



IARIW 2025

IARIW 2025

Thursday, October 2 & Friday, October 3

Multidimensional Fusion

Minh Nguyen

Ben Brunckhorst

Christopher Lakner

Hai-Anh Dang

Nishant Yonzan

(World Bank)

Paper prepared for the IARIW–World Bank–UEB/VNU Conference on “Improving Well-being Measurement in Data-challenged Environments in Developing Countries for Better Evidence-based Policies” October 2-3, 2025

Parallel Session 6B Surveys and Methods
Time Slot: Friday, October 3, 1:00-3:00 PM

Multidimensional Fusion

Ben J. Brunckhorst, Hai-Anh Dang, Christoph Lakner, Minh C. Nguyen,
Nishant Yonzan¹

September 19, 2025

*Preliminary draft prepared for the IARIW-World Bank-UEB/VNU Special Conference
“Improving Well-being Measurement in Data-Challenged Environments in Developing
Countries for Better Evidence-based Policies.” Please do not cite.*

Multidimensional poverty metrics offer a more complete picture of deprivation than monetary poverty, yet their adoption has been limited by data constraints. This paper introduces a novel approach to estimate multidimensional welfare metrics by fusing summary statistics for a subset of indicators derived from different data sources. We validate the method by simulating typical missing data scenarios using 573 national household surveys from 112 countries (1989–2024) that collect the indicators that are included in the World Bank’s Multidimensional Poverty Measure. The fusion method accurately predicts multidimensional metrics in typical missing data scenarios even in the presence of (simulated) sample bias reflecting differences between surveys in the real world. For infrastructure indicators, the approach yields an average absolute error of about 1 percentage point in estimating the share of households—roughly 45 percent—lacking at least one basic service, and about 0.5 percentage point error for the multidimensional poverty headcount of 26 percent. Sensitivity analyses show that the mean absolute error is minimal when pseudo-data match the true mean and rises asymmetrically as sampling or data-source bias increases, with upward bias producing larger errors than equivalent downward bias. The method offers a way to estimate multidimensional poverty for more countries, consider more dimensions, and improves the ability to predict multidimensional welfare metrics over time.

Keywords: Welfare measurement, multidimensional poverty, missing data, data fusion, conditional independence, joint distribution

JEL codes: I32, C81, C82, C13

¹ All authors are with the World Bank. The authors gratefully acknowledge financial support from the UK government through the Data and Evidence for Tackling Extreme Poverty (DEEP) Research Programme.

1. Introduction

Poverty is commonly understood as a lack of income or consumption. Monetary poverty alone provides an incomplete picture of the poor. Poverty is multifaceted and has been conceptualized as a deprivation in capabilities (Sen, 1993). Multidimensional measures of poverty have emerged that consider deprivations across several dimensions, such as health, education and access to basic infrastructure, and how these deprivations overlap (Alkire et al. 2015). These measures claim to inform more impactful multisectoral interventions by better revealing what makes people poor, how deprivations are related, and the intensity of poverty.

While a multidimensional approach to poverty measurement is clearly preferable from a conceptual point of view, it has not been widely adopted. The lack of adoption in practice is partly a result of the lack of data on multiple dimensions at the household or individual level in the same household survey. This is especially severe at the global level, where the same indicator needs to be available across a large variety of countries. Hence, only a handful of indicators representing few dimensions can be used to monitor multidimensional poverty at a global level. In addition to the variation across countries, the information collected by household surveys differs also within countries over time. Another challenge is that comprehensive surveys are conducted with large lags in developing countries, making it difficult to update multidimensional measures for years or even decades.

This paper studies the World Bank's Multidimensional Poverty Measure (MPM), which consists of three dimensions: income (or consumption), access to education, and access to basic infrastructure. In its latest year of 2021, the measure could only be estimated for 3 out of 7 regions in the world using data collected between 2018 and 2024. The latest MPM covers less than a third of the population in East Asia & Pacific, Middle East & North Africa, South Asia, and Sub-Saharan Africa. These four regions account for 95 percent of the extreme poor (as defined as individuals living on less than \$2.15 per day in 2017 PPP) in 2021 (World Bank 2024a, 2024b). Globally, roughly 4 in 10 households are represented by the MPM circa 2021, most of them in countries with relatively higher living standards. This is not due to a lack of data in general. In the same timeframe when considering each indicator by itself, information is available for about three-quarters of the global population regarding monetary poverty, 70 percent of the global population regarding education, and 60 percent of the global population regarding basic infrastructure. But the key challenge is that these dimensions are not all observed in the same survey.

In addition, there has been progress on methods to fill gaps over time in monetary poverty data. The World Bank's Poverty and Inequality Platform (PIP) extrapolates and interpolates country-level poverty rates using growth in national accounts to report global and regional statistics for a common year (see Mahler et al. 2022). No equivalent "lineup" methodology is used for the MPM. Instead, estimates use data collected anytime in a seven-year window (+/- 3 years) centered on a "circa" reporting year.

In what follows, we outline a method to estimate multidimensional welfare metrics that relies only on aggregate level statistics and incomplete information about the dependency between deprivations, or their partial joint distributions. Specifically, the share of the population deprived on

any combination of dimensions can be decomposed as the product of the marginal probability of being deprived on a single dimension and conditional probabilities of also being deprived on other dimensions using the chain rule. We show that accurate estimates of multidimensional poverty can be recovered with partial knowledge of the joint distribution of deprivations by assuming some degree of conditional independence.

Our approach is less data intensive than more complex survey imputation and simulation methods. It can be feasibly applied in data constrained contexts. The aggregate deprivation rates and the partial joint distributions can be derived from multiple data sources or modelled, rather than relying on one comprehensive household survey. In addition, our simple probabilistic fusion approach can be used to line-up multidimensional welfare measures across countries to a common reference year. Multidimensional metrics for a line-up year can be estimated by combining gap-filled aggregate deprivation rates with conditional probabilities between a subset of deprivations derived from existing microdata.

In summary, the method we propose: (1) provides a way to estimate multidimensional poverty metrics when some dimensions are missing, (2) allows for incorporating additional dimensions that were previously excluded due to data limitations, and (3) improves the ability to fill gaps in time series of multidimensional welfare metrics.

We validate the method using harmonized microdata from the World Bank's Global Monitoring Database (GMD) and report the accuracy across various scenarios where data might be missing.

The literature has proposed various methods to impute missing data or variables. Ummel et al (2024) use a statistical data fusion model to combine information from different surveys at the household level. They show an example for the US where both surveys sample an underlying population. They predict households' response in one survey using demographic and other characteristics from the other survey. Matching various household characteristics, Roy and van der Weide (2024) impute consumption aggregate for India from a survey with consumption to another without consumption to estimate the trends in consumption in years when only the latter survey is available. Piketty, Saez, and Zucman (2018) propose a method to fuse income in household surveys with income from administrative data sources to generate an income distribution that is consistent with the national accounts.

Our approach to estimate multidimensional poverty metrics puts less emphasis on imputing an indicator at the household level but instead focuses on predicting prevalence of that indicator at some population group level – that is, national, rural/urban, regional, or other population sub-group where statistics are representative. This is the statistic that is reported and used to inform policy.

2. Method

Theory

The method to fuse datasets for multidimensional poverty estimates is based on probability theory. The share of households deprived on two indicators X_1, X_2 can be expressed as a joint probability $P(X_1, X_2)$. This joint probability can be expressed as a product of conditional probabilities using the chain rule:

$$P(X_1, X_2) = P(X_1) \cdot P(X_2|X_1) = P(X_2) \cdot P(X_1|X_2)$$

This chain rule factorization underlies many models (e.g. Bayesian networks) and is often used to break a complex joint probability into simpler pieces. The joint probability of three indicators can similarly be decomposed as:

$$P(X_1, X_2, X_3) = P(X_1, X_2) \cdot P(X_3|X_1, X_2) = P(X_1) \cdot P(X_2|X_1) \cdot P(X_3|X_1, X_2)$$

Notice that the order of conditional probabilities in the equations above does not matter. We can equally calculate the joint probability of three indicators as:

$$P(X_1, X_2, X_3) = P(X_1, X_3) \cdot P(X_2|X_1, X_3) = P(X_1) \cdot P(X_3|X_1) \cdot P(X_2|X_1, X_3)$$

Calculating the joint probability of any combination of indicators using the chain rule, or calculating a multidimensional poverty metric such as the MPM, requires knowledge of all conditional probabilities in this expression. In practice, these are known only when all indicators are measured for the same households or individuals (in the same survey). When indicators are measured in different data sources, more information on the conditional independence is necessary to estimate this joint probability expression.

Suppose X_1 and X_3 are not available from the same (nationally representative) surveys but come from different (also nationally representative) surveys which collect information on X_2 . If we assume X_3 is conditionally independent of X_1 given X_2 :

$$P(X_3|X_1, X_2) = P(X_3|X_2)$$

The joint probability can now be expressed as:

$$P(X_1, X_2, X_3) = P(X_1, X_2) \cdot P(X_3|X_2)$$

Here, conditional independence implies that among those deprived on X_2 , the probability of being deprived on X_3 is the same regardless of whether one is deprived on X_1 .

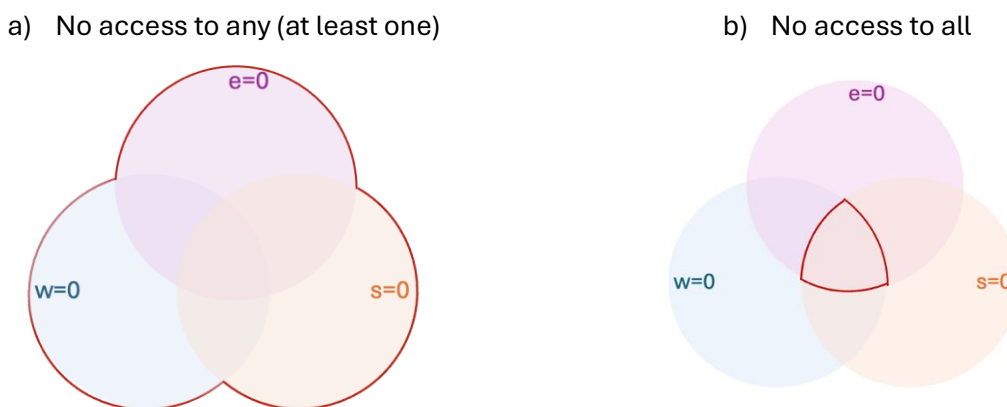
We hypothesize that assuming some degree of conditional independence is reasonable to estimate multidimensional poverty metrics because dimensions of poverty are often related or driven by common underlying factors. Recent research has confirmed that the structure of multidimensional poverty networks is similar across countries and stable over time (Stojkoski et al. 2025). Therefore, it might be possible to accurately recover the joint distribution of deprivations when only a subset of conditional probabilities between indicators are known.

