependence structure.

An important observation concerns the properties that the made-up measures do not inherit from the ideal measure. Consider, for instance, a "Multidimensional Transfer" property requiring that the poverty measure records a worsening after an *association-increasing rearrangement* (Atkinson and Bourguignon, 1982), which does not change the number of deprivations experienced in the population but concentrates these deprivations on fewer individuals. Clearly, the made-up measures violate "Multidimensional Transfer" even when their underlying ideal measure satisfies this property. The made-up measures cannot be sensitive to such rearrangement because it assumes a fixed value for the missing

[9] We suggest using a *fixed* value for simplicity, noting that this choice also has benefits, as shown in Proposition 1.

joint distribution data. Many other properties that also rely on this missing data are similarly lost by the made-up measures. Even the basic "Focus" property could be lost, at least when the ideal identification of the multidimensionally poor relies on the missing data.[10] We note, however, that these properties are bound to be lost by all constrained poverty measures because the necessary data are missing. Now, the value of these ordinal properties lies in their implications for poverty comparisons. Therefore, the violation of these properties is less severe for constrained measures whose poverty comparisons are closer to those made by the ideal measure, which is precisely what we investigate in our empirical results and simulations.[11]

More Sophisticated Multi-Survey Measures

There are several alternative ways to refine our multi-survey poverty measures. One strategy would be to impute monetary welfare (income or consumption) from the monetary survey into the non-monetary survey. Such an imputation model could be constructed in a dataset containing all poverty-relevant outcomes. Variables that are common to the two surveys can also provide useful information for the imputation (Christiaensen et al., 2012; Dang and Lanjouw, 2023). The purpose of this strategy is to construct a "synthetic" dataset that contains all poverty-relevant outcomes for each individual. The ideal poverty index can then be computed in the synthetic dataset obtained. This strategy consists in estimating, rather than assuming, the missing joint distribution data. For the ideal measure H, a value $\bar{c}$ is estimated instead of assuming a fixed $\hat{c}$, yielding a multi-survey measure $\overline{H}_{\bar{c}}$ for which the value for $\bar{c}$ is society-specific, unlike $\hat{H}_{\hat{c}}$, which relies on a fixed overlap parameter $\hat{c}$. Although this approach is more demanding for the analyst, it should, in principle, deliver more accurate results. However, some argue that it remains vulnerable to the external validity critique, namely that the context in which the imputation model is trained may differ from the context in which it is applied. In particular, it is unclear whether the estimated $\bar{c}$ provides a good approximation of the true, but unobserved, value of c. While this concern cannot be fully resolved, our results show that this external validity critique is not sufficient to discard these survey-to-survey imputations, as explained in the Introduction.

We select the mash-up measure $\hat{H}^{1/2}$ and the made-up measures $\hat{H}_{\hat{c}}$ for two reasons. The first and most important is that this choice is conservative. Indeed, they are unsophisticated measures whose social comparisons are likely to be worse than those of more

[10] The Focus property is not lost when using the identification defined in Equation (1). However, Focus is lost for the identification method used by the depth-sensitive index we define in Section 3.4.2. That is, our made-up measures G might record a worsening when an individual who is not multidimensionally poor receives a slightly worse vector of outcomes.

[11] We note that it is possible for made-up measures to inherit the decomposability properties of their ideal measure, like Subgroup Decomposability or Dimensional Breakdown. This is, for instance, the case when the made-up data is used to generate a synthetic microdataset, as is done in Section 3.4.2 to compute a depth-sensitive poverty index.

sophisticated multi-survey measures. As a result, if our unsophisticated multi-survey measures outperform single-survey measures, then we can reasonably assume that more sophisticated multi-survey measures would also outperform them.[12] The second reason is that they are simple measures. If the social comparisons based on simple multi-survey measures are very close to those of the ideal measure, then it may not be worth incurring the cost of constructing more sophisticated measures.

Finally, the parameter $\hat{c}$ of the made-up measure $\hat{H}_{\hat{c}}$ offers an additional benefit. Varying the value of $\hat{c}$ makes it possible to assess how sensitive social comparisons are to the assumption that $\hat{c}$ is close to the true value of c. In intuitive terms, if $\hat{H}_{\hat{c}}$ outperforms single-survey measures, even for implausible values of $\hat{c}$, then the external validity critique is not a sufficient reason to discard measures based on such assumptions.

The Need for Data

Ex-ante, it is conceptually unclear whether a multi-survey measure that estimates the missing joint distribution data, such as $\overline{H}_{\overline{c}}$, will necessarily outperform single-survey measures. We illustrate this in Table 1, where two regions A and B are compared. Region A has a higher level of multidimensional poverty ($H(A) > H(B)$). Both single-survey measures H_M and H_{NM} correctly rank these two regions. However, the two survey-measure $\overline{H}_{\overline{c}}$ makes an incorrect social poverty comparison because its estimate $\overline{c}(B) = -1$ is incorrect ($c(B) = 1$). This suggests that data might be necessary to assess whether, on average, a multi-survey measure outperforms single-survey measures.

Table 1: Multi-survey measures that estimate the missing joint distribution data may make worse poverty comparisons than one-survey measures.

	H_M (%)	H_{NM} (%)	c	H (%)	$\overline{c}$	$\overline{H}_{\overline{c}}$ (%)
Region A	25	20	1	25	1	25
Region B	20	15	1	20	-1	35

Notes: The two one-survey measures H_M and H_{NM} make the correct cross-region comparison $H(A) > H(B)$, while the two-survey measure $\overline{H}_{\overline{c}}$ incorrectly makes the opposite comparison because the estimated $\overline{c}(B)$ is incorrect.

We highlight a benefit of our proposal to use a *fixed* value $\hat{c}$ for the missing joint distribution data c. This proposal implies that $\hat{H}_{\hat{c}}$ cannot be affected by the issue illustrated in Table 1. Specifically, whenever both single-survey measures H_M and H_{NM} make the same social comparison, $\hat{H}_{\hat{c}}$ necessarily produces the same comparison, as stated in the

[12] By showing that made-up measures outperform single-survey measures even when the value assumed for the missing data is implausible, our results provide a powerful defense (against the external validity critique) for the use of survey-to-survey imputations, which are very likely to outperform made-up measures.

next Proposition. As a result, $\hat{H}_{\hat{c}}$ can never yield more rank reversals than both H_M and H_{NM}.

Proposition 1. *If* $H_M(A) > H_M(B)$ *and* $H_{NM}(A) > H_{NM}(B)$*, then* $\hat{H}_{\hat{c}}(A) \geq \hat{H}_{\hat{c}}(B)$ *for all* $\hat{c} \in [-1, 1]$*.*

Proof. See Appendix A.2 for the proof. ∎

2.4 Three Criteria for Social Comparisons Performance

We now turn to our third question, namely, how to assess empirically whether multi-survey measures provide more accurate social poverty comparisons than single-survey measures. The accuracy of any constrained measure $\tilde{P}$ refers to how close social comparisons with $\tilde{P}$ are to those with the ideal measure P.[13] We use three alternative criteria to quantify this proximity. Each criterion has its own limitations, but these differ across the three. Using all three, therefore, provides robustness to results that are consistent across them.

Our first criterion is the Spearman correlation, which measures the rank correlation between $\tilde{P}$ and P across societies. A higher Spearman correlation indicates better performance of $\tilde{P}$.

The second criterion, which we call the rank-reversal share, quantifies the fraction of all pairs of societies for which $\tilde{P}$ provides a social comparison opposite to that of P. Let x and y denote the outcome distributions of two different societies. For the pair of societies (x, y), measure $\tilde{P}$ provides a social comparison opposite to that of P when either

$$\tilde{P}(x) > \tilde{P}(y) \text{ and } P(x) < P(y), \text{ or}$$
$$\tilde{P}(x) < \tilde{P}(y) \text{ and } P(x) > P(y).$$

A smaller rank-reversal share indicates better performance of $\tilde{P}$.

The third criterion, which we call the budget-reallocation share, computes the fraction of a counterfactual social protection budget that must be reallocated when switching the poverty measure from P to $\tilde{P}$. The underlying idea is that some fixed monetary amount B is allocated across all societies proportionally to the poverty present in these societies. When the poverty index is the headcount ratio, budget B is allocated so that all poor individuals receive the same fraction of B, regardless of the society in which they live. For other indices, budget B is allocated so that all poor individuals with the same poverty contribution $p(x_i)$ receive the same fraction of B, and the amount received is proportional to $p(x_i)$. Observe that if a society has the same measured poverty but twice the population size of another society, then the former receives a fraction of B twice as large as that of the latter. The allocation of B across all societies according to measure $\tilde{P}$ will differ

[13] About notation, recall that multi-survey measures ($\hat{P}$) and single-survey measures (P_M or P_{NM}) are two different types of constrained measures ($\tilde{P}$).

from the allocation according to P. This criterion quantifies the minimal fraction ρ of budget B that must be reallocated across all societies so that the allocation according to P is transformed into the allocation according to $\tilde{P}$. For this reallocation, a total amount $\rho \times B$ must thus be taken away from a subset of societies for which the amount received is reduced, and the same total amount $\rho \times B$ must be given to a subset of societies for which the amount received is increased. The smaller the budget-reallocation share, the better the performance of $\tilde{P}$.

One criticism often expressed against our second criterion, the rank-reversal share, is that it is sensitive to measurement error. In particular, when two societies have outcome vectors x and y such that $P(x) \approx P(y)$, even small measurement errors may cause $\tilde{P}$ to yield a social comparison opposite to that of P. We acknowledge this limitation, but several considerations mitigate its importance for our analysis. First, and most importantly, our two other criteria are much less sensitive to measurement error. This is especially true for the budget-reallocation share, which depends primarily on the *magnitude* of the difference in poverty levels between the two societies rather than on the direction of the comparison. In other words, this criterion relies more on the cardinal values of the poverty measures than on their ordinal rankings. Second, our main focus is not the exact value of the rank-reversal share, but rather the difference in this share across two different constrained poverty measures, $\tilde{P}_1$ and $\tilde{P}_2$. A priori, measurement error is likely to affect the rank-reversal shares of $\tilde{P}_1$ and $\tilde{P}_2$ similarly. If this is the case, then the sensitivity of the rank-reversal share to measurement error is not an issue for our purposes. Third, our simulation results in Section 4 contain no measurement error by construction, so this concern does not apply in that context.

The relevance of each of these different criteria depends on the purpose of the social poverty comparisons. If the question is whether there has been progress over time, or whether a country is poorer than another, then the rank-reversal share is helpful. If the question is how to allocate policy efforts across countries or regions, then the budget-reallocation share is helpful.

3 Empirical Results

3.1 Data and Parameters

We exploit six recent cross-sectional household surveys from Bolivia, Brazil, Ecuador, Ethiopia, Ghana, and Uganda. The surveys, conducted from 2012 to 2016, are nationally representative and offer the unique advantage of providing both monetary and non-monetary outcomes for the same households. This allows us to evaluate each household's extreme (monetary) poverty status, as defined by the World Bank, and its non-monetary poverty status, as defined by the global MPI. These data also allow us to compute the ideal poverty index.

Given that our main goal is to contrast the social poverty comparisons from different poverty measures, we measure poverty at the regional level, distinguishing between urban and rural areas. The number of regions is 9 in Bolivia, 27 in Brazil, 4 in Ecuador, 11 in Ethiopia, 10 in Ghana, and 4 in Uganda, for a total of 129 geographical areas.

We define two thresholds, one for each type of poverty. We use \$1.90 as the threshold for monetary poverty, and we set $k = 1/3$ as the non-monetary poverty cutoff. We test the robustness of our results to alternative thresholds (see Section 3.4.1).

3.2 Descriptive Statistics

Next, we present descriptive statistics on the monetary and non-monetary poverty measures and the coefficient of overlap across the six countries.