To test our hypothesis, we use a large set of harmonized microdata from the World Bank’s Global Monitoring Database (GMD).² We compare the true value of multidimensional poverty metrics to those estimated in scenarios where the relationship between deprivations is only partially known. These scenarios simulate situations where the indicators needed to measure multidimensional poverty are not collected in a single survey, but available from different sources. This allows us to identify the conditions under which our method produces reliable multidimensional poverty estimates (see Data and Validation for more details).

Consider the following illustration to clarify the method

Consider the share of the population without access to basic infrastructure by using three binary indicators: access to electricity (e), improved drinking water (w), and improved sanitation (s). Specifically, we would like to know the share of the population without access to *at least one* of these services (Figure 1, panel a), and the share without access to *all* three of them (Figure 1, panel b). Understanding how these deprivations overlap is important because investments in infrastructure need to be coordinated.

Figure 1 – three partially overlapping deprivations



Note: the solid red line encircles the population share of interest.
 Source: Authors’ illustration.

At first glance, it appears the recent household consumption survey only collected data on access to electricity. However, we have statistics on access to drinking water and sanitation from separate

² The Global Monitoring Database (GMD) is the World Bank’s repository of multitopic income and expenditure household surveys used to monitor global poverty and shared prosperity. The household survey data are typically collected by national statistical offices in each country, and then compiled, processed, and harmonized. The process is coordinated by the Data for Goals (D4G) team and supported by the six regional statistics teams in the Poverty and Equity Global Practice. Global Poverty & Inequality Data Team (GPID) in Development Economics Data Group (DEC DG) also contributes historical data from before 1990, and recent survey data from Luxemburg Income Studies (LIS). Selected variables have been harmonized to the extent possible such that levels and trends in poverty and other key sociodemographic attributes can be reasonably compared across and within countries over time. The GMD’s harmonized microdata are currently used in Poverty and Inequality Platform (PIP), World Bank’s Multidimensional Poverty Measures (WB MPM), the Global Database of Shared Prosperity (GDSP), and Poverty, Prosperity and Planet Reports.

reports. We will refer to this scenario e-w-s, since information on each indicator comes from different places as shown in Figure 2.

Figure 2 – scenario e-w-s



Source: Authors’ illustration; calculations by the authors.

As explained above, the joint probability of not having access to all three services can be expressed using the chain rule:

$$P(e = 0, w = 0, s = 0) = P(e = 0) \cdot P(w = 0|e = 0) \cdot P(s = 0|e = 0, w = 0)$$

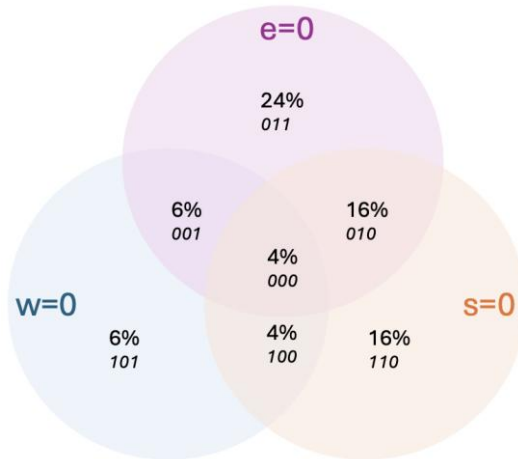
This may be written in any other order. In the e-w-s scenario, we don’t know anything about the conditional probability of (not) having access to one type of infrastructure given access to another type. We know only $P(e)$, $P(w)$ and $P(s)$. If we assume all indicators are statistically independent, e.g., $P(w = 0) = P(w = 0|e = 0)$, it is possible to simplify the expression as the product of these marginal probabilities:

$$P(e, w, s) = P(e) \cdot P(w) \cdot P(s)$$

This is a strong assumption, since someone without access to electricity seems more likely not to have access to other basic infrastructure. Nonetheless, assuming independence allows us to crudely estimate the joint probability distribution and calculate multidimensional metrics of interest (Figure 3). We estimate the share of the population without access to at least one service is 76% and the share without access to all three is 4%.

Figure 3 – scenario e-w-s joint probability distribution

e	w	s	P(e)*P(w)*P(s)	P(e,w,s)
0	0	0	0.5*0.2*0.4	0.04
1	0	0	0.5*0.2*0.4	0.04
0	1	0	0.5*0.8*0.4	0.16
0	0	1	0.5*0.2*0.6	0.06
1	1	0	0.5*0.8*0.4	0.16
1	0	1	0.5*0.2*0.6	0.06
0	1	1	0.5*0.8*0.6	0.24
1	1	1	0.5*0.8*0.6	0.24

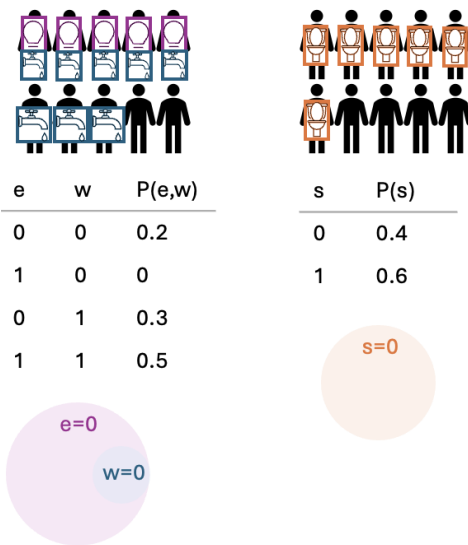


Source: Authors’ illustration; calculations by the authors.

Assume instead that the access to drinking water indicator can be constructed from existing variables in the household survey with access to electricity. In this scenario, we know the population jointly deprived in electricity and water $P(e = 0, w = 0)$, and the conditional probability $P(w = 0|e = 0) = \frac{P(w=0,e=0)}{P(e=0)}$, because data was collected on both indicators from the same households (Figure 4). We can estimate the joint distribution in this “ew-s” scenario assuming only that access to sanitation is not correlated with access to electricity or drinking water, or formally $P(s) = P(s|e, w)$:

$$P(e, w, s) = P(e, w) \cdot P(s) = P(e) \cdot P(w|e) \cdot P(s)$$

Figure 4 – scenario ew-s

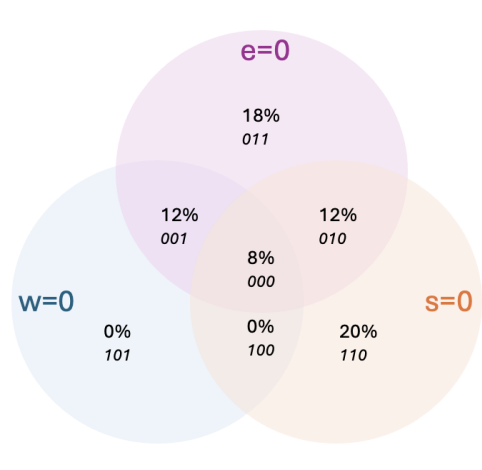


Source: Authors’ illustration; calculations by the authors.

In this specific example, the data indicate all 20% of the population without access to drinking water also lack access to electricity. Compared to the previous scenario, our new results show that a lower share of the population lack access to any of the services (70%), but a higher share lack access to all three (8%), shown in Figure 5.

Figure 5 – scenario ew-s joint probability distribution

e	w	s	P(e,w)*P(s)	P(e,w,s)
0	0	0	0.2*0.4	0.08
1	0	0	0*0.4	0.00
0	1	0	0.3*0.4	0.12
0	0	1	0.2*0.6	0.12
1	1	0	0.5*0.4	0.20
1	0	1	0*0.6	0.00
0	1	1	0.3*0.6	0.18
1	1	1	0.5*0.6	0.30

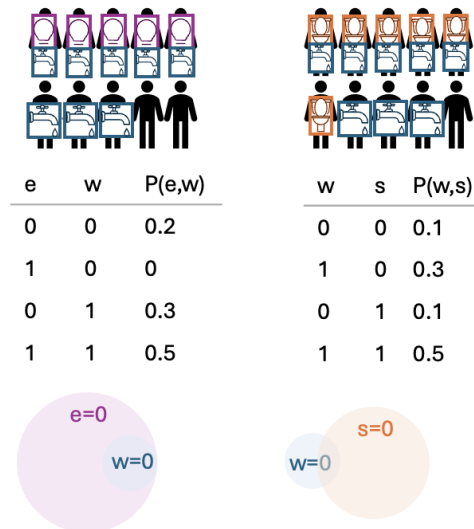


Source: Authors’ illustration; calculations by the authors.

Assume further that the survey with information on access to sanitation also collected information on access to drinking water (Figure 6). The scenario can now be described as “ew-ws” and we can use information about how e and s are related to w to better estimate the joint distribution:

$$P(e, w, s) = P(e, w) \cdot P(s|w) = P(e, w) \cdot \frac{P(w, s)}{P(w)}$$

Figure 6 – scenario ew-ws

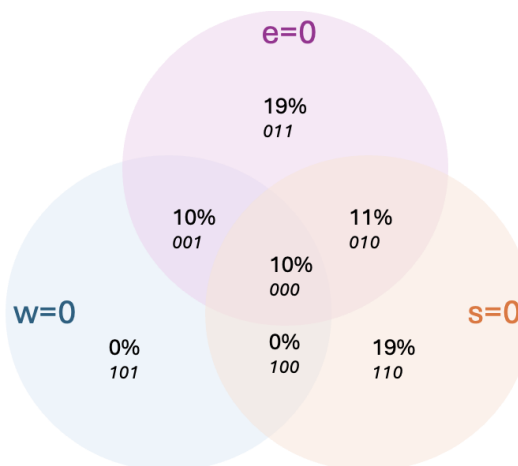


Source: Authors’ illustration; calculations by the authors.

Here we make a weaker assumption of *conditional* independence. Formally, we assume $P(s|e, w) = P(s|w)$, which implies that there is no added information on access to sanitation from information on access to electricity, given we know about access to water. With this assumption, 69% of the population do not have access to at least one service and 10% do not have access to all three.

Figure 7 – scenario ew-s joint probability distribution

e	w	s	$P(e,w)*P(s w)$	$P(e,w,s)$
0	0	0	$0.2*0.1/0.2$	0.10
1	0	0	$0*0.1/0.2$	0.00
0	1	0	$0.3*0.3/0.8$	0.11
0	0	1	$0.2*0.1/0.2$	0.10
1	1	0	$0.5*0.3/0.8$	0.19
1	0	1	$0*0.1/0.2$	0.00
0	1	1	$0.3*0.5/0.8$	0.19
1	1	1	$0.5*0.5/0.8$	0.31



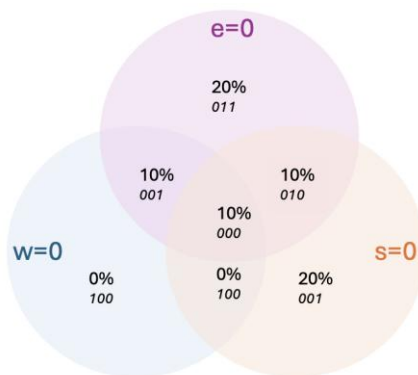
Source: Authors’ illustration; calculations by the authors.

With all three indicators in the same survey, we can compare the results of the predictions above making assumptions of (conditional) independence to the actual data. We find that the error in the share of the population without access to any or all basic services was no more than one percentage point in the “ew-ws” scenario with the access to water indicator in both datasets (Figure 8).

The validation exercises we conduct in this paper follow the same logic to understand the data scenarios when a probabilistic fusion method, assuming conditional independence between a subset of indicators, produces reliable multidimensional poverty metrics.

Figure 8 – true joint probability distribution and comparison to missing data scenarios

e	w	s	$P(e,w,s)$
0	0	0	0.10
1	0	0	0.00
0	1	0	0.10
0	0	1	0.10
1	1	0	0.20
1	0	1	0.00
0	1	1	0.20
1	1	1	0.30



Target	Scen.	Prob.	Error
Any	e-w-s	0.76	0.06
Any	ew-s	0.70	0.00
Any	ew-es	0.69	-0.01
All	e-w-s	0.04	-0.06
All	ew-s	0.08	-0.02
All	ew-es	0.10	0.00

Source: Authors’ illustration; calculations by the authors.

Probabilistic fusion algorithm

This section describes an algorithm that generalizes the method illustrated above to estimate the joint probability distribution of many indicators from multiple datasets. This algorithm uses all information about the conditional probabilities between indicators that are available in the data and otherwise makes assumptions of conditional independence.

Suppose we have L datasets D_1, D_2, \dots, D_L , each containing probability distributions over subsets of indicators. The fusion process proceeds as follows:

1. Start with the primary dataset D_1 , which provides the baseline joint probability $P(X)$ over a set of indicators X_1, X_2, \dots, X_K .
2. For each subsequent dataset D_l ($l = 2, \dots, L$),
 - a. Identify the set of common indicators X_C in D_l and the baseline joint distribution, such that X_C is a subset of X (e.g., X_2, X_3).
 - b. Calculate the conditional probabilities for the remaining indicators Y in D_l given the common indicators X_C :

$$P(Y|X_C) = P(Y, X_C)/P(X_C)$$

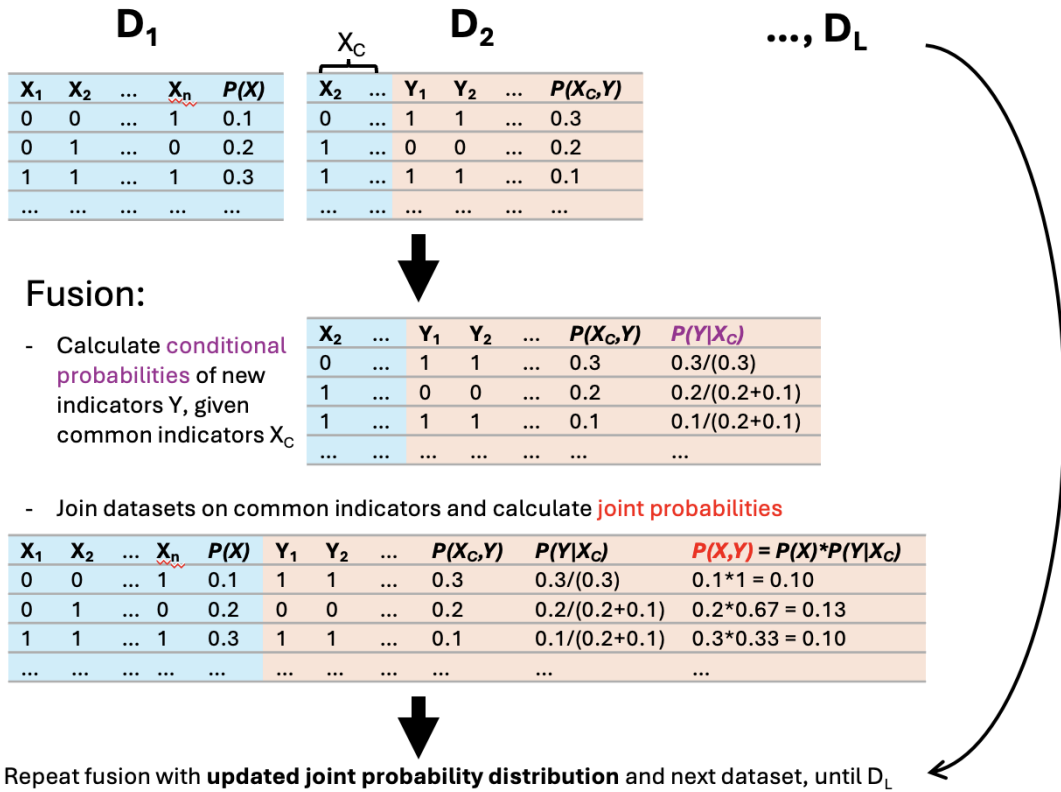
- c. Join the baseline joint distribution and D_l on the common indicators X_C . This maps out every combination of X and Y defined by the probability distribution.
- d. Calculate the joint probabilities for every combination of X and Y by multiplying the corresponding marginal and conditional probabilities:

$$P(X, Y) = P(X) \cdot P(Y|X_C)$$

3. Repeat step 2 for all datasets, updating the baseline joint distribution at each iteration until all indicators are included.

Figure 9 visualizes the fusion algorithm.

Figure 9 – probabilistic fusion algorithm



Source: Authors' illustration; calculations by the authors.

3. Data and Validation

Overview

This section describes the data and approach used to validate the performance of the probabilistic fusion method to estimate multidimensional poverty metrics based on independent datasets. We use a large set of harmonized microdata from the World Bank’s Global Monitoring Database (GMD), which is used to estimate the MPM in practice. We then compare the true value of multidimensional poverty metrics to those estimated in “missing data” scenarios where indicators and the relationship between them are derived from different sources.

Data

We use harmonized microdata from the World Bank’s Global Monitoring Database (GMD), constructing the variables used by the MPM and extracting standard design and population group variables (weights, strata, PSUs, rural/urban, subnational ADM1). We use data from 573 surveys covering 112 countries collected between 1989 and 2024. The GMD data are harmonized and standardized as much as possible, but differences exist with regards to the method of data collection, different survey instruments, and whether the welfare aggregate is based on income or consumption.

The MPM is composed of six indicators: consumption or income, educational attainment, educational enrollment, drinking water, sanitation, and electricity. These are mapped into three dimensions of well-being: monetary, education, and basic infrastructure services. The MPM dimensions, indicators and weights are shown in Table 1.

Table 1 – MPM dimensions and indicators

Dimension	Indicator	Weight	Ref
Monetary	Daily consumption or income is less than \$ 3.00 per person (2021 PPP).	1/3	p
Education	No adult in the household (age of grade 9 or above) has completed primary education.	1/6	c
	At least one school-age child up to the age of grade 8 is not enrolled in school.	1/6	r
Access to basic infrastructure	The household has no access to electricity.	1/9	e
	The household lacks access to limited-standard drinking water.	1/9	w
	The household lacks access to limited-standard sanitation.	1/9	s

We aggregate the harmonized microdata to population group statistics describing the joint probability distribution of the MPM indicators, that is the share of the population with each 0/1/missing combination (see Figure 8 for example). These statistics are calculated at several levels of aggregation (national, rural/urban, subnational and quintile) to better understand how disaggregation of summary statistics influences the performance of multidimensional fusion. All calculations use survey weights in line with standard practice; when aggregating subgroup predictions to the national level we use the appropriate population shares.

Target variables

The validation proceeds as follows. First, we calculate several multidimensional metrics of interest from the true probability distribution derived from GMD. These are the true values of the target variables we will use to evaluate the performance of the fusion method across different scenarios and are calculated at the national level. We consider two sets of target variables defined in Table 2, where $i = 1$ denotes being deprived on indicator i . Iverson notation is used in formulas to express a sum over a subset of joint probabilities, defined by the condition in square brackets.

The first set represents the simpler case considering the joint distribution of three basic infrastructure indicators, similar to the illustration in the method section. The second are derived from the joint distribution of all indicators used by the MPM and include the official MPM (a weighted headcount metric), an application of the Alkire-Foster method (Alkire and Foster 2011) multiplying the weighted headcount by the average deprivation rate of the poor, and the share of people deprived on at least n dimensions. Since these target variables summarize the probability distribution in different ways, they might be more or less sensitive to the same assumptions of conditional independence made by the probabilistic fusion method.

Table 2 – Target variables

Three infrastructure indicators	
Deprived in at least 1 (any)	$= \sum_{e,w,s \in \{0,1\}} P(e, w, s)[e + w + s \geq 1] = 1 - P(0,0,0)$
Deprived in at least 2	$= \sum_{e,w,s \in \{0,1\}} P(e, w, s)[e + w + s \geq 2]$ $= P(1,1,0) + P(1,0,1) + P(0,1,1) + P(1,1,1)$
Deprived in all three	$= \sum_{e,w,s \in \{0,1\}} P(e, w, s)[e + w + s = 3] = P(1,1,1)$
Six MPM indicators	
MPM (weighted headcount)	$= \sum_{i=p,c,r,e,w,s \in \{0,1\}} P(i) \left[\sum_{i=p,c,r,e,w,s \in \{0,1\}} i \cdot w_i \geq \frac{1}{3} \right]$
AF method (headcount * intensity)	$= \sum_{i=p,c,r,e,w,s \in \{0,1\}} P(i) \cdot i \cdot w_i \left[\sum_{i=p,c,r,e,w,s \in \{0,1\}} i \cdot w_i \geq \frac{1}{3} \right]$
Deprived in at least 1 (any)	$= \sum_{i=p,c,r,e,w,s \in \{0,1\}} P(i) \left[\sum_{i=p,c,r,e,w,s \in \{0,1\}} i \geq 1 \right]$
Deprived in at least 3	$= \sum_{i=p,c,r,e,w,s \in \{0,1\}} P(i) \left[\sum_{i=p,c,r,e,w,s \in \{0,1\}} i \geq 3 \right]$

The three MPM dimensions are weighted equally, and within each dimension each indicator is also weighted equally. Individuals are considered multidimensionally deprived if they fall short of the threshold in at least one dimension or in a combination of indicators equivalent in weight to a full dimension (Table 1). In other words, households will be considered poor if they are deprived in indicators whose weight adds up to 1/3 or more (see Table 2). Because the monetary dimension is measured using only one indicator, anyone who is income poor is also poor under the MPM. If indicators for any one dimension are partially missing, the weights for the remaining indicators for that dimension are adjusted so they sum to 1/3 and are equal. If all indicators for a dimension are missing, the MPM cannot be defined. In practice, the MPM is therefore calculated from the joint probability distribution by first calculating the weighted deprivation for each 0/1/missing combination, respecting the weight adjustments when indicators are missing, and then taking the sum of the population share over combinations with a weighted deprivation of at least 1/3, divided by the sum of the population share for which the MPM is defined.