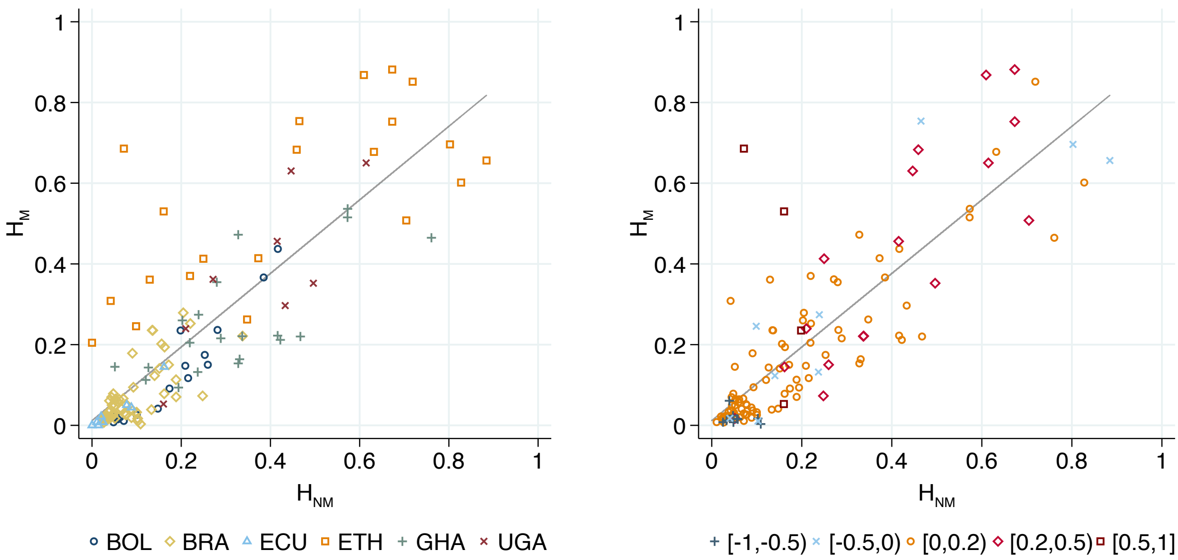
We first plot the monetary poverty rate H_M against the non-monetary poverty rate H_{NM} across all regions (see Figure 1a). The two poverty measures are positively correlated across regions, with a correlation coefficient of 0.8 in our sample. However, we observe that both types of poverty are heterogeneously distributed across regions. Neither form is irrelevant, nor does one systematically dominate the other: some regions experience more monetary poverty, while others experience more non-monetary poverty. For instance, most regions in Ethiopia have higher monetary poverty rates than non-monetary ones, while the opposite is true for most regions in Ghana.

Figure 1b plots again the monetary poverty rate H_M against the non-monetary poverty rate H_{MN} across all regions, but with points categorized by ranges of the coefficient of overlap c. We observe that regions with very different poverty profiles display similar coefficients of overlap. Conversely, some regions with high poverty levels show low overlap, while others with moderate poverty display higher overlap. This suggests that, although the two single-survey measures are positively correlated across regions, relying on only one may overlook important differences in who is classified as poor.

Figure 2 displays the average coefficient of overlap (c) by geographical area. Each area includes values for the region as a whole, urban areas, and rural areas, allowing for visual comparison within and across countries. While there is some variation across areas, the majority of coefficients lie within a relatively narrow band. Specifically, 81% of areas have c values between 0 and 0.4. The average coefficient of overlap is 0.03, and the median is 0.08.

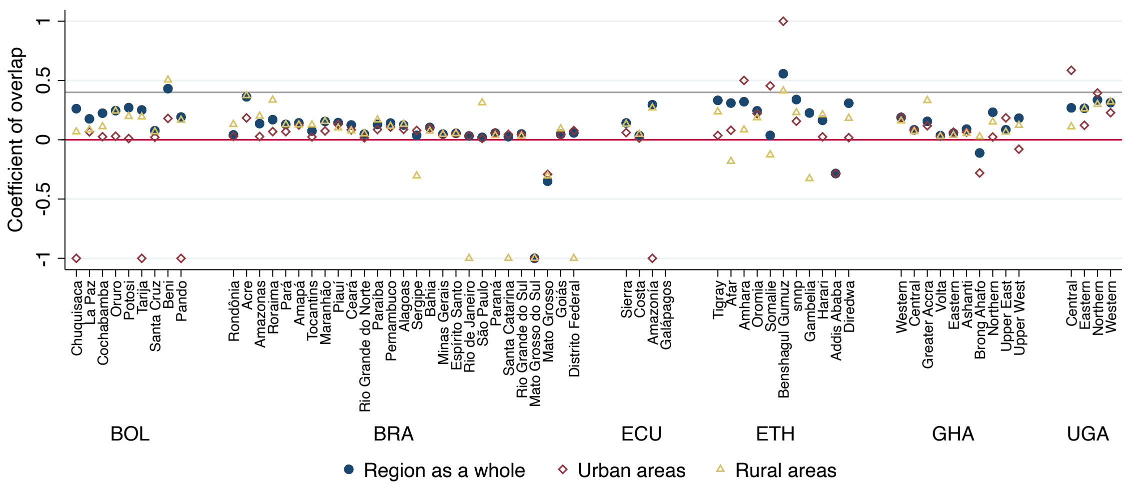
Figure 1: Single Poverty Measures

(a) H_{MN} and H_M **(b)** H_{MN} and H_M and Ranges of c



Notes: The figures display the relationship between the monetary poverty rate H_M and the non-monetary poverty rate H_{NM} across regions in six countries. Panel (a) plots H_M against H_{NM} by region. Panel (b) shows the same scatterplot with points categorized by ranges of the coefficient of overlap c.

Figure 2: Coefficient of Overlap by Area



Notes: The figure plots the coefficient of overlap c for each geographical area, including regional, urban, and rural values, for all six countries.

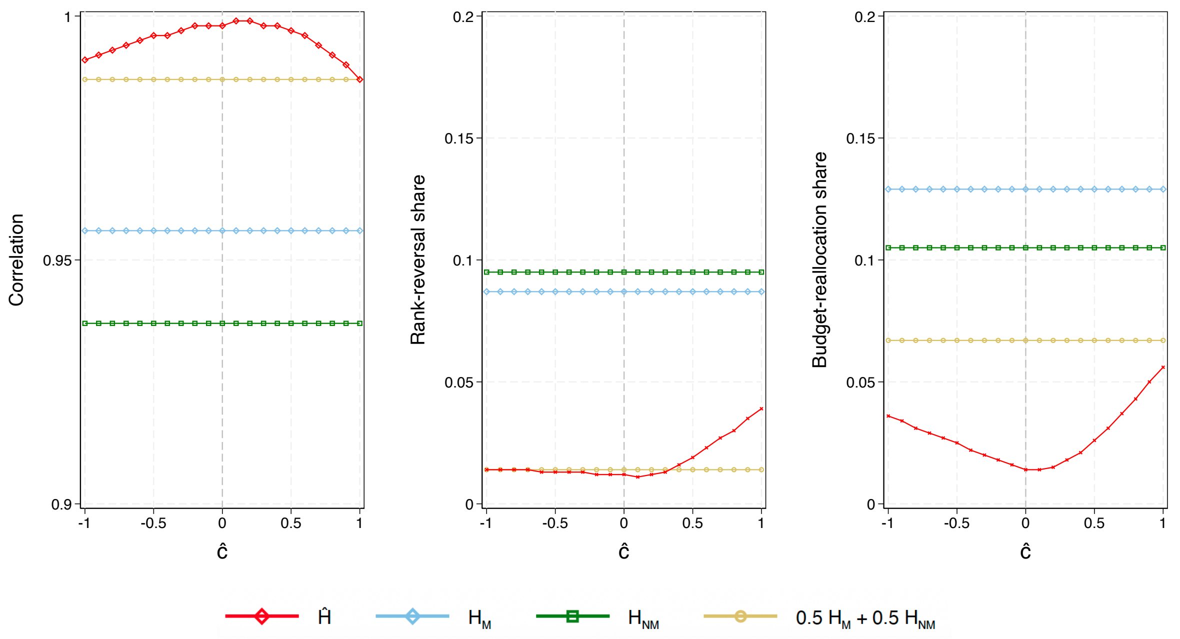
3.3 Performance of Single- vs Multi-Survey Measures

In this section, we analyze the performance of single- versus multi-survey poverty measures using our three evaluation criteria, as defined in Section 2.4: the correlation with the ideal headcount, the rank-reversal share, and the budget-reallocation share.

Figure 3 presents a comparative evaluation of different poverty measures using our three criteria for assessing how well they approximate the ideal multidimensional poverty

Figure 3: Performance of Single- Versus Multi-survey Poverty Measure

(a) Correlations **(b)** Rank-reversal shares **(c)** Budget-reallocation shares



Notes: The figure plots three evaluation criteria for single- and multi-survey poverty measures as functions of the assumed overlap coefficient ($\hat{c}$). Panel (a) shows correlations with the ideal measure, Panel (b) the rank-reversal shares, and Panel (c) the budget-reallocation shares.

headcount (H).[14] Each panel in the figure corresponds to one of the three evaluation criteria. The figure includes results for single-survey measures: H_M (monetary only) and H_{NM} (non-monetary only), a mash-up measure: $\hat{H}^{1/2}$, defined as the average of the two single-survey headcounts ($\frac{1}{2}H_M + \frac{1}{2}H_{NM}$), and our made-up measures: estimates of $\hat{H}_{\hat{c}}$ using various assumed values of the overlap coefficient $\hat{c}$ (ranging from -1 to $+1$).

Results from Figure 3 demonstrate that both types of simple multi-survey poverty measures (mash-up and made-up) yield social comparisons that are systematically closer to those based on the ideal measure H than those derived from single-survey poverty measures. This pattern holds across all three evaluation criteria. Specifically, the share of rank reversals is about 9-10% for single-survey measures, compared with below 2% for the mash-up measure and for the made-up measures under most values of $\hat{c}$. Similarly, the budget reallocation share falls from between 11 and 13% for single-survey measures to 7% for the mash-up measure, and to below 4% for most values of $\hat{c}$ used in the made-up measures. Even in the worst-case scenario, made-up measures require less than 6% budget reallocation. These improvements are meaningful, suggesting that incorporating both monetary and non-monetary dimensions, even through simple methods, can substantially improve the accuracy of social poverty comparisons.

The superiority of our made-up measures persists even under implausible values as-

[14] Appendix Figure A1 replicates Figure 3 but excludes observations with extreme values of c (c=-1 and c=1), which can be identified in Figure 2.

sumed for c. This suggests that the external validity critique (i.e., concerns about using an incorrect value for c) is not a sufficient reason to dismiss the made-up approach when the goal is to improve social poverty comparisons. In fact, even extreme assumptions about the coefficient of overlap (e.g., $\hat{c} = -1$ or $\hat{c} = 1$) lead to more accurate social comparisons than relying on single-dimension poverty measures alone.

Moreover, the results show that made-up measures (and, to a lesser extent, the mash-up measure) already resolve most of the shortcomings associated with single-survey approaches. This suggests that more sophisticated techniques, such as survey-to-survey imputation or synthetic data generation, may deliver only marginal improvements over these simpler alternatives, at least when the goal is to improve social poverty comparisons.

To illustrate the good performance of made-up measures, Figure 4 shows scatterplots comparing the made-up measures $\hat{H}_{\hat{c}}$ to the observed ideal headcount H by geographical area, for values of $\hat{c}$ spanning the entire range $[-1, 1]$. These plots reveal that $\hat{H}_{\hat{c}}$ aligns closely with H even when $\hat{c}$ takes on extreme values (e.g., -1 or $+1$), and across a variety of country contexts. This suggests that the made-up measures perform well even when the assumed dependence between poverty dimensions is far from the truth. For comparison, panels (f) and (g) of Figure 4 further plot the single-survey measures and the mash-up measure against H. We observe that $\hat{H}_{\hat{c}}$ aligns more closely with H than do the single-survey measures and the mash-up measure.