Validation scenarios

Validation of multidimensional poverty measures begins with core infrastructure indicators—access to electricity, access to drinking water, and access to sanitation—before extending to the complete Multidimensional Poverty Measure (MPM) framework. These infrastructure variables are typically collected together in standardized household surveys such as the Demographic and Health Surveys (DHS), allowing early consistency checks and cross-source comparisons. By contrast, income, consumption, education, and other welfare dimensions are often drawn from Household Budget Surveys (HBS) or Living Standards Measurement Study (LSMS) instruments, which differ in sampling design and timing. Prioritizing infrastructure validation therefore leverages the shared data environment of these indicators, establishes confidence in foundational inputs, and provides a stable basis for subsequent verification of all MPM dimensions.

We define several validation scenarios. These scenarios simulate cases where MPM indicators were not collected by a single survey so we cannot observe the full joint distribution of deprivations, but specific indicators are available from different sources. To simulate these scenarios, we split the true population group statistics describing the complete joint distribution (each 0/1 combination) into datasets each containing a subset of the deprivation indicators. To refer to these scenarios, we combine the reference for each indicator from table X (p, c, r, e, w and s) with a hyphen to denote a separate dataset. For example, scenario “ew-ws” refers to the situation where access to electricity and improved drinking water indicators are available from one dataset (and their joint distribution is known), with access to improved drinking water and improved sanitation available from a secondary dataset.

Table 3 indicates the scenarios considered in the validation. In the simpler case considering only three infrastructure deprivations, we consider seven “block-missingness” patterns commonly observed in the GMD database. When considering all indicators used by the MPM, we also restrict the validation to seven scenarios. In the first two scenarios, each indicator and each dimension comes from a separate dataset. The remaining five scenarios represent the current state of data availability given indicators that are more frequently missing in GMD and those typically available from alternate datasets, such as Demographic Health Surveys. Together, they account for patterns of missing MPM indicators observed in 54% of GMD surveys.

Table 3 – Validation scenarios

Infrastructure scenarios	MPM scenarios	
e-w-s	p-c-r-e-w-s	Each deprivation is independently measured
e-ws	p-cr-ews	Each dimension is independently measured
es-w	pcr-ews	Missing pattern in 3% of GMD surveys
ew-s	pcre-ws	Missing pattern in 6% of GMD surveys
es-ws	pce-rews	Missing pattern in 4% of GMD surveys
ew-es	pcew-rews	Missing pattern in 9% of GMD surveys
ew-ws	pcews-rews	Missing pattern in 32% of GMD surveys

For each scenario, we fuse the respective datasets describing the joint distribution for a subset of MPM indicators using the algorithm described in the method. The result is a predicted probability distribution over all indicators with the same format as the true probability distribution we started with. The target multidimensional metrics (any, MPM, etc.) are then calculated in the same way as those from the true probability distribution (Table 2). We aggregate predictions $\hat{y}_{c,s,g}$ for each population group g based on survey s in country c to a national level target variable $\hat{y}_{c,s}$, using subgroup population weights so results are comparable across countries and scenarios:

$$\hat{y}_{c,s} = \sum_g w_{c,s,g} \hat{y}_{c,s,g}, \quad \sum_{s,g} w_{c,s,g} = 1$$

where w_{cg} are subgroup population shares for population group g (from the same GMD survey).

Validation metrics

We calculate several validation metrics that summarize the differences between true and predicted target variables in each scenario. Because survey frequency varies widely across countries, regions, and income groups, we calculate these metrics by averaging across countries rather than surveys. Specifically, for country c with n_c eligible surveys, we assign weight $1/n_c$ to each of its surveys so that each country is weighted equally in the validation results.

To evaluate the performance of the method in various scenarios, we primarily use the mean absolute error (MAE) between predicted and true multidimensional target variables.

$$MAE = \frac{1}{C} \sum_{c=1}^C \frac{1}{n_c} \sum_{s=1}^{n_c} |\hat{y}_{c,s} - y_{c,s}|$$

where $c = 1, \dots, C$ countries; $y_{c,s}$ is the true value of the target variable for survey s and $\hat{y}_{c,s}$ is the predicted value.

The MAE is used rather than the (root) mean squared error because the latter gives more weight to larger errors. We prefer to evaluate the method where the objective is to minimize the difference between the true and predicted values while giving equal weight to all errors. Percentage point errors are of primary interest rather than percentage error since the latter can be very large for countries

with low multidimensional poverty rates. If a country has a true poverty rate of 1 percent and the fusion method predicts a multidimensional poverty rate of 2 percent, the error in percentage terms would be 100 percent which would give this observation a large impact. The focus on percentage point errors gives a larger emphasis to countries with higher multidimensional poverty rates. However, we still report the mean absolute percentage error (MAPE) and root mean squared error (RMSE) across scenarios for comparison:

$$MAPE = \frac{1}{C} \sum_{c=1}^C \frac{1}{n_c} \sum_{s=1}^{n_c} |(\hat{y}_{c,s} - y_{c,s}) / (y_{c,s} + \varepsilon)| * 100$$

where ε is a small positive constant, so that the MAPE is defined for $y_{c,s} = 0$.

$$RMSE = \sqrt{\frac{1}{C} \sum_{c=1}^C \frac{1}{n_c} \sum_{s=1}^{n_c} (\hat{y}_{c,s} - y_{c,s})^2}$$

We also consider the association between predicted and true values by calculating Pearson and Spearman correlation coefficients in each scenario. These are useful to assess the correlation and rank association between predicted and true multidimensional metrics in each scenario, but should not be interpreted as a measure of bias. Surveys are weighted equally for these correlation metrics. The Pearson correlation coefficient is defined as:

$$r = \frac{[\sum_i (y_i - \bar{y})(\hat{y}_i - \hat{\bar{y}})]}{\sqrt{[\sum_i (y_i - \bar{y})^2][\sum_i (\hat{y}_i - \hat{\bar{y}})^2]}} = \frac{\text{cov}(y, \hat{y})}{\sigma_y \sigma_{\hat{y}}}$$

where \bar{y} is the mean of the true metric y , and $\hat{\bar{y}}$ is the mean of the predicted metric \hat{y} .

Spearman's rank correlation coefficient is the Pearson correlation of the rank-transformed target variables:

$$\rho = \frac{[\sum_i (R_i - \bar{R})(\hat{R}_i - \hat{\bar{R}})]}{[\sum_i (R_i - \bar{R})^2][\sum_i (\hat{R}_i - \hat{\bar{R}})^2]} = \frac{\text{cov}(R, \hat{R})}{\sigma_R \sigma_{\hat{R}}}$$

where R_i and \hat{R}_i are the ranks of y_i and \hat{y}_i (using average ranks for ties).

Simulating sample bias

Finally, we also test how sensitive our method is to sampling error by simulating bias in one of the validation data sets before fusion. This resembles real world settings of margin incompatibility in which the population share deprived on the same indicator estimated from two different surveys varies due to differences in the underlying sample (or sampling weights). To simulate bias, we adjust the population share deprived (on any indicator) in the secondary dataset by a factor ranging between 0.8 and 1.2. The share of the population not deprived is adjusted in the opposite direction so that the total population share (joint probabilities) adds to one. Surveys are excluded from the sample bias

validation scenarios if the adjusted population shares are below zero or above one. In these cases, we cannot update population shares by the full extent of the simulated bias in a valid way. This augments the number of scenarios in Table 3, as we simulate 10 levels of sample bias for each one ($\pm 2.5\%$, $\pm 5\%$, $\pm 10\%$, $\pm 15\%$, $\pm 20\%$).

Figure 10 summarizes the validation steps.

Figure 10 – Summary of validation steps

1. True joint distribution (by population group)

Country	Survey	Group	p	c	r	e	w	s	$P(p,c,r,e,w,s)$
AAA	2020	Urban	0	0	0	0	0	0	0.05
...	1	0	0	0	0	0	0.06
...	1	1	0	0	0	0	0.07
...

2. True multidimensional poverty (national level)

Country	Survey	MPM	Any	...
AAA	2020	0.3	0.7	...
AAA	2015	0.4	0.9	...
BBB	2017	0.2	0.4	...
...

3. Split true distribution into validation datasets for each scenario (e.g. pcr-e-w-s)

Country	Survey	Group	p	c	r	e	$P(p,c,r,e)$	Country	Survey	Group	e	w	s	$P(e,w,s)$
AAA	2020	Urban	0	0	0	0	0.08	AAA	2020	Urban	0	0	0	0.09
...

6. Compare and compute validation metrics

Scenario	Group ^a	Target	MAE	RMSE	...
pcr-e-w-s	Urban	MPM	0.02	0.03	...
...	...	Any	0.01	0.05	...
...

4. Apply fusion algorithm to predict joint distribution for each scenario (by population group)

Scenario	Country	Survey	Group	p	c	r	e	w	s	$P^{\wedge}(p,c,r,e,w,s)$
pcr-e-w-s	AAA	2020	Urban	0	0	0	0	0	0	0.06
...	1	0	0	0	0	0	0.07
...	1	1	0	0	0	0	0.08
...

5. Predicted multidimensional poverty (national level)

Scenario	Country	Survey	Group	MPM [^]	Any [^]	...
pcr-e-w-s	AAA	2020	Urban	0.33	0.7	...
...	AAA	2020	Rural	0.44	0.9	...
...

Source: Authors' illustration; calculations by the authors.

4. Results

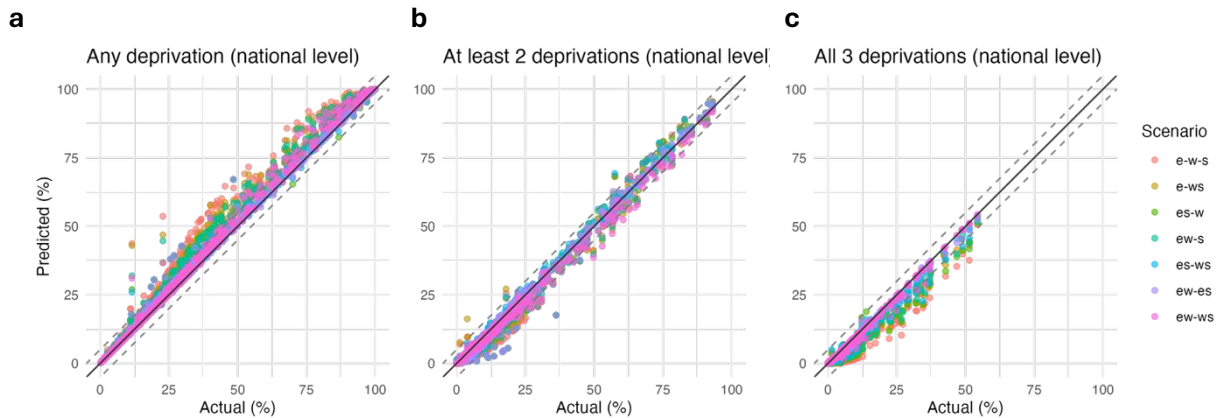
We first validated the core infrastructure indicators—household access to electricity, improved drinking water, and sanitation—using Global Monitoring Database (GMD) microdata. After establishing this infrastructure baseline, we extended the analysis to the full multidimensional poverty measures comprising all six indicators. Although all variables originate from the same household surveys, we simulated conditions in which indicators appear to come from separate data sources and are available only at different population-group levels. Starting from the complete joint distribution of binary deprivations, the survey data were partitioned into subsets, each containing a restricted set of indicators. Each validation scenario was labeled by concatenating the indicator codes from Table 3 (p, c, r, e, w, s) with hyphens to denote the distinct pseudo-datasets. For example, scenario “ew-ws” represents a setting where electricity (e) and improved drinking water (w) are treated as jointly observed in one pseudo-dataset, while improved drinking water (w) and sanitation (s) are treated as jointly observed in another.

For each indicator and scenario, we regenerated deprivation and multidimensional estimates from these pseudo population-group datasets and compared them with the “true” estimates derived from the full survey. Validation involved (i) *visual assessment*, plotting observed versus predicted deprivations in national-level scatter plots and repeating the exercise after aggregating subgroup predictions (urban/rural, income quintile, and subnational units) to examine consistency across population structures; and (ii) *tabular evaluation*, computing mean absolute error (MAE), mean absolute percentage error (MAPE), root mean squared error (RMSE), and Pearson (r) and Spearman (ρ) correlation coefficients between simulated and true estimates.

Multidimensional fusion with three infrastructure indicators

At every disaggregation we examined—national, wealth quintile, urban/rural, and subnational ADM1—the predicted deprivation rates in access to three types of basic infrastructure closely track the survey “ground truth”, as shown in Figure 11. Bias in predictions of the population deprived in at least one infrastructure (any) due to assumptions of conditional independence is almost always in the positive direction (panel a), while bias tends to be in the negative direction for predictions of the population deprived in all three indicators. Results for the “at least 2 deprivations” target variable (panel b) indicate a middle ground, with no clear direction in the bias. Validation scenarios with at least one indicator in common across datasets appear to accurately estimate the true value of each national level target indicators with very few outliers.

Figure 11 – Actual vs predicted joint deprivations in access to infrastructure

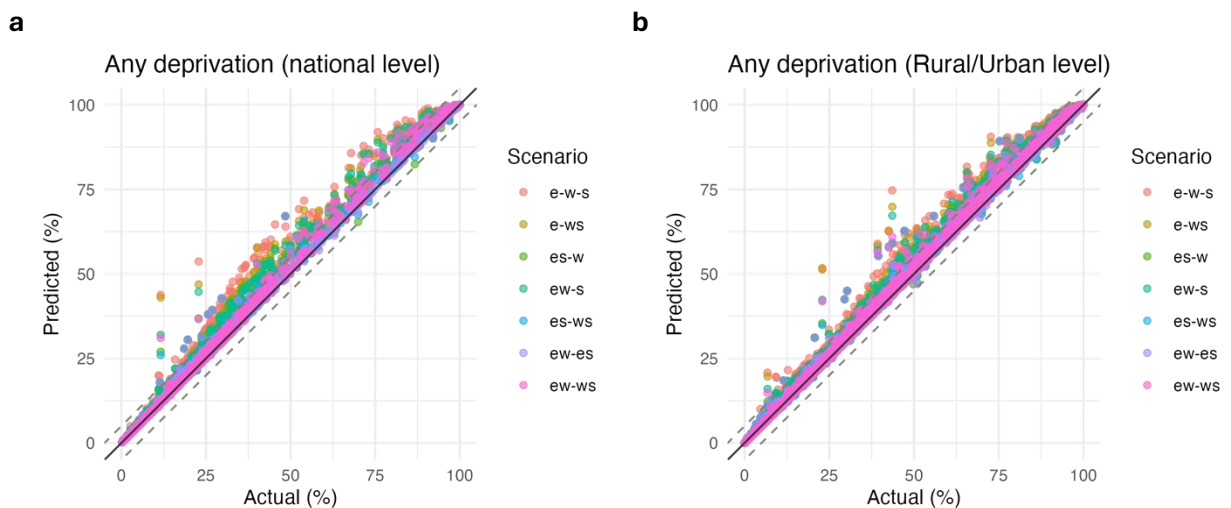


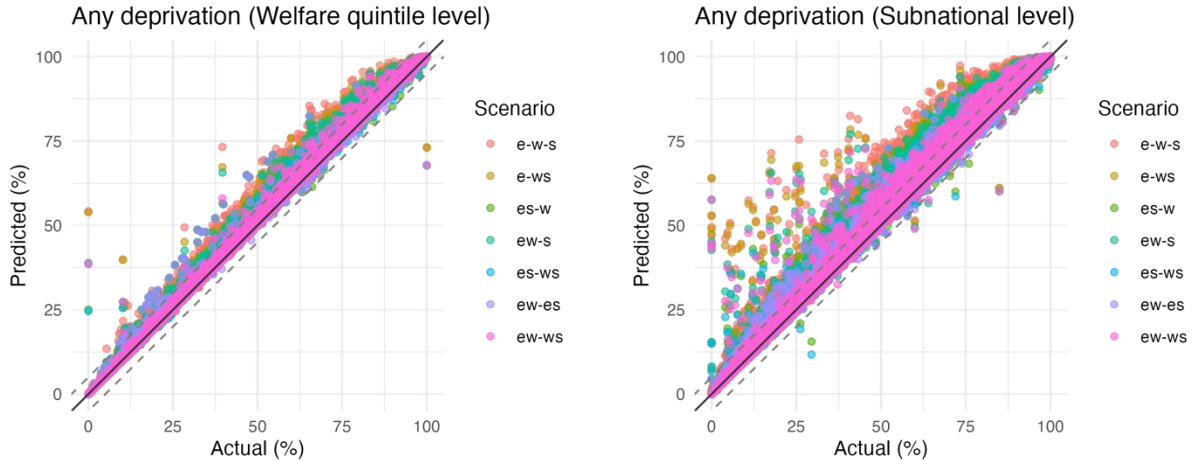
Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Notes: The figures show actual vs predicted values of three target metrics calculated at national level. Fusion is performed using national level population statistics. No sample bias is simulated. Dashed lines denote ± 5 percentage-point tolerance bands around the 45-degree reference line.

Figure 12 shows actual and predicted population shares without access to at least one type of infrastructure aggregated at different levels of the population. Points from all groupings cluster near the 45-degree line, indicating strong agreement overall. The range of errors is systematically smaller for national, urban/rural, and quintile data levels, compared to at the subnational level. Points representing subnational ADM1 areas display slightly wider scatter, suggesting greater variability when predictions are fused at finer subnational scales. As above, we find that the population share deprived on any indicator tends to be overestimated in scenarios with stronger assumptions of independence, regardless of the data level and similar patterns in the bias across the range of true values. The largest errors (in the worst performing scenario) occur when the true population share is approximately 50 percent.

Figure 12 – Actual vs predicted population deprived in any infrastructure, by population level





Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Notes: The figures show actual vs predicted values of the “any infrastructure deprivation” target variable calculated at four different population group levels. Fusion is performed using population statistics at this level. No sample bias simulated. Dashed lines denote ± 5 percentage-point tolerance bands around the 45-degree reference line.

Table 4 shows the mean of each target variable. National “true” estimates from the full survey underscore the substantial burden of infrastructure deprivation. Using 573 surveys covering 112 countries from 1989 to 2024 and weighting countries equally, over 45 percent of households lack access to at least one basic service—electricity, improved drinking water, or sanitation—while 24 percent lack any two of these services and nearly 8 percent are deprived in all three indicators. The magnitude varies markedly across population groups: in some subnational regions, deprivation rates exceed 50 percent. These baseline figures highlight the scale of the challenge and provide the benchmark against which all simulated scenarios are evaluated.

Table 5 shows validation metrics across different missingness scenarios when aggregating predictions at different levels of the population to the national “any infrastructure deprivation” target variable (using population weights). The main pattern is monotone with information overlap: scenarios that share at least one indicator across sources (e.g., es-ws, ew-es, ew-ws) yield the lowest MAEs, mixed-overlap cases (e-ws, es-w, ew-s) perform in the middle, and the fully split case e-w-s (three distinct sources) has the highest MAE. This ranking holds across the three target variables, with the “all 3 indicators deprived” predictions typically having the largest MAE (Table A2) and the “ ≥ 1 indicator deprived” prediction the most accurate on average (smallest MAE). Importantly, the gains from cross-source fusion attenuate with finer granularity: improvements (lower MAE) are largest at the urban/rural level, followed by quintile, then subnational (ADM1), and are smallest at the national level. This ordering is consistent with increasing sampling noise and measurement heterogeneity as groups become larger.

Between “true” and predicted deprivation, rank association (ρ) are high (more than 0.99) across data level and most data scenarios. Across the seven cross-source missingness scenarios and four fusion levels, the algorithm consistently reconstructs deprivation patterns with low error and high concordance at coarse aggregation, and with gradually weaker performance as granularity

decreases. Scenarios with overlapping indicators across sources outperform fully split sources, and the “all 3 indicators deprived” target remains the most difficult, while “deprived in at least one of the three indicators” is the most stable. Aggregating subgroup predictions with population weights yields small bias of national estimates in our experiments, supporting the use of cross-source fusion for practical monitoring when single-source coverage is incomplete.

Table 4 – True infrastructure deprivation (%)

	Deprivation in any indicator	Deprivation in any two indicators	Deprivation in all three indicators
National	45.61	24.18	7.86

Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Note: The table shows the true deprivation metrics of every survey data for a country, then average over all countries. Aggregated numbers weigh each country by the inverse of its number of surveys, such that the total weight of each country equals one. Sample: 573 surveys for 112 countries from 1989 to 2024.