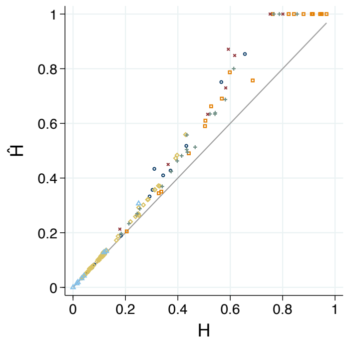
The descriptive statistics, particularly Figures 1 and 2, shed light on why the made-up approach performs well in practice. In our data, the primary drivers of cross-regional variation in multidimensional poverty are H_M and H_{NM}, which display substantial heterogeneity across regions. By contrast, the coefficient of overlap c is relatively stable and contributes less to the observed variation in overall poverty H. Consequently, even rough assumptions about c can yield reasonably reliable estimates of H when combined with accurate information on H_M and H_{NM}.

Finally, we further notice that made-up measures based on plausible assumptions enjoy some advantages over mash-up measures: the former typically slightly outperform the latter. Moreover, the former better approximate *the level of* the ideal headcount ratio than the latter.[15] This can already be deduced from Figure 4 and is further illustrated in Figure A2 in Appendix A.5.

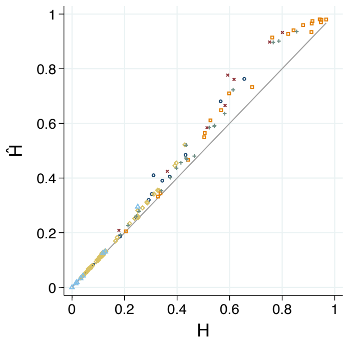
[15] All our results contrast differences in levels, not directly the levels themselves.

Figure 4: Accuracy of $\hat{H}_{\hat{c}}$, H_M, H_{NM}, and $\hat{H}^{1/2}$

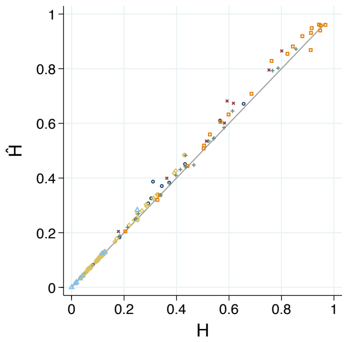
(a) $\hat{H}_{\hat{c}=-1}$ and H



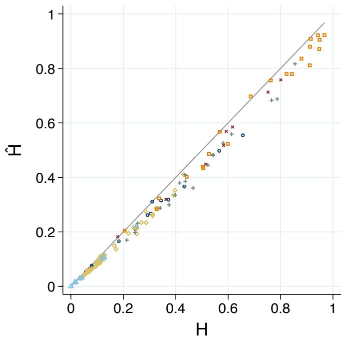
(b) $\hat{H}_{\hat{c}=-0.5}$ and H



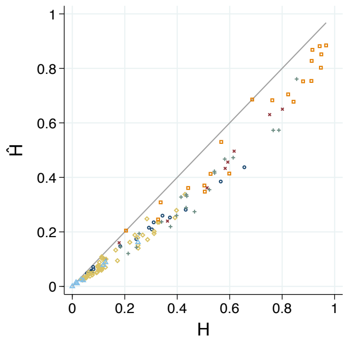
(c) $\hat{H}_{\hat{c}=0}$ and H



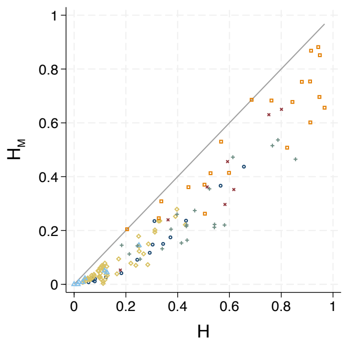
(d) $\hat{H}_{\hat{c}=0.5}$ and H



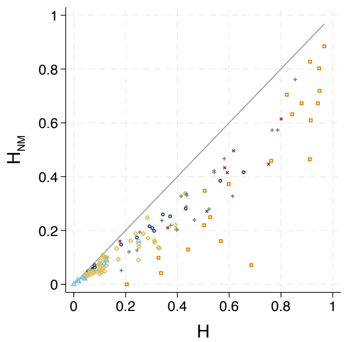
(e) $\hat{H}_{\hat{c}=1}$ and H



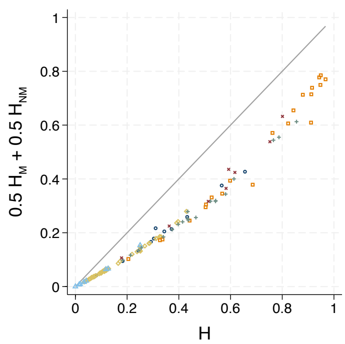
(f) H_M and H



(g) H_{NM} and H



(h) $\hat{H}^{1/2}$ and H



∘ BOL ◇ BRA △ ECU □ ETH + GHA × UGA

Notes: This figure shows different poverty measures against the ideal measure H by geographical area defined jointly by region and settlement type (urban or rural). Panels (a) to (e) show made-up measures ($\hat{H}_{\hat{c}}$) for different values of $\hat{c}$ against H. Panels (f) to (h) show H_M, H_{NM}, and $\hat{H}^{1/2}$ against H.

3.4 Robustness

This section examines the robustness of our main findings to alternative parameter choices and methodological assumptions. Specifically, we assess whether the relative performance of single- and multi-survey measures holds under different poverty cutoffs and index specifications.

3.4.1 Alternative Poverty Cutoffs

We show that the results are robust to selecting different values for the monetary poverty line z and the non-monetary poverty cutoff k. We change the value of z from \$1.9 to \$3.2 a day (the former is the international poverty line for low-income countries, while the latter is the international poverty line for low- and middle-income countries). We change the value of k from 0.33 to 0.2 (the former corresponds to the value used by the global MPI). Both changes correspond to "less austere" poverty cutoffs, i.e., more people are identified as monetary poor (respectively, non-monetary poor).

Appendix Figure A3 shows that our findings on the relative performance of single- and multi-survey poverty measures are robust to these changes in poverty thresholds. Performance rankings across the three evaluation criteria remain stable: as in our main findings shown in Figure 3, the mash-up and made-up measures consistently outperform single-survey measures. This reinforces the reliability of our conclusions.

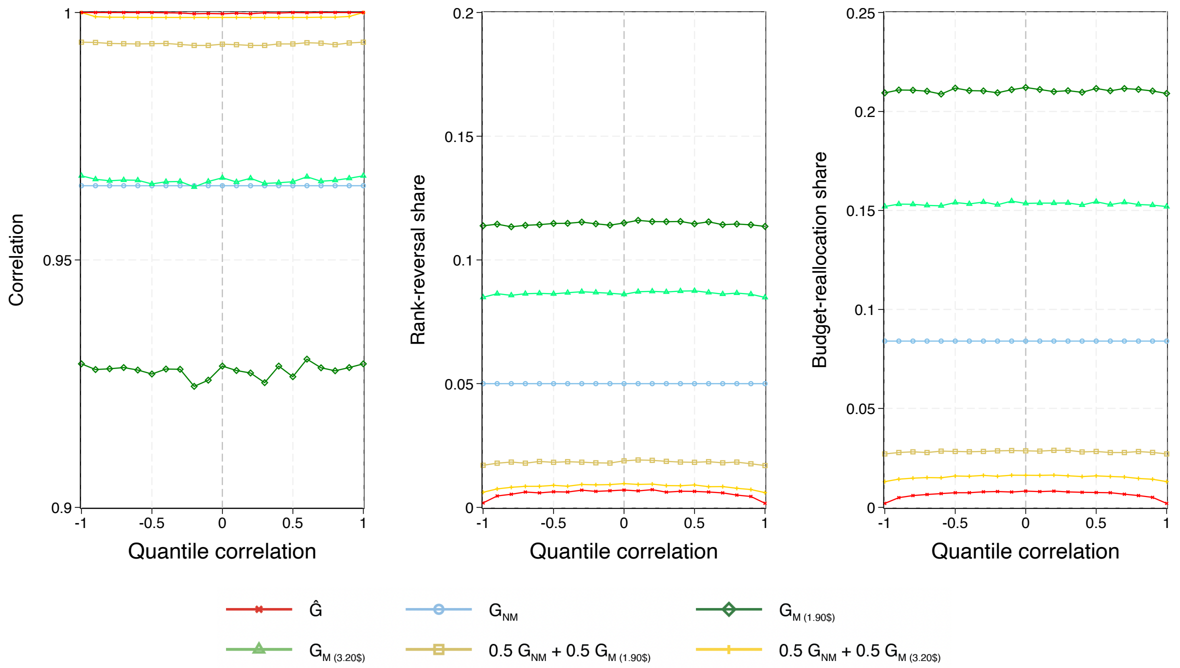
3.4.2 Depth-Sensitive Poverty Index

We next assess robustness to using a depth-sensitive ideal poverty measure rather than the headcount index. As single-survey measures, we adopt the poverty gap ratio for monetary poverty (G_M, under two poverty lines) and the adjusted headcount ratio for non-monetary poverty (G_N). Multi-survey variants are defined in parallel to before. i) Mash-up measures take the simple average of the two single-survey indices. Because the monetary poverty gap ratio is computed under two poverty lines, this yields two mash-up measures (one for each monetary gap specification). ii) Made-up measures ($\hat{G}$) approximate the ideal multidimensional gap index by combining each individual's non-monetary deprivation score with a weighted monetary gap, while assuming the unobserved joint distribution of monetary and non-monetary outcomes. To construct the latter, we generate synthetic microdata using a Gaussian copula that embeds a pre-specified correlation between the two dimensions (details can be found in Appendix A.4). To limit noise from the stochastic assignment, we bootstrap this procedure 20 times and average the resulting estimates.

Results for our Three Criteria.—We assess the performance of single- versus multi-survey depth-sensitive poverty measures using our three evaluation criteria. As shown in Figure 5, the multi-survey depth-sensitive measures substantially outperform their single-

Figure 5: Performance of Single- Versus Multi-survey Depth-Sensitive Poverty Measure

(a) Correlations **(b)** Rank-reversal shares **(c)** Budget-reallocation shares



Notes: The figure plots three evaluation criteria (correlations, rank-reversal shares, and budget-reallocation shares) for depth-sensitive poverty measures as functions of the assumed quantile correlation between dimensions.

survey counterparts across all three evaluation criteria.[16]

For the rank-reversal shares, single-survey measures record values of 5% for G_{NM}, and around 10% for the two monetary gap ratios, compared to less than 2% for the mash-up measures and under 1% for the made-up measures. The budget-reallocation shares show a similar pattern: from around 15–20% for G_M and 8% for G_{NM}, to below 3% for the mash-up measures, and less than 1% for the made-up measures. The consistency of these patterns (even when the assumed dependence between dimensions is extreme) suggests that adopting simple multi-survey measures can deliver large improvements in the accuracy of social poverty comparisons, even under a depth-sensitive index.

4 Simulation Results

In the previous section, we showed that multi-survey measures substantially outperform single-survey measures in six countries. While these findings are encouraging, they may partly reflect the specific characteristics of those data. For instance, although the countries analyzed have very different poverty profiles, we observe that H_M and H_{NM} are highly

[16]The slight variation in monetary estimates (G_M) reflects randomness in the copula-based simulation. With small regional samples, the random assignment of synthetic monetary values introduces minor fluctuations in both the monetary and multidimensional poverty estimates across iterations.

heterogeneous across regions, while c is not. It is natural to ask whether multi-survey measures still outperform single-survey measures under different conditions. To assess the generality of our results, we turn to simulation exercises that allow us to systematically vary the underlying data-generating process. By controlling key parameters, simulations allow us to test whether the advantages of multi-survey measures extend beyond the particular contexts observed in our empirical analysis. Additionally, they allow us to investigate which features drive the relative performance of multi-survey measures.