Table 5 – Validation of infrastructure deprivation (any) metrics by scenario and population level of fused data

	MAE	MAPE	RMSE	r	ρ	N
e-w-s						
National	4.909	13.810	6.913	0.990	0.994	573
Quintile	3.606	11.404	5.478	0.994	0.994	573
RuralUrban	3.150	10.452	4.786	0.995	0.995	573
Subnational	3.693	11.469	5.455	0.993	0.994	566
e-ws						
National	3.584	8.557	5.466	0.994	0.996	573
Quintile	2.623	7.096	4.540	0.995	0.996	573
RuralUrban	2.365	6.647	4.095	0.996	0.996	573
Subnational	2.735	7.139	4.526	0.995	0.996	566
es-w						
National	2.438	7.816	3.514	0.996	0.997	573
Quintile	1.942	7.092	3.440	0.996	0.996	573
RuralUrban	1.548	6.115	2.610	0.998	0.997	573
Subnational	1.837	6.665	2.909	0.997	0.997	566
es-ws						
National	0.805	2.066	1.507	0.999	0.999	573
Quintile	0.807	2.496	2.436	0.998	0.997	573
RuralUrban	0.659	2.083	1.739	0.999	0.998	573
Subnational	0.684	1.980	1.626	0.999	0.998	566
ew-es						
National	0.900	3.622	1.786	0.998	0.998	573
Quintile	0.923	4.000	2.218	0.998	0.997	573
RuralUrban	0.769	3.417	1.585	0.999	0.998	573
Subnational	0.806	3.452	1.609	0.999	0.998	566
ew-s						
National	3.759	10.228	5.263	0.995	0.996	573

Quintile	2.783	8.653	4.321	0.996	0.995	573
RuralUrban	2.519	8.050	3.871	0.997	0.996	573
Subnational	2.899	8.684	4.292	0.996	0.996	566
ew-ws						
National	2.298	4.727	3.686	0.998	0.998	573
Quintile	1.713	4.164	3.341	0.997	0.997	573
RuralUrban	1.679	4.112	3.140	0.998	0.997	573
Subnational	1.835	4.136	3.253	0.998	0.997	566

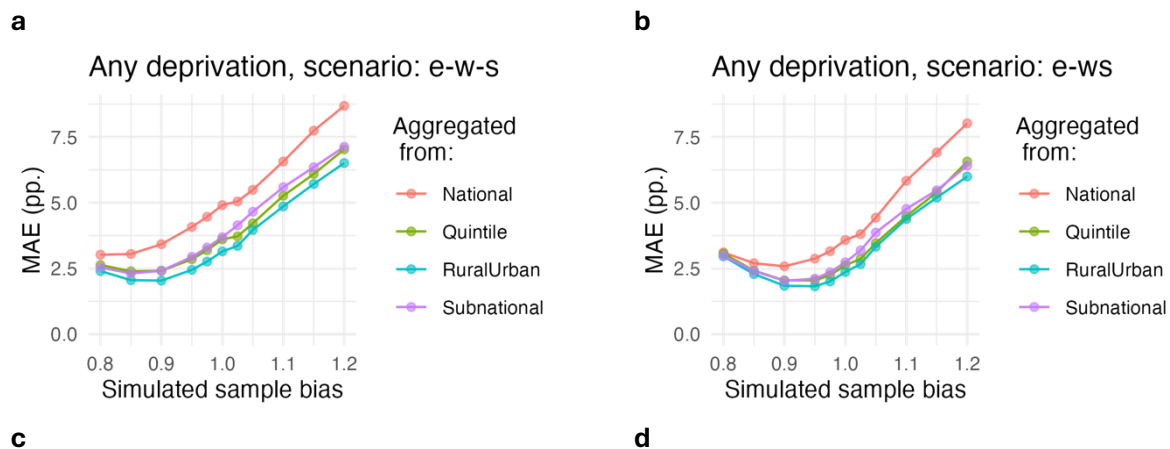
Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

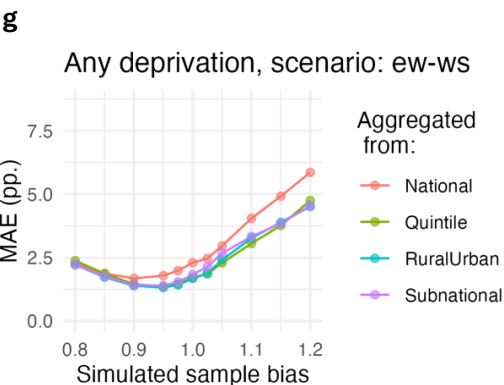
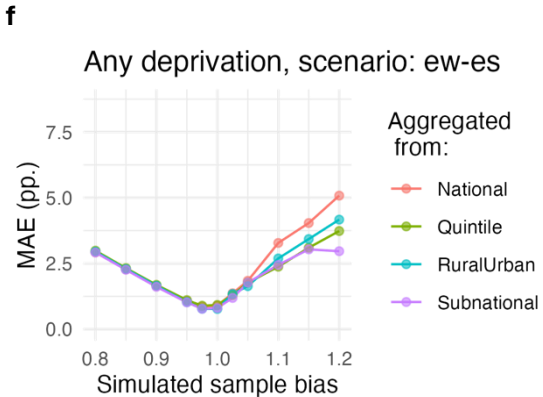
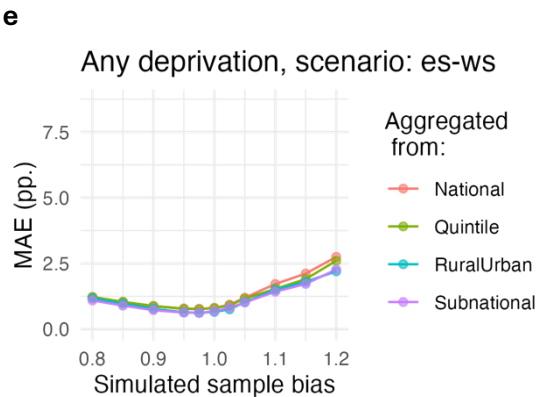
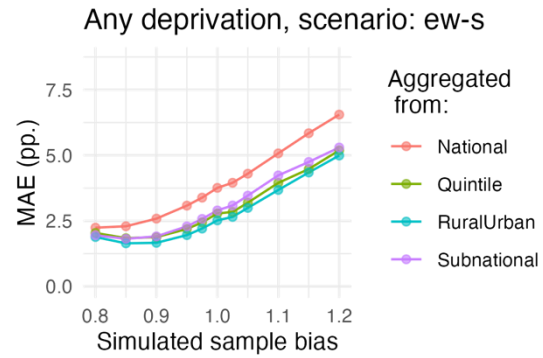
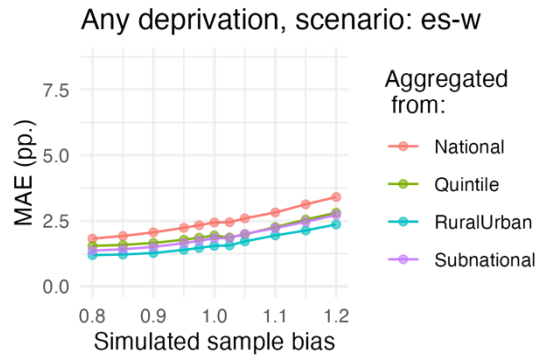
Note: The table shows the mean absolute error at the national level aggregated from different fusion levels for all countries. In each scenario and each fusion level, mean absolute error (MAE) is calculated for every survey data for a country, then average over all countries using weight as the inverse of its number of surveys.

To mimic real-world conditions in which indicator estimates may come from different data sources with distinct sampling frames or weighting schemes, we calculated MAE under a range of imposed upward and downward biases in the pseudo-data. This resembles real-world settings of margin incompatibility, where the population share deprived on the same indicator—when estimated from two independent surveys—can differ because of both data-source heterogeneity and sampling differences.

Figure 13 illustrates this relationship, plotting MAE against 10 levels of data bias ($\pm 2.5\%$, $\pm 5\%$, $\pm 10\%$, $\pm 15\%$, $\pm 20\%$) across all scenarios and indicator sets. The validation results reveal a clear and consistent pattern in mean absolute error (MAE) as the pseudo-data mean or joint probability depart from the true population mean or joint probability. MAE is lowest when the simulated samples reproduce the actual mean, but it rises progressively as the assumed means deviate because of these cross-source or sampling discrepancies. The increase is asymmetric: upward-biased pseudo-means generate larger MAE than downward-biased ones of similar magnitude, indicating greater sensitivity to overestimation of deprivation. This pattern is consistent across every validation configuration and subgroup populations.

Figure 13 – MAE vs simulated sample bias for any infrastructure deprivation across scenarios





Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

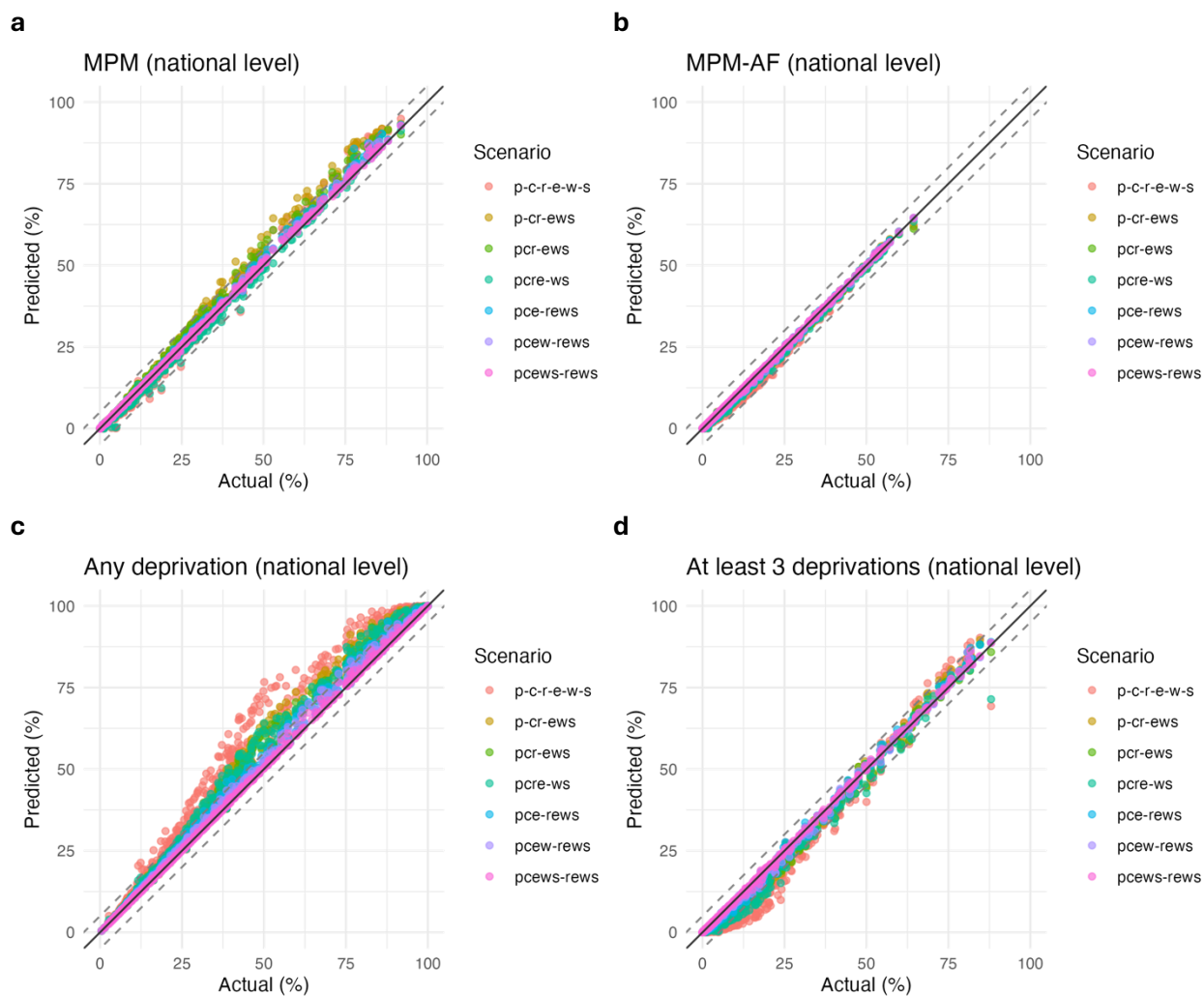
Notes: The figures report the mean absolute error (MAE) in any infrastructure deprivation under simulated sample-bias factors ranging from 0.8 to 1.2 across all validation scenarios. Population-group estimates are aggregated to the national level to enable cross-group comparison.

Multidimensional fusion with 6 MPM indicators

Figure 14 shows actual and predicted multidimensional poverty metrics for four target variables across the seven validation scenarios that consider all six indicators used to construct the MPM from national level population statistics. While we generally observe very strong agreement, the results reveal some interesting patterns. Firstly, results vary significantly depending on the target multidimensional variable of interest. The error between the actual and predicted population share

deprived on any indicator is larger than for other metrics on average (panel c) and in scenarios where less is known about the joint distribution, the fusion method tends to significantly overestimate this metric. The MPM also tends to be overestimated, but to a much smaller extent for the same data scenario (panel a). The Alkire-Foster method of calculating multidimensional poverty appears remarkably insensitive to assumptions of independence in the fusion method (panel b), with predicted values for almost every country and data scenario falling within five percentage points of the true value. There is some indication the MPM-AF tends to be underestimated for lower true values of this metric. The share of the population deprived on at least three deprivations is often underestimated in the more data constrained scenarios, particularly for lower true values (panel d). Overall, the fusion scenarios with at least one indicator in common across datasets (pce-rews, pcew-rews, pcews-rews) accurately estimate the true value of all four national level target indicators with very few outliers.

Figure 14 – Actual vs predicted multidimensional poverty by target variable

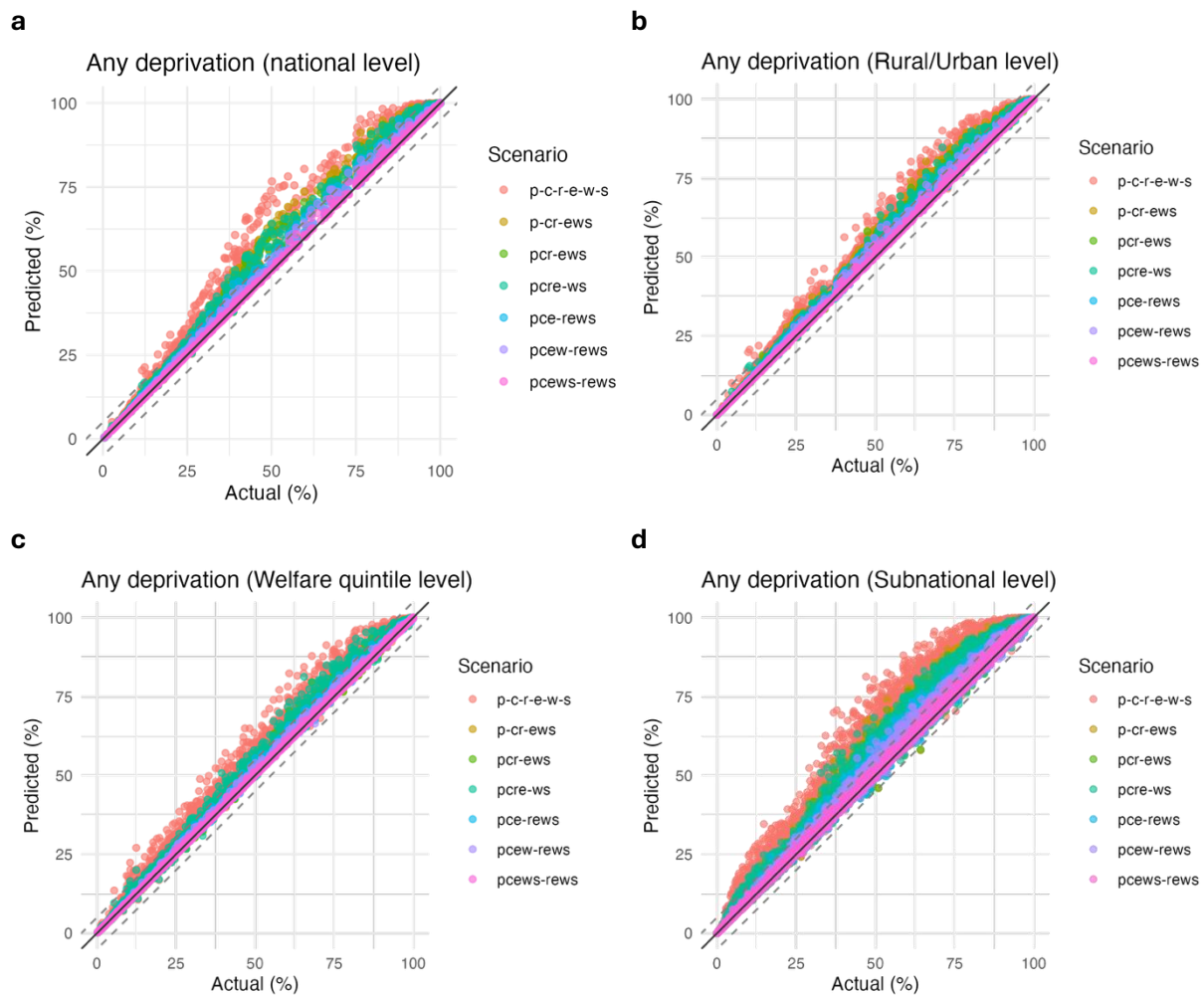


Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Notes: The figures show actual vs predicted values of four target multidimensional poverty metrics calculated at national level. Fusion is performed using national level population statistics. No sample bias is simulated. The dashed lines show $\pm 5\%$ tolerance bands.

Figure 15 shows the actual and predicted share of population deprived in any MPM indicator calculated at different levels of the population by fusing statistics at the same level. As above, we find that the population share deprived on any indicator tends to be overestimated in scenarios with stronger assumptions of independence, regardless of the data level. There appears to be slightly fewer outliers when predicting this metric at rural/urban or welfare quintile level, compared to national level (panel b and c). There are many more observations available for subnational level statistics, and higher errors in prediction using the fusion method in some scenarios, although the pattern of bias is very similar. The largest errors (in the worst performing scenario) occur when the true population share is approximately 60 percent.

Figure 15 – Actual vs predicted population deprived in any MPM indicator, by population level



Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Notes: The figures show actual vs predicted values of the “any deprivation” target variable calculated at four different population group levels. Fusion is performed using population statistics at this level. No sample bias simulated. The dashed lines show $\pm 5\%$ tolerance bands.

The panels shown above in Figure 15 should not be compared on a like-for-like basis, since they estimate multidimensional poverty at different levels of the population. To assess how much difference it makes to fuse data at different population group levels for national statistics, we aggregate these predictions to the same national level target variable. Table 6 shows the mean of true target variables and Table 7 shows validation metrics across target variables, scenarios and aggregated from different population levels.

Table 6 – True multidimensional metrics (%)

	MPM (weighted headcount)	AF method (headcount * intensity)	Deprived in at least 1 (any)	Deprived in all 6 indicators (all)
National	28.67	16.06	52.77	0.59

Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Note: The table shows the true multidimensional metrics of every survey data for a country, then average over all countries. Aggregated numbers weigh each country by the inverse of its number of surveys, such that the total weight of each country equals one. Sample: 573 surveys for 112 countries from 1989 to 2024.

Table 7 – Validation of MPM (weighted headcount) metrics by scenario and population level of fused data

	MAE	MAPE	RMSE	r	ρ	N
p-c-r-e-w-s						
National	1.696	11.228	2.701	0.998	0.994	573
Quintile	1.365	10.310	2.013	0.999	0.995	573
RuralUrban	1.009	8.959	1.594	0.999	0.996	573
Subnational	1.219	9.074	1.912	0.999	0.995	566
p-cr-ews						
National	2.260	6.380	3.598	0.999	1.000	573
Quintile	0.282	2.312	0.477	1.000	1.000	573
RuralUrban	1.298	4.430	2.027	0.999	1.000	573
Subnational	1.558	4.614	2.421	0.999	1.000	566
pce-rews						
National	0.864	3.602	1.453	1.000	1.000	573
Quintile	0.199	2.174	0.414	1.000	1.000	573
RuralUrban	0.661	3.219	1.071	1.000	1.000	573
Subnational	0.710	3.090	1.175	1.000	1.000	566
pcew-rews						
National	0.555	2.617	0.995	1.000	1.000	573
Quintile	0.166	1.769	0.361	1.000	1.000	573
RuralUrban	0.485	2.579	0.828	1.000	1.000	573
Subnational	0.467	2.316	0.815	1.000	1.000	566
pcews-rews						

National	0.379	2.119	0.690	1.000	1.000	573
Quintile	0.125	1.528	0.307	1.000	1.000	573
RuralUrban	0.352	2.079	0.625	1.000	1.000	573
Subnational	0.332	1.935	0.586	1.000	1.000	566
pcr-ews						
National	1.369	4.566	2.299	0.999	1.000	573
Quintile	0.298	2.349	0.496	1.000	1.000	573
RuralUrban	0.729	3.183	1.187	1.000	1.000	573
Subnational	0.931	3.313	1.529	1.000	1.000	566
pcre-ws						
National	0.862	7.837	1.413	0.999	0.996	573
Quintile	0.995	7.492	1.528	0.999	0.996	573
RuralUrban	0.550	6.435	0.920	1.000	0.997	573
Subnational	0.649	6.154	1.040	0.999	0.997	566