Specifically, we explore the performance of our proposed made-up measures using Monte Carlo simulations that generate distributions of the three key aggregate poverty statistics (H_M, H_{NM}, and the coefficient of overlap c) under varying assumptions about their variances, the correlation between H_M and H_{NM}, and the mean absolute gap between H_M and H_{NM}. These simulations, therefore, abstract from individual-level data and instead model how combinations of (H_M, H_{NM}, c) vary across hypothetical units, such as regions or societies. The design allows us to isolate how each of these features (the variance of c and of H_M and H_{NM}, the strength of correlation between these rates, and the difference in their levels) affects the accuracy of social poverty comparisons.

4.1 Construction of Simulated Scenarios

We build three simulation scenarios (A–C), whose main characteristics are summarized in Table 2. In Scenario A (variance-only), we vary the extent to which cross-unit variation comes from the coefficient of overlap c versus the two (single-survey) poverty measures H_M and H_{NM}. To this end, we set the monetary and non-monetary poverty-rate distributions to be uncorrelated and to have equal means. We consider three variance configurations: (1) low variance in c and high variance in H_M and H_{NM}; (2) high variance in c and low variance in H_M and H_{NM}; and (3) high variance in all three. Specifically, H_M and H_{NM} are drawn independently from symmetric Beta distributions with either low or high variance, and the overlap coefficient c is drawn independently from a rescaled Beta distribution on $[-1, 1]$ with low or high variance. In addition to imposing zero correlation between H_M and H_{NM} ($\rho = 0$), we eliminate the mean absolute gap between them by recentring H_{NM} to the mean of H_M (clipping values to $(0, 1)$).

For scenarios B and C, we fix a low variance for c and a high variance for H_M and H_{NM}. We choose this variance configuration because it is favorable for the performance of our made-up measures, and hence, it allows us to examine how other parameters affect their performance.

In scenario B (correlation-only), besides setting a low variance for c and a high variance for H_M and H_{NM}, we keep the mean absolute gap between H_M and H_{NM} at zero, and only vary the correlation between them from moderate ($\rho = 0.5$) to near-perfect ($\rho = 0.99$). This allows us to assess how the correlation by itself shapes performance. Specifically, we generate H_M and H_{NM} using a Gaussian copula calibrated to a target Spearman correlation

Table 2: Scenario Design to Isolate Variance, Correlation, and Mean Absolute Gap Effects

Scenario	Correlation (ρ)	Mean absolute gap	$Var(c)$	$Var(H_M/H_{NM})$
A: variance-only	0	0	low / high	low / high
B: correlation-only	{0.5, 0.99}	0	low	high
C: gap-only	0	{0.2, 0.8}	low	high

Notes: Scenario A varies the variance of H_M and H_{NM}, and of c. It assumes zero correlation between H_M and H_{NM} ($\rho = 0$) and equal means. Scenario B sets a low variance for c and high, equal variances for H_M and H_{NM} (with equal means), while varying the correlation between H_M and H_{NM} from moderate ($\rho = 0.5$) to near-perfect ($\rho = 0.99$). Scenario C also fixes a low variance for c and a high variance for H_M and H_{NM}, while varying the mean absolute gap between H_M and H_{NM} from small (0.2) to large (0.8).

of 0.5 or 0.99, and then map these to Beta distributions with high variance via the inverse Beta cumulative distribution function (CDF). We remove the mean gap between H_{NM} and H_M by centering H_{NM} on the mean of H_M and clipping values to $(0, 1)$. In turn, we draw the overlap coefficient c independently from a rescaled Beta on $[-1, 1]$ with low variance.

Scenario C (gap-only) again fixes a low variance for c and a high variance for H_M and H_{NM}, but keeps H_M and H_{NM} independent. The mean absolute gap between them varies from small (0.2) to large (0.8). We reparameterize the Beta distributions so that both H_M and H_{NM} share the same target (high) variance at each gap level (with values obtained via the inverse Beta CDF). This isolates the effect of mean absolute differences. In turn, c is drawn as in scenario B.

This exercise yields a total of seven variants across three scenarios. For each variant, we compute the ideal poverty index (H), the single-survey measures (H_M and H_{NM}), the mash-up measure ($\hat{H}^{1/2}$), and our made-up measures ($\hat{H}_{\hat{c}}$) based on different assumed values of $\hat{c}$ ranging from -1 to $+1$. Finally, we assess the performance of the simulated single- versus multi-survey poverty measures using the three evaluation criteria introduced earlier: correlation with the ideal headcount, rank-reversal shares, and budget-reallocation shares.

Considering these three scenarios in combination allows us to cover a broad set of empirically relevant cases. First, examining the variance of H_M, H_{NM}, and c together is useful because the informativeness of overlap assumptions could depend on how much independent variation is already captured by the single-survey measures. If H_M and H_{NM} are highly dispersed, they might explain most of the cross-unit variation in multidimensional poverty, so the role of the coefficient of overlap could be comparatively limited. By contrast, if H_M and H_{NM} exhibit little variation, differences in overlap assumptions may become more consequential.

Second, if H_M and H_{NM} are highly correlated across units, then either dimension could already serve as a reasonable proxy for multidimensional poverty. In that case, single-survey measures might perform reasonably well, and the gains from making up the missing joint distribution data could be limited. Instead, if the correlation is weak or moderate, each dimension might provide distinct information, and making up the missing joint distribution data could become much more valuable.

Third, the mean absolute gap between H_M and H_{NM} captures systematic level differences between them and measures the extent to which one dimension consistently exceeds the other across units, on average. This feature can also shape how informative the joint distribution is. When the mean absolute gap is large, one single-survey poverty measure effectively dominates. In such cases, the made-up measures are largely driven by the dominant single-survey measure, so combining both dimensions could add little information. Conversely, when the mean absolute gap is small, the two single-survey measures are closer in levels, and the made-up measures could better reflect a balanced contribution from both them.

4.2 Performance Based on Budget-Reallocation Shares

For expositional clarity, this section presents results using the budget-reallocation share. We provide the corresponding results for the other two performance criteria in Appendix A.5 (see Figures A4 and A5). Most of our results hold under the three performance criteria, and we highlight any relevance differences across them.

Figure 6 shows the results. Panel A displays the results for scenario A and it shows that the made-up measures systematically outperform the single-survey measures. The smallest improvements for the made-up measures appear in panel A(b), where the variance of c is high and that of H_M and H_{NM} is low. The intuition for this result is that when the variance is concentrated in c, any approach that fixes c across units cannot capture its heterogeneity, so the advantage of the made-up measures over single-survey measures shrinks. Conversely, when the variance of c is low, and that of H_M and H_{NM} is high, the made-up measures deliver their largest gains (see Panel A(a)). Each single-survey measure carries useful but incomplete information, so setting a fixed overlap allows the made-up measures to combine these signals and improve performance. Finally, there are also large gains in Panel A(c), where the variance of both the single-survey poverty measures and the coefficient of overlap is high.

Panel B displays the results for scenario B, which only varies the correlation between H_M and H_{NM}. The uncorrelated baseline ($\rho = 0$) matches the scenario A variant with low variance of c and high variances for H_M and H_{NM}. Accordingly, we refer to panel A(a) for $\rho = 0$, and Panel B for $\rho \in \{0.5, 0.99\}$. As expected, the made-up measures perform best when the two dimensions are uncorrelated ($\rho = 0$). Interestingly, under moderate positive correlation ($\rho = 0.5$; panel B(d)) they still exhibit strong gains. With near-perfect correlation ($\rho = 0.99$; panel B(e)), the benefits naturally diminish.[17] However, for the budget-reallocation share, which is a cardinally sensitive criterion, the made-up measures

[17] This is particularly true for the rank-based metric, as shown in Appendix Figure A5.

continue to deliver gains. This underscores an important policy implication: even when the single-survey measures are (almost) perfectly correlated, capturing both can improve decisions that depend on intensity rather than rank order.

Panel C presents the results for scenario C, in which we vary the mean absolute gap between the two partial distributions from small to large. When the gap is large, the performance of the made-up measures is comparable to that of the higher-mean single-survey measure (i.e., H_M by construction). Their performance curves are nearly flat: once H_M and H_{NM} differ sharply in level, most cross-unit variation is captured by the dominant measure, and the assumed overlap plays little role. By contrast, when the gap is smaller, the made-up measures still outperform the single-survey measures. We further notice that in the large-gap scenario, the made-up measures exhibit a clearer advantage over the mash-up measure. While the made-up measures perform similarly to H_M, the mash-up measure performs worse, effectively adding noise rather than improving accuracy.

Across all three scenarios, a coherent picture emerges. The made-up measures generally outperform the single-survey measures (and the mash-up measures as well) across values of $\hat{c}$ corresponding to non-extreme assumptions. It is worth emphasizing that in all simulation scenarios, the coefficient of overlap c is drawn from a symmetric distribution on $[-1, 1]$ with mean zero. Consequently, the best simulated performance of the made-up measures is typically observed when the assumed value $\hat{c} = 0$, simply because this matches the average true dependence in the simulated data—not because zero overlap would be optimal in empirical applications.

The gains of made-up measures over single-survey measures are largest when H_M and H_{NM} each contain meaningful but distinct information—for example, when the variance is concentrated in H_M and H_{NM} rather than in c, or when their correlation is moderate. By contrast, when the variance is mainly driven by c or when the correlation between H_M and H_{NM} is very high, the two partial distributions move closely together, leaving little room for improvement, and the made-up measures perform similarly to one or both single-survey measures. Likewise, when the mean absolute gap between H_M and H_{NM} is very large, the single-survey measure with the higher mean becomes the primary driver of the results, and the made-up measures converge toward it.

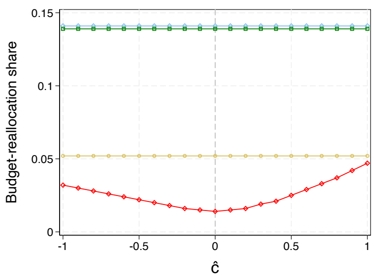
Moreover, our results show that made-up measures that are based on plausible assumptions are almost never outperformed by mash-up measures, while they sometimes substantially outperform the latter. This is, for instance, the case for the rank-reversal share when the deprivation captured in one survey is substantially more prevalent than that captured in the other survey (see panel C(g) in Appendix Figure A5).

Finally, we notice that in some settings our made-up measures recover nearly all the errors made by single-survey measures (for instance, in panels A(a), B(d), and C(f) of Appendix Figure A4), so there is little added value in performing a complex imputation to close the remaining gap.

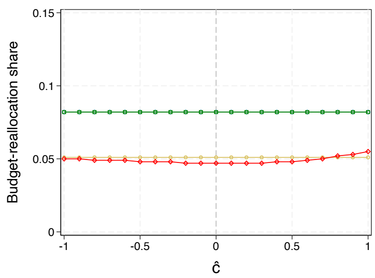
Figure 6: Budget-Reallocation Shares in Scenarios A-C

Panel A: Scenario A (variance-only)

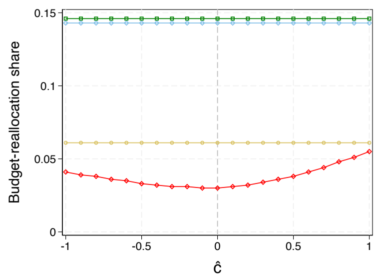
(a) Low c — High H_M/H_{NM}



(b) High c — Low H_M/H_{NM}

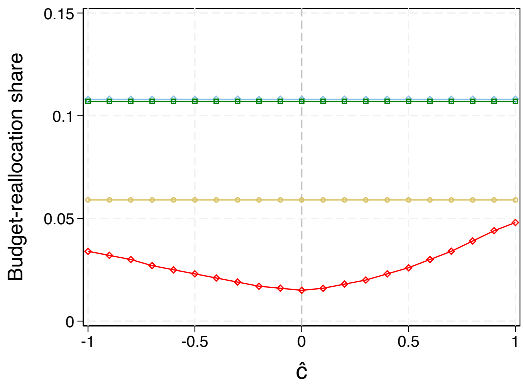


(c) High c — High H_M/H_{NM}

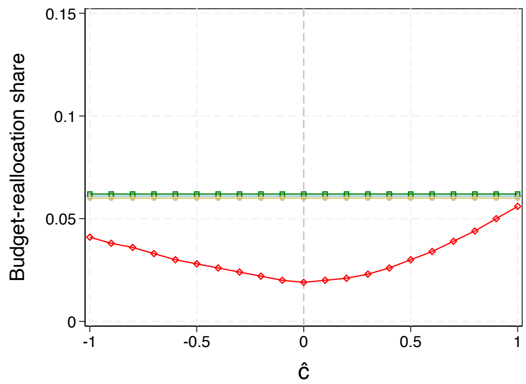


Panel B: Scenario B (correlation-only)

(d) $\rho = 0.5$

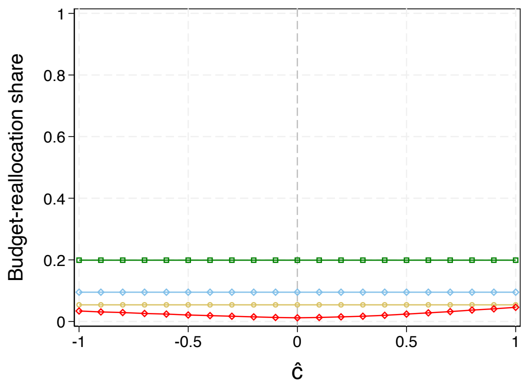


(e) $\rho = 0.99$

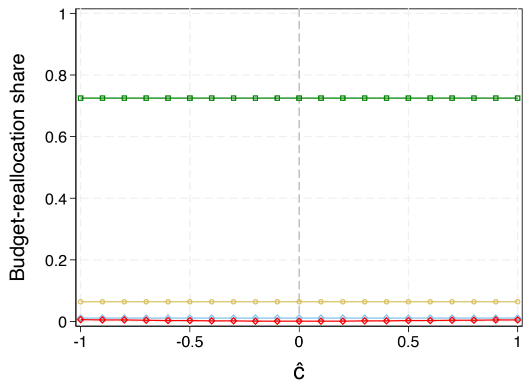


Panel C: Scenario C (gap-only)

(f) Gap 0.2



(g) Gap 0.8



Notes: The figure displays the budget-reallocation share for made-up, mash-up, and single-survey poverty measures as a function of the assumed overlap coefficient $\hat{c}$. Panel A corresponds to scenario A (variance-only), Panel B presents results from scenario B (correlation-only), and Panel C displays scenario C (gap-only). Subpanels (a)–(g) correspond to the scenario variants described in Section 4.1.

5 Concluding Remarks

This paper proposes a simple but conceptually grounded way to improve multidimensional poverty comparisons when data on different dimensions come from separate surveys. We develop multi-survey measures, which we call made-up measures, that combine information on monetary and non-monetary outcomes by assuming the missing part of the joint distribution. Unlike mash-up indices, which are based on arbitrary aggregation formulas, our approach allows the use of a conceptually grounded multidimensional poverty index. Using household surveys from six countries where both types of outcomes are observed jointly, we show that the made-up measures approximate the ideal benchmark much more closely than single-survey alternatives, across correlations, rank-reversal shares, and budget-reallocation shares. Robustness checks confirm that these findings are robust to alternative poverty thresholds and depth-sensitive indices. Moreover, simulation exercises prove that our results hold in a set of scenarios. Our results demonstrate that cross-context poverty comparisons are more accurate when combining partial distributions from separate surveys than when relying on a single survey alone. The broader message is that the relevant question when dealing with a split-surveys constraint is whether social poverty comparisons are correct, not whether the assumed joint distribution is itself correct. Taken together, the empirical and simulation evidence show that assuming (or estimating) the unobserved part of the joint distribution can substantially improve the reliability of multidimensional poverty monitoring.

Our results also provide some reasons to favor made-up measures over mash-up measures. Besides avoiding arbitrary aggregation formulas, made-up measures that are based on plausible assumptions are almost never outperformed by mash-up measures, while they sometimes substantially outperform the latter. Moreover, when the ideal index is the headcount ratio, there is another advantage that made-up measures may enjoy over mash-up measures. In our data, the former better approximate *the level of* the ideal headcount ratio than the latter.[18]

Our made-up measures are particularly simple when assuming that the outcomes collected from one survey are independent of the outcomes collected in the other survey.[19] Interestingly, in our data, $\hat{c} = 0$ is among the values for which our made-up measures perform the best. The simplifying independence assumption might thus not necessarily imply too high a price.

Our approach comes with some limitations. First, it does not allow for direct interpersonal poverty comparisons. This restricts its usefulness for applications such as within-region targeting, where identifying and ranking individual households is essential. Our

[18] See last paragraph in Section 3.3.

[19] Indeed, when assuming $\hat{c} = 0$, we get $\hat{H}_0 = H_M + H_{NM} - H_M H_{NM}$ (from Equation (4)). To take an example illustrating this simplicity, if a society has $H_M = 1/3$ and $H_{NM} = 1/4$, then $\hat{H}_0$ assumes that the fraction of individuals who are both monetary and non-monetary poor is $1/3 \times 1/4$.

solution is not intended for targeting but remains valuable for monitoring and comparing poverty across groups and regions. Second, the made-up measures may lose some desirable properties of their underlying poverty index. For the ordinal properties that are lost, whose purpose is to ensure that social poverty comparisons make sense, our results show that violations of these properties by made-up measures have less severe consequences than the corresponding violations by the alternative measures, whose comparisons are more biased. In other words, this limitation is also shared by the alternative solutions, and even more so in their case.

The methodology we develop can be adapted to different applications, depending on the number and nature of the available surveys. Our paper illustrates one practical way to integrate outcomes from two separate surveys, both for the headcount ratio and for a depth-sensitive poverty index. Ongoing work by Brunckhorst et al. (2025) illustrates the flexibility of our framework: they apply our approach to expand coverage and produce more timely estimates of the World Bank's Multidimensional Poverty Measure, developing a 'probabilistic fusion algorithm' that implements our solution with more than two surveys when the ideal index is the headcount ratio. This application shows that our framework is already gaining traction and highlights its practical relevance for policymakers.

While our analysis focuses on poverty measurement, the proposed method has a broader potential scope. In particular, it may be applied to other domains such as empowerment, capability assessment, inequality of opportunity, or related multidimensional outcomes. In general, our approach could be of interest for any social indicator in which dimension-specific outcomes are aggregated at the individual level into individual scores, which are then aggregated across individuals into a social measure, as is the case for mainstream additively decomposable multidimensional poverty measures. This opens avenues for future work in multidimensional welfare analysis. In this sense, the flexibility of our framework extends well beyond poverty analysis, offering a general strategy for combining fragmented data to obtain robust multidimensional indicators.

References

Alkire, S. and J. Foster (2011): "Counting and Multidimensional Poverty Measurement,," *Journal of Public Economics*, 95, 476–487.

Alkire, S., U. Kanagaratnam, and N. Suppa (2022): "The Global Multidimensional Poverty Index (MPI) 2022: Disaggregation Results and Methodological Note," MPI Methodological Note 53, Oxford Poverty and Human Development Initiative (OPHI), Oxford.

Anand, S. and A. Sen (1994): "Human Development Index: Methodology and Measurement," Occasional Paper 12, Human Development Report Office (HDRO), United Nations Development Programme (UNDP), New York.

Atkinson, A. B. and F. Bourguignon (1982): "The Comparison of Multi-dimensioned Distributions of Economic Status," *The Review of Economic Studies*, 49, 183–201.

Bourguignon, F. and S. R. Chakravarty (2003): "The Measurement of Multidimensional Poverty," *The Journal of Economic Inequality*, 1, 25–49.

Brunckhorst, B. J., H.-A. Dang, C. Lakner, M. C. Nguyen, and N. Yonzan (2025): "Multidimensional Fusion," Tech. rep.

Chakravarty, S. R., J. Deutsch, and J. Silber (2008): "On the Watts multidimensional poverty index and its decomposition," *World Development*, 36, 1067–1077.

Christiaensen, L., P. Lanjouw, J. Luoto, and D. Stifel (2012): "Small Area Estimation-based Prediction Methods to Track Poverty: Validation and Applications," *The Journal of Economic Inequality*, 10, 267–297.

Dang, H.-A. H. and P. F. Lanjouw (2023): "Regression-Based Imputation for Poverty Measurement in Data-Scarce Settings," in *Research Handbook on Measuring Poverty and Deprivation*, Edward Elgar Publishing, 141–150.

Decancq, K. (2023): "Cumulative deprivation: identification and aggregation," in *Research Handbook on Poverty and Inequality*, Edward Elgar Publishing, 52–67.

Decancq, K., M. Fleurbaey, and F. Maniquet (2019): "Multidimensional Poverty Measurement with Individual Preferences," *The Journal of Economic Inequality*, 17, 29–49.

Decancq, K. and V. Jorda (2024): "Evolution of Global Inequality in Well-being: A Copula-based Approach," *arXiv preprint arXiv:2410.22892.*

Decerf, B. (2023): "A Preference-Based Theory Unifying Monetary and Non-Monetary Poverty Measurement," *Journal of Public Economics*, 222, 104898.

——— (2024): "Multidimensional Well-Being Measurement Practices: A Review Focused on Improving Global Multidimensional Poverty Indicators," Policy Research Working Paper 10800, World Bank, Washington, DC.

——— (2025): "A Quick-Fix for Perverse Incentives Inherent in Mainstream Multidimensional Poverty Measures," Policy Research Working Paper 11113, World Bank, Washington, DC.

Diener, E., S. Oishi, and L. Tay (2018): "Advances in Subjective Well-Being Research," *Nature Human Behaviour*, 2, 253–260.

Duclos, J., D. E. Sahn, and S. D. Younger (2006): "Robust Multidimensional Poverty Comparisons," *The Economic Journal*, 116, 943–968.

Evans, M., R. Nogales, and M. Robson (2024): "Monetary and multidimensional poverty: correlation, mismatches, and a combined approach," *The Journal of Development Studies*, 60, 147–170.

Fleurbaey, M. (2009): "Beyond GDP: The Quest for a Measure of Social Welfare," *Journal of Economic Literature*, 47, 1029–1075.

Gethin, A. (2025): "Who Benefits From Public Services? Novel Evidence and Implications for Inequality Measurement," *Journal of Development Economics*, 103627.

Mathiassen, A. (2013): "Testing Prediction Performance of Poverty Models: Empirical Evidence from Uganda," *Review of Income and Wealth*, 59, 91–112.

Mathiassen, A. and B. K. G. Wold (2021): "Predicting Poverty Trends by Survey-to-Survey Imputation: The Challenge of Comparability," *Oxford Economic Papers*, 73, 1153–1174.

Newhouse, D. L., S. Shivakumaran, S. Takamatsu, and N. Yoshida (2014): "How Survey-to-Survey Imputation Can Fail," Policy Research Working Paper 6961, World Bank, Washington, DC.

Piacquadio, P. G. (2026): "Lack of opportunities, non-paternalism, and priority to the worst off: PG Piacquadio," *Social Choice and Welfare*, 66, 255–274.

Ravallion, M. (2011): "On Multidimensional Indices of Poverty," *The Journal of Economic Inequality*, 9, 235–248.

Sen, A. (1976): "Poverty: an Ordinal Approach to Measurement," *Econometrica*, 44, 219–231.

Seth, S. and M. E. Santos (2019): "On the interaction between focus and distributional properties in multidimensional poverty measurement," *Social Indicators Research*, 145, 503–521.

Stiglitz, J. E., A. Sen, and J.-P. Fitoussi (2009): "Report by the Commission on the Measurement of Economic Performance and Social Progress," Tech. rep., Commission on the Measurement of Economic Performance and Social Progress, Paris.

Tsui, K.-y. (2002): "Multidimensional poverty indices," *Social choice and welfare*, 19, 69–93.

UNDP (1990): *Human Development Report 1990*, New York: Oxford University Press.

——— (1997): *Human Development Report 1997: Human Development to Eradicate Poverty*, New York: Oxford University Press.

Watkins, K. (2006): *Human Development Report 2006: Beyond Scarcity-Power, Poverty and the Global Water Crisis*, New York: United Nations Development Programme (UNDP) and Oxford University Press.

A Appendix

A.1 Intuition for the Definition of c

We provide intuition on why the mathematical expression for c captures linearly the distance to independence between the two poverty statuses.

The starting point of the definition for c is that, when the two statuses are *independent*, we should have that

$$\frac{Proba\left(\mathbb{1}^{NM} = 1 \cap \mathbb{1}^{M} = 1\right)}{Proba\left(\mathbb{1}^{M} = 1\right)} = Proba\left(\mathbb{1}^{NM} = 1\right) = H_{NM}.$$

The definition for c normalizes the deviation between this independent probability and the maximal (respectively minimal) value that this probability can take for positive (respectively negative) dependence between the two statuses.

We provide intuition for Equation (4), namely

$$H = \begin{cases} H_M + H_{NM} - H_M \times (H_{NM} + c \times (\min(1, H_{NM}/H_M) - H_{NM})) \text{ if } c \geq 0 \\ \\ H_M + H_{NM} - H_M \times (H_{NM} + c \times (H_{NM} - \max(0, (H_M + H_{NM} - 1)/H_M))) \text{ else} \end{cases}$$

Consider the case $c \geq 0$, for which there is positive dependence between the two statuses. This case is such that $H_{NM} \leq \frac{H_M + H_{NM} - H}{H_M}$, where the numerator $H_M + H_{NM} - H$ captures the fraction of individuals that are *both* monetary poor and non-monetary poor, so that

$$\frac{H_M + H_{NM} - H}{H_M} = \frac{Proba(\mathbb{1}^{NM} = 1 \cap \mathbb{1}^{M} = 1)}{Proba(\mathbb{1}^{M} = 1)}.$$

First, consider the subcase for which there are enough balls in the M-poor urn to draw all the $H_{NM} \times 100$ balls, which implies $H_{NM} \leq H_M$. This subcase implies that $\min(1, H_{NM}/H_M) = H_{NM}/H_M$. The mathematical expression for H in Equation (4) becomes

$$H = H_M + H_{NM} - H_M \times \left(H_{NM} + c \times H_{NM} \times \frac{1 - H_M}{H_M}\right)$$

which yields

$$c = 1 - \frac{H - H_M}{H_{NM} \times (1 - H_M)}$$

where

$$\frac{H - H_M}{(1 - H_M)} = \frac{Proba(\mathbb{1}^{NM} = 1 \cap \mathbb{1}^{M} = 0)}{Proba(\mathbb{1}^{M} = 0)}$$

and as $H_{NM} = Proba(\mathbb{1}^{NM} = 1)$ we get

$$c = 1 - \frac{Proba(\mathbb{1}^{NM} = 1 \cap \mathbb{1}^{M} = 0)}{Proba(\mathbb{1}^{NM} = 1) \times Proba(\mathbb{1}^{M} = 0)} \tag{7}$$

where in the last term the numerator is the number of balls drawn in the M-rich urn and the denominator is the number of balls that should be drawn in the M-rich urn if the two statuses were independent. The value of c decreases linearly with the number of balls drawn in the M-rich urn.

Second, consider the subcase for which there are not enough balls in the M-poor urn to draw all the $H_{NM} \times 100$ balls, which implies $H_{NM} > H_M$. This subcase implies that $\min(1, H_{NM}/H_M) = 1$. The mathematical expression for H in Equation (4) becomes

$$H = H_M + H_{NM} - H_M \times (H_{NM} + c \times (1 - H_{NM}))$$

which yields

$$\begin{aligned} c =& \frac{(H_M + H_{NM} - H) - H_M \times H_{NM}}{H_M - H_M \times H_{NM}} \\ =& \frac{Proba(\mathbb{1}^{NM} = 1 \cap \mathbb{1}^{M} = 1) - Proba(\mathbb{1}^{M} = 1) \times Proba(\mathbb{1}^{NM} = 1)}{Proba(\mathbb{1}^{M} = 1) - Proba(\mathbb{1}^{M} = 1) \times Proba(\mathbb{1}^{NM} = 1)} \end{aligned}$$

where the first term in the numerator is the number of balls that are drawn from the M-Poor urn and the second term in the numerator is the number of balls that would have been drawn from the M-Poor urn if the two statuses were independent. In turn, the first term in the denominator is the number of balls in the M-Poor urn and its second term is again the number of balls that would have been drawn from the M-Poor urn if the two statuses were independent. The value of c increases linearly with the number of balls drawn in the M-poor urn.

The case $c < 0$, where the dependence between the two statuses is negative is such that $H_{NM} \geq \frac{H_M + H_{NM} - H}{H_M}$, where $H_M + H_{NM} - H$ captures the fraction of individuals that are *both* monetary poor and non-monetary poor. This case comes with a similar intuition and is thus omitted.

A.2 Proof of Proposition 1

It is sufficient to show that for all $\hat{c} \in [-1, 1]$, $H_M \in [0, 1)$ and $H_{NM} \in [0, 1)$ we have $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_M} \geq 0$ and $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_{NM}} \geq 0$. From the mathematical expression of $\hat{H}_{\hat{c}}$ given in Equation (4), we build a case by case proof:

- CASE 1: $\hat{c} \in [0, 1]$ and $H_{NM} > H_M$:

  This case is such that $\min(1, H_{NM}/H_M) = 1$.
  We have $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_M} = (1 - \hat{c}) \times (1 - H_{NM})$ and $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_{NM}} = 1 - H_M \times (1 - \hat{c})$.

- CASE 2: $\hat{c} \in [0, 1]$ and $H_{NM} \leq H_M$:

  This case is such that $\min(1, H_{NM}/H_M) = H_{NM}/H_M$.
  We have $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_M} = 1 - H_{NM} \times (1 - \hat{c})$ and $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_{NM}} = (1 - \hat{c}) \times (1 - H_M)$.

- CASE 3: $\hat{c} \in [-1, 0)$ and $H_{NM} + H_M \leq 1$:

  This case is such that $\max(0, (H_M + H_{NM} - 1)/H_M) = 0$.
  We have $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_M} = 1 - H_{NM} \times (1 + \hat{c})$ and $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_{NM}} = 1 - H_M \times (1 + \hat{c})$.

- CASE 4: $\hat{c} \in [-1, 0)$ and $H_{NM} + H_M > 1$:

  This case is such that $\max(0, (H_M + H_{NM} - 1)/H_M) = (H_M + H_{NM} - 1)/H_M$. We have $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_M} = (1 + \hat{c}) \times (1 - H_{NM})$ and $\frac{\partial \hat{H}_{\hat{c}}}{\partial H_{NM}} = (1 + \hat{c}) \times (1 - H_M)$.

In all cases, both partial derivatives are non-negative, which concludes the proof.

A.3 Definition of the Global MPI

Table A1: Succinct definition of the global MPI.

Dimension	Nutrition	Child mortality	Years of schooling	School attendance	Cooking fuel	Sanitation	Drinking water	Electricity	Housing	Assets
Weight w_j	1/6	1/6	1/6	1/6	1/18	1/18	1/18	1/18	1/18	1/18

Notes: See Alkire et al. (2022) for a complete definition of the Global MPI.

A.4 Depth-Sensitive Measures

In this Appendix, we formally define the depth-sensitive poverty measures used for the robustness analysis of Section 3.4.2.

Definition of Ideal and Single-Survey Depth-sensitive Poverty Measures.—We define a depth-sensitive ideal poverty measure, a characteristic not shared by H. A natural candidate is to use a poverty index that is *linear* in the depth of poverty. In the case of monetary poverty, the *poverty gap ratio* is linear in the depth of poverty, i.e.,

$$G_M(x^M) = \frac{1}{n'} \sum_{i \in Q'_z} \frac{z - m(x_i^M)}{z} \tag{8}$$

where Q'_z is the subset of individuals in N' whose monetary welfare is below the poverty line z, namely $Q'_z = \{i \in N' | m(x_i^M) < z\}$. In the case of non-monetary poverty, the *adjusted headcount ratio* is also linear in the depth of poverty, i.e.,

$$G_{NM}(x^{NM}) = \frac{1}{n''} \sum_{i \in Q''_k} s(x_i^{NM}) \tag{9}$$

where the deprivation score of individual i is $s(x_i^{NM}) = \sum_{j=k+1}^{\ell} w_j x_{ij}^{NM}$ and the subset of individuals in N'' whose deprivation score is at least k is denoted by $Q''_k = \{i \in N'' | s(x_i^{NM}) \geq k\}$. Both of these indices are used in practice. The latter is, for instance, the index defining the global MPI.

We now describe how we define our ideal multidimensional gap index, which we denote by G. This index is defined based on an individual *total deprivation score*, which we define as

$$\begin{aligned} d(x_i) &= s(x_i^{NM}) + w_m \times \frac{z - m(x_i^M)}{z} && \text{if } z \geq m(x_i^M) \\ d(x_i) &= s(x_i^{NM}) && \text{else} \end{aligned} \tag{10}$$

which sums i's deprivation score to her monetary gap, which is weighed with $w_m \geq 0$. The weight w_m thus tunes the importance for total deprivation of the monetary gap (relative to the non-monetary deprivation score). Let k^T denote the cutoff for the total deprivation score above which an individual is identified as multidimensionally poor. That is, i is identified as multidimensionally poor when $d(x_i) \geq k^T$.

We must select values for parameters z, k, w_m and k^T. Clearly, these choices are to some extent arbitrary. We select these values with the objective of maintaining some consistency with the analysis done in Section 3. More precisely, we select these values in such a way that:

(i) an individual who is identified as multidimensionally poor under H is also identified as multidimensionally poor under G

(ii) an individual who is not identified as multidimensionally poor under H is not identified as multidimensionally poor under G *unless* this individual is "almost" monetary poor *and* "almost" non-monetary poor.

The "unless" proviso in (ii) follows from the idea that someone who cumulates moderately low achievements in both monetary and non-monetary outcomes should be identified as poor, even though her low achievements in monetary (respectively non-monetary) outcomes are insufficient for her to qualify as monetary (respectively non-monetary) poor.

We select $k^T = k = 0.33$ to satisfy (i). Indeed, we have $k^T \leq k$ to ensure that an individual who is non-monetary poor is identified as multidimensionally poor even when her income is large. We also have $k^T \geq k$ to ensure that an individual whose income is large is not identified as multidimensionally poor when $s(x_i^{NM}) \leq k$.

We then select $w_m = k^T \times \frac{z}{z - z^{extr}}$ where $z^{extr} = 1.9$ and $z = 3.2$ to satisfy (i). These values imply that any individual whose income is below z^{extr} is identified as multidimensionally poor, even if she does not suffer from any non-monetary deprivations. To see this, observe that if individual i's income is $m(x_i^M) = z^{extr}$ and her deprivation score is $s(x_i^{NM}) = 0$, then her total deprivation score is $d(x_i) = k^T$. These values thus respect (i).

Observe that these values are also consistent with (ii). Some individuals are identified as multidimensionally poor under G, even though they are *not* identified as multidimensionally poor under H. Any individual i among these *additional multidimensionally poor* is such that

- $z^{extr} \leq m(x_i^M) < z$, and
- $0 < s(x_i^{NM}) \leq k$,

and of course $d(x_i) \geq k^T$. One could select for z another value other than 3.2. We take $z = 3.2$ for consistency with subsection 3.4.1. Overall, the values selected for our parameters are inspired by a preference-based definition of the multidimensionally poor (Decerf, 2023) (see section 3.2 in Decerf, 2024 for more details).

We can now define our *ideal multidimensional gap index* as

$$G(x) = \frac{1}{n} \sum_{i \in Q_{k^T}} d(x_i) \tag{11}$$

where the subset of multidimensionally poor individuals in N is denoted by $Q_{k^T} = \{i \in N | d(x_i) \geq k^T\}$.

Given the definition of G, the natural choice for *single-survey depth-sensitive poverty measures* are the poverty gap ratio G_M and the non-monetary adjusted headcount ratio G_{NM}. Given that G considers two different poverty lines, namely z and z^{extr}, we consider two different single-survey monetary poverty measures, namely G_M^z and $G_M^{z^{extr}}$.[20] One remark is on order. By definition, all non-monetary poor are multidimensionally poor, i.e., $Q_k \subseteq Q_{k^T}$. However, for some outcome vectors, some monetary poor are *not* multidimensionally poor, i.e., $Q_z \nsubseteq Q_{k^T}$. This is, for instance, the case of an individual i for whom $z^{extr} \leq m(x_i^M) < z$ and $s(x_i^{NM}) = 0$. Indeed, for the reasons explained above, we use a higher poverty line z than in Section 3. We now explain why we do not take $z = 1.9$ as in Section 3 together with a seemingly natural alternative definition for the multidimensional gap index G', which is

$$G'(x) = \frac{1}{n} \sum_{i \in Q_k \cup Q_z} d(x_i) \tag{12}$$

where $d(x_i)$ is again defined by Equation (10). The definition of G' again requires selecting a value for w_m, but observe that the definition of G' does not require two of the additional parameters that enter the definition of G, namely k^T and the second extreme poverty line z^{extr}. G' seems natural because the set of multidimensionally poor is $Q_k \cup Q_z$, which is exactly the same set as under H. One of the problems with G' is that a multidimensionally poor individual $i \in Q_k$ who has a large income and the minimal deprivation score to be non-monetary poor, namely $s(x_i^{NM}) = k$ and $m(x_i^M) > z$, has a much larger total deprivation score than another multidimensionally poor individual $j \in Q_z$ who has no non-monetary deprivations and the minimal income to be monetary poor, namely $m(x_j^{NM}) = z - \epsilon$ and $s(x_j^{NM}) = 0$. Indeed, we have $d(x_i) = k$ but $d(x_j) \approx 0$ when ϵ comes close to zero. This shows that G' attributes an infinitely larger importance to the non-monetary poverty *status* than to the monetary poverty *status*, regardless of the value selected for w_m. The problem comes from the fact that the value for the adjusted headcount ratio G_{NM} "jumps" when an individual escapes non-monetary poverty status (Decerf, 2025). (In contrast, the value for the monetary gap G_M changes continuously when an individual escapes monetary poverty status, because her monetary gap tends to zero when her income tends to z.)

Definition of Multi-survey Variants.—We now turn to the definition of our multi-survey depth-sensitive poverty measures. We start with our two different mash-up poverty measures. Similarly to Section 3, the mash-up poverty measures attribute equal weights to the two single-survey poverty measures, namely

$$\hat{G}_z^{1/2}(x^M, x^{NM}) = \frac{1}{2} G_M^z(x^M) + \frac{1}{2} G_{NM}(x^{NM}), \tag{13}$$

$$\hat{G}_{z^{extr}}^{1/2}(x^M, x^{NM}) = \frac{1}{2} G_M^{z^{extr}}(x^M) + \frac{1}{2} G_{NM}(x^{NM}). \tag{14}$$

Finally, we turn to our made-up poverty measures. As in Section 3, made-up measures

[20] Both $G_M^z(x^M)$ and $G_M^{z^{extr}}(x^M)$ are defined from Equation (8), but the latter replaces z by z^{extr}.

correspond to the ideal measure when computed based on an "assumed" value for the missing parts of the joint distribution. In a nutshell, a complete "synthetic" database (where both monetary and non-monetary outcomes are available for each individual) is generated from the two surveys by assuming a value for the (unobserved) dependence, captured by a correlation coefficient $corr \in [-1,1]$, between monetary gaps ($\frac{z-m(x_i^M)}{z}$) and non-monetary gaps ($s(x_i^{NM})$). We thus have

$$\hat{G}_{\hat{corr}}(x^M, x^{NM}) = G\left(\hat{X}\left(x^M, x^{NM}, \hat{corr}\right)\right) \tag{15}$$

where the mapping $\hat{X}$ generates a complete "synthetic" database based on its three inputs.

Imputation Method.—In contrast to headcount-based poverty measures, computing $\hat{G}_{\hat{corr}}(x^M, x^{NM})$ requires micro-level information on the joint distribution of monetary and non-monetary deprivation. Because this joint distribution is not observed when data sources are split, we simulate it using a copula-based imputation strategy. First, we assign each household a quantile rank based on its observed monetary gap or non-monetary score. Next, we generate pairs of correlated uniform random variables using a Gaussian copula, parameterized to reflect a target correlation structure. We then match these quantiles back to the observed distributions to assign each household a synthetic monetary score that corresponds to its non-monetary rank.[21]

This procedure yields a fully synthetic joint distribution that preserves the observed partial distributions while embedding a controlled degree of dependence. As an alternative, we also explore generating monetary scores from a fitted log-normal distribution, though this yields results nearly identical to the re-shuffling approach. Finally, we compute depth-sensitive poverty measures based on the synthetic data and evaluate their sensitivity to changes in correlation.

A.4.1 Details on Imputation and Simulation Procedure

To explore how varying the correlation between monetary and non-monetary dimensions affects depth-sensitive poverty measures, we use a copula-based imputation strategy. Since the joint distribution of these two dimensions is typically unobserved in practice — due to separate survey sources — we impute monetary (e.g., income or consumption) data into the non-monetary survey. Unlike conventional survey-to-survey imputation techniques, which rely on complex statistical matching based on covariates and conditional distributions, our approach is tailored for hypothetical scenarios and relies on a simple, transparent Monte Carlo procedure.

The simulation process involves four steps. First, we assign each household a quantile rank in both the monetary and non-monetary dimensions, based on their position within

[21]For further details on the imputation and simulation procedure, see A.4.1.

the national (or regional) distribution.[22] Second, we generate two correlated uniform random variables using the NORTA (NORmal To Anything) method.[23] These uniform variables preserve the target correlation and are used to link synthetic monetary scores to observed non-monetary ranks.

Third, using the generated pair of correlated uniform draws, we assign each household a new monetary quantile that corresponds to its non-monetary rank. This synthetic rank is then matched to an actual monetary value drawn from the empirical distribution.[24] As an alternative, we also explore assigning synthetic monetary values by sampling from a log-normal distribution fit to the observed monetary data.[25] Both methods yield nearly identical results, though the re-ranking approach more directly preserves the observed empirical partial distributions.

Finally, we compute household-level poverty indices using the synthetic monetary and observed non-monetary scores. These are then aggregated to produce regional poverty measures across a range of assumed correlation scenarios. This approach allows us to evaluate the sensitivity of multidimensional poverty metrics to alternative dependence structures between dimensions, while holding the partial distributions constant.

[22]We use 1,000 quantile bins. In cases where values are clustered and ties arise at quantile thresholds, we randomly assign households across the relevant bins to ensure an even distribution.

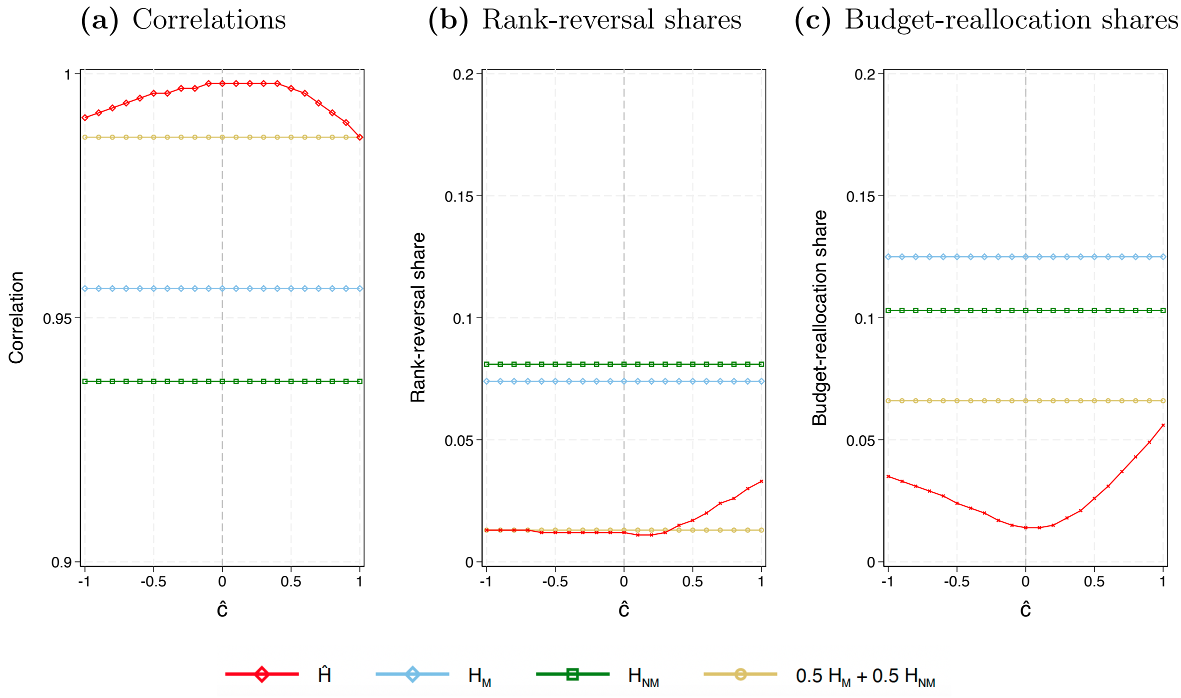
[23]This involves drawing from a bivariate normal distribution with a pre-specified correlation coefficient, then transforming the partial distributions to uniform via the standard normal cumulative distribution function (CDF).

[24]When multiple monetary values correspond to a given quantile, one is drawn at random. In rare cases where no exact match is found — especially in small samples — we fallback to a coarser 100-quantile binning.

[25]We estimate the parameters μ_M and σ_M of the log-normal distribution from the observed monetary scores and simulate accordingly.

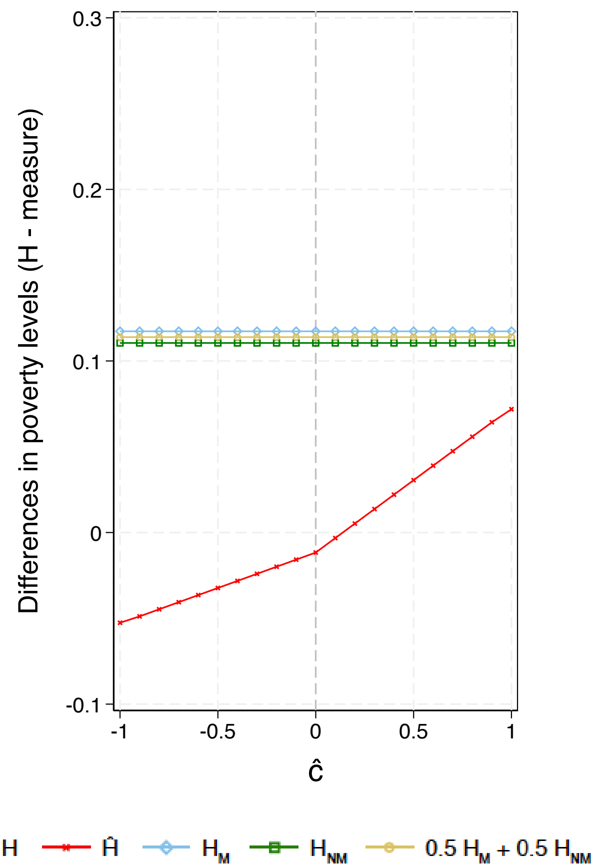
A.5 Additional Tables and Figures

Figure A1: Performance of Single- Versus Multi-Survey Poverty Measure Excluding Extreme Values of c



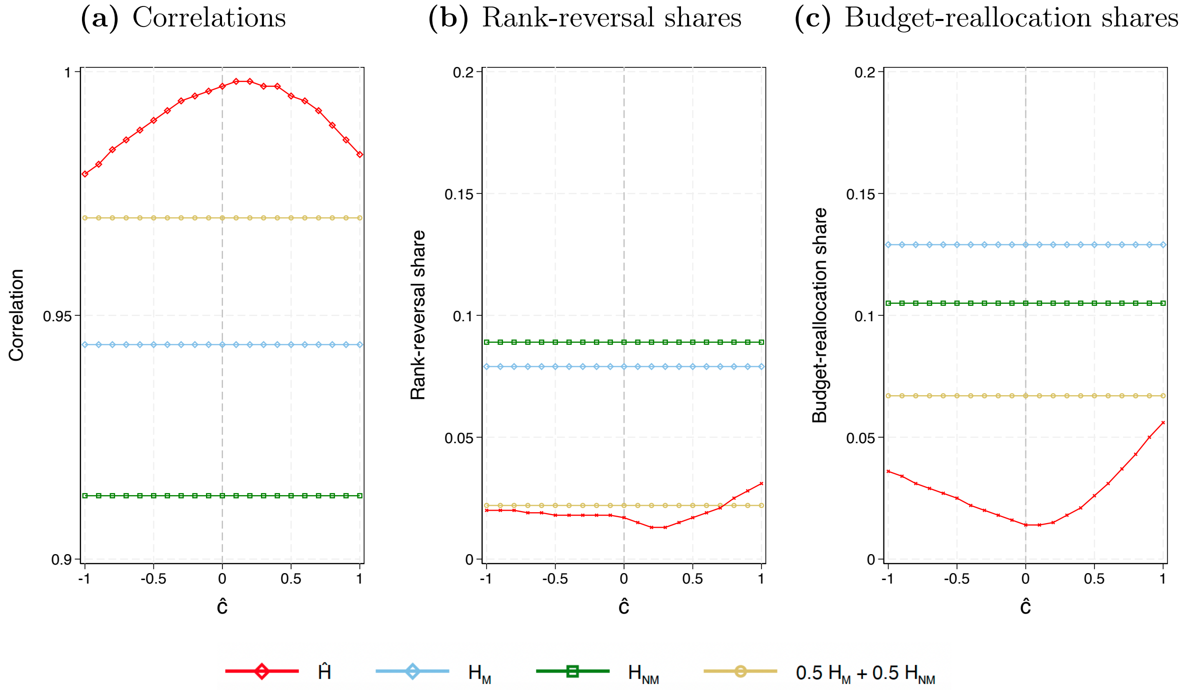
Notes: The figure replicates Figure 3 but excludes observations with extreme values of c (i.e., 1 and -1).

Figure A2: Accuracy of Poverty Measures by Assumed Overlap Coefficient



Notes: The figure plots the difference between the ideal multidimensional headcount H and each constrained poverty measure as a function of the assumed overlap coefficient $\hat{c}$. Positive values indicate that the constrained measure underestimates poverty relative to the ideal headcount, while negative values indicate overestimation. To construct the level differences, we first use household survey weights to compute regional indicators. The differences are then calculated on the unweighted pooled sample of regions, giving each region equal weight.

Figure A3: Performance of Single- Versus Multi-survey Poverty Measures. z = \$3.20, k = 20%

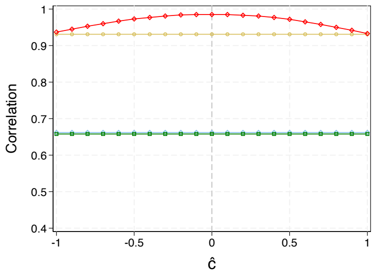


Notes: The figure plots correlations, rank-reversal shares, and budget-reallocation shares for alternative poverty thresholds, as functions of the assumed overlap coefficient ($\hat{c}$).

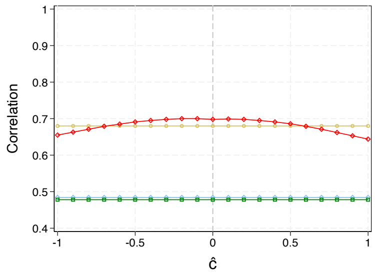
Figure A4: Correlations in Scenarios A-C

Panel A: Scenario A (variance-only)

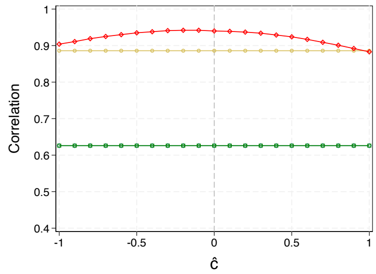
(a) Low c — High H_M/H_{NM}



(b) High c — Low H_M/H_{NM}

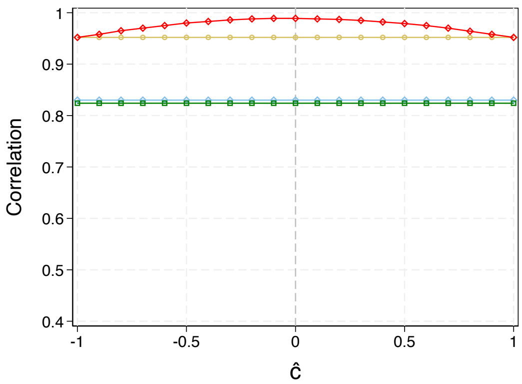


(c) High c — High H_M/H_{NM}

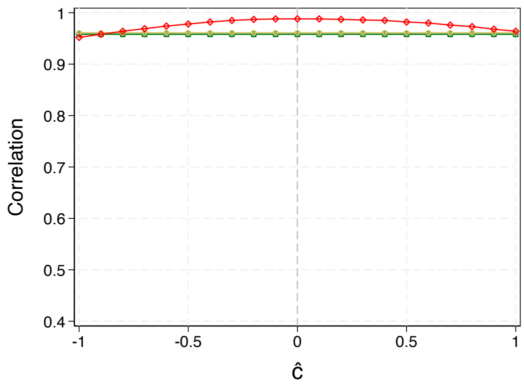


Panel B: Scenario B (correlation-only)

(d) $\rho = 0.5$

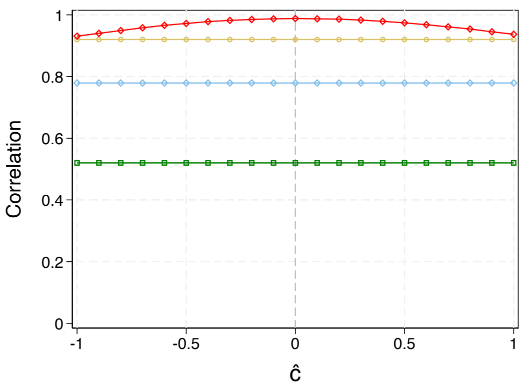


(e) $\rho = 0.99$

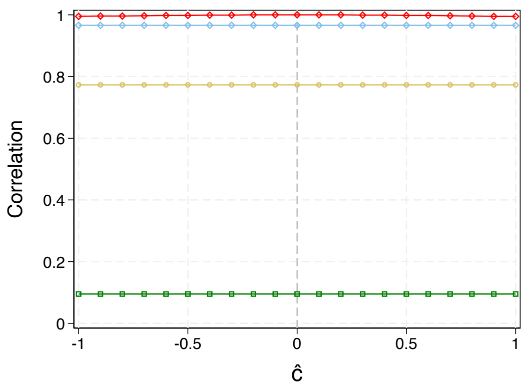


Panel C: Scenario C (gap-only)

(f) Gap 0.2

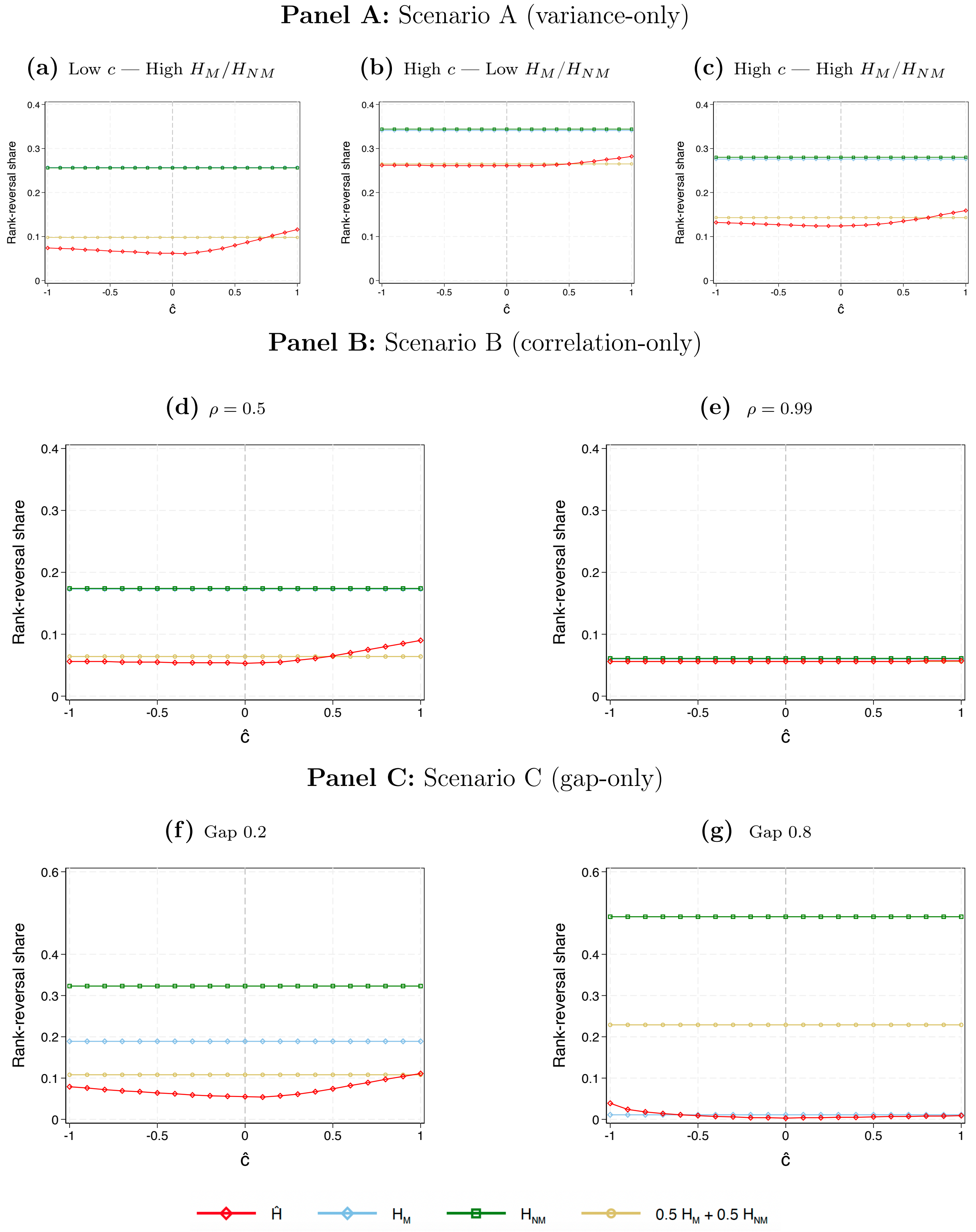


(g) Gap 0.8



Notes: The figure displays the correlation between the ideal multidimensional headcount H and each constrained measure (made-up, mash-up, and single-survey poverty measures) as a function of the assumed overlap coefficient $\hat{c}$. Panel A corresponds to scenario A (variance-only), Panel B presents results from scenario B (correlation-only), and Panel C displays scenario C (gap-only). Subpanels (a)–(g) correspond to the scenario variants described in Section 4.1.

Figure A5: Rank-Reversal Shares in Scenarios A-C



Notes: The figure displays the share of rank reversals between the ideal multidimensional headcount H and each constrained measure (made-up, mash-up, and single-survey poverty measures) as a function of the assumed overlap coefficient $\hat{c}$. Panel A corresponds to scenario A (variance-only), Panel B presents results from scenario B (correlation-only), and Panel C displays scenario C (gap-only). Subpanels (a)–(g) correspond to the scenario variants described in Section 4.1.