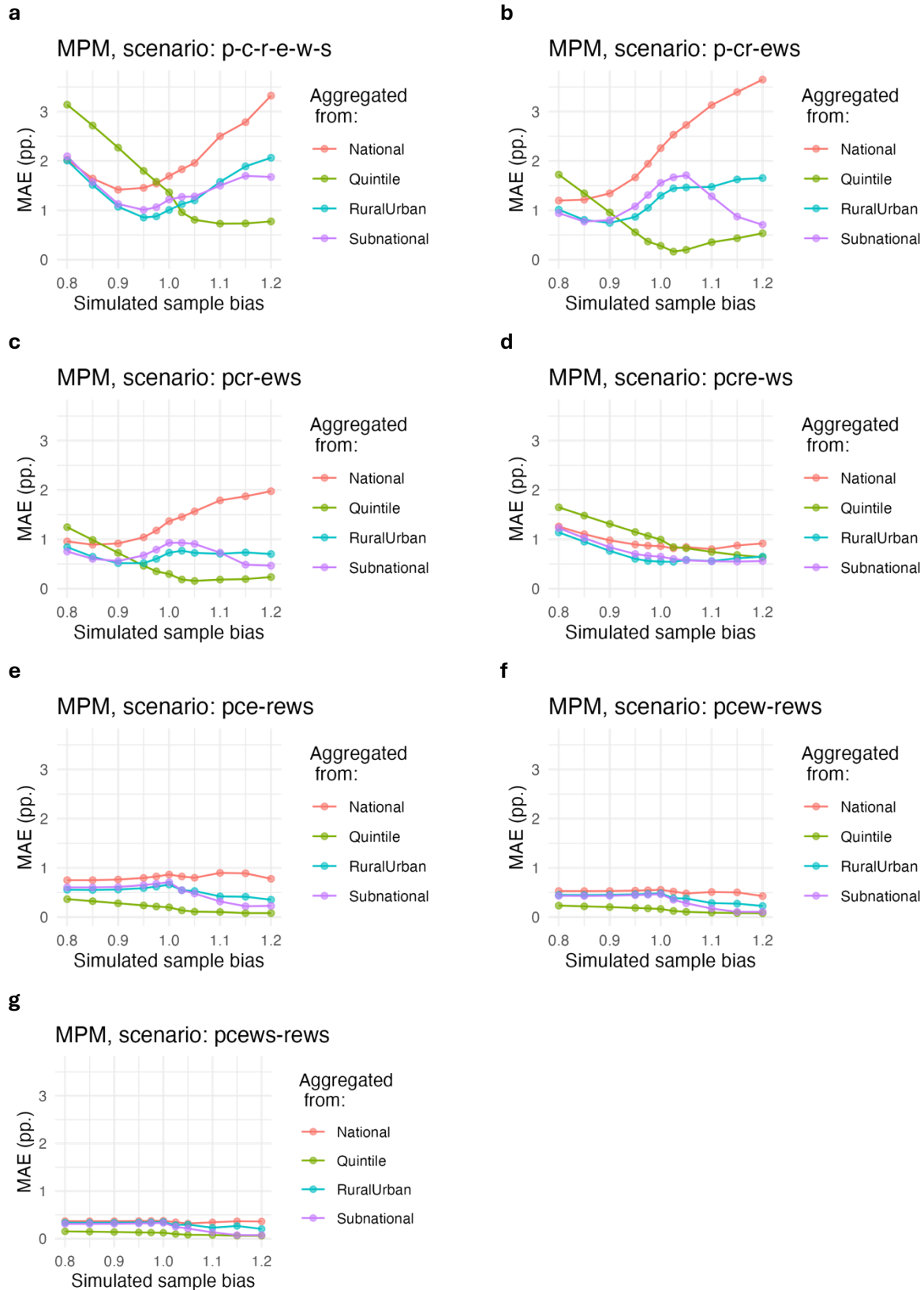
Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Note: The table shows different validation metrics of every survey data for a country, then average over all countries. Aggregated numbers weigh each country by the inverse of its number of surveys, such that the total weight of each country equals one. Results from different population groups are aggregated to the national target level.

Figure 16 shows the MAE when predicting the MPM in validation scenarios where sampling bias has been simulated in the fused data, as described in the method. This resembles real world settings of margin incompatibility in which the population share deprived on the same indicator estimated from two different surveys varies due to differences in the underlying sample. Interestingly, we find the average error in predictions to be smaller in some cases of sample bias in the negative direction compared to no adjustment (sample bias = 1). This is because the lower estimates of the deprived population partly offset the effect of assuming (conditional) independence between indicators, which tends to overestimate the MPM in more constrained scenarios. Aggregating national level target variables from Rural/Urban or subnational level fusion predictions consistently performs better across the range of simulated sample bias compared to starting with national level statistics. Fusing quintile level statistics performs best in some scenarios, but mostly in cases of positive sampling bias – when secondary datasets estimate higher deprivation rates. In scenarios with at least one common indicator across datasets, the MAE is less than one percentage point for all simulated bias scenarios and there is little to split the results aggregating to the national level from different population subgroups. This should provide confidence to apply the multidimensional fusion method to missing data scenarios that are commonly observed in the GMD, and using a secondary dataset such as a Demographic Health Survey drawn from a different representative sample.

Figure 17 shows analogous results when predicting the share of the population deprived in any MPM indicator across sampling bias scenarios. This target variable – the union of all deprivations - is more sensitive to sample bias in the fused indicators. Across most scenarios and data levels, the MAE decreases when there is less sample bias (closer to one) and errors are larger for positive sample bias compared to the same extent of negative bias in fused deprivations. The scenario in panel g, with three common indicators and therefore weaker assumptions of conditional independence, is significantly less sensitive to sample bias compared to other scenarios.

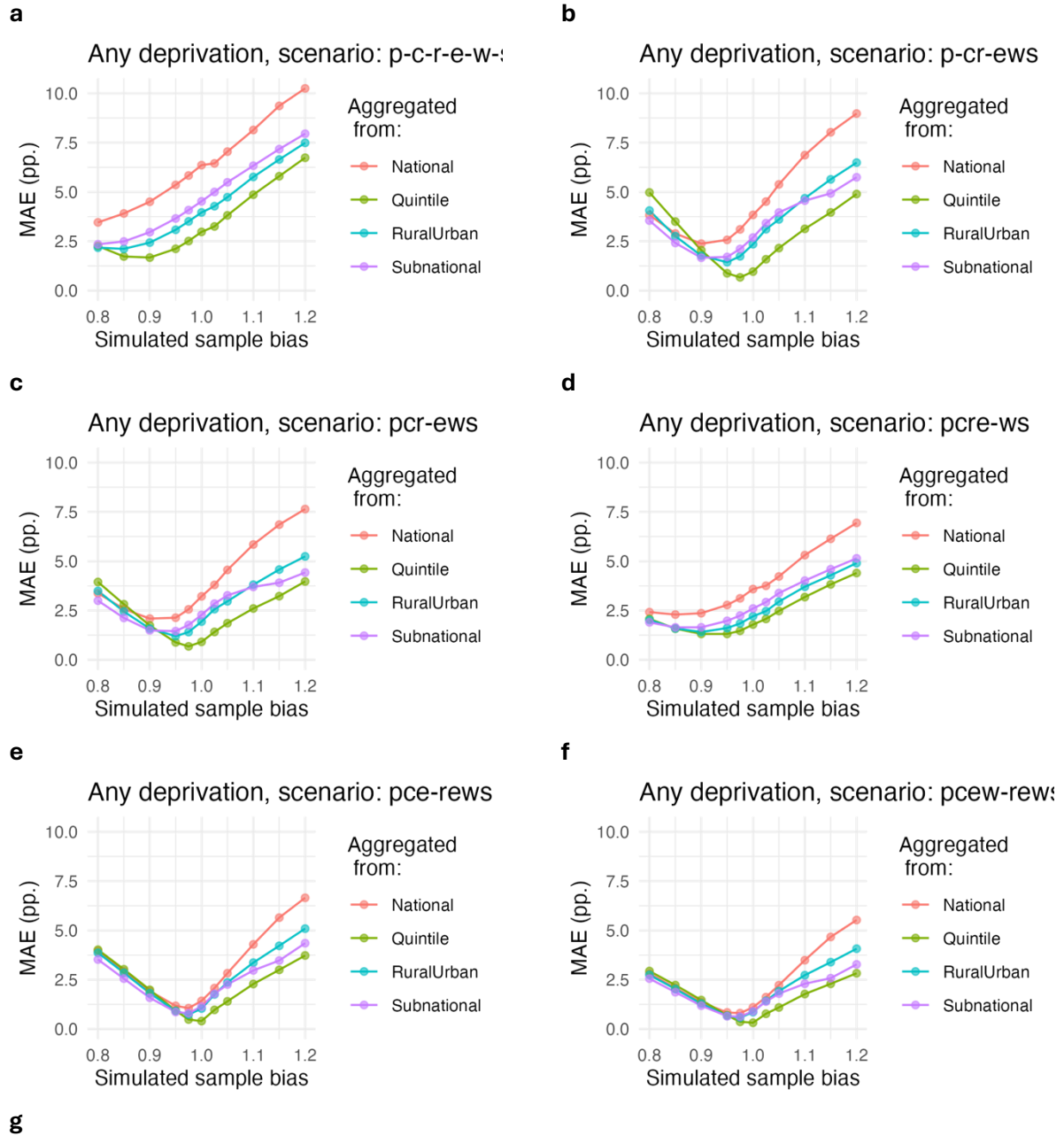
Figure 16 – MAE vs simulated sample bias for MPM across scenarios

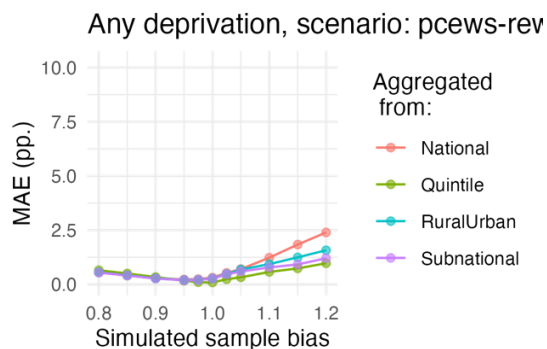


Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Notes: The figures report the mean absolute error (MAE) in any infrastructure deprivation under simulated sample-bias factors ranging from 0.8 to 1.2 across all validation scenarios. Population-group estimates are aggregated to the national level to enable cross-group comparison.

Figure 17 – MAE vs simulated sample bias for any deprivation across scenarios





Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Notes: The figures report the mean absolute error (MAE) in any infrastructure deprivation under simulated sample-bias factors ranging from 0.8 to 1.2 across all validation scenarios. Population-group estimates are aggregated to the national level to enable cross-group comparison.

5. Discussion and Conclusion

This study demonstrates that multidimensional poverty metrics can be estimated with high accuracy even when indicators originate from multiple data sources and are available only as population-group aggregates. Drawing on 573 surveys from 112 countries collected between 1989 and 2024, the proposed probabilistic fusion method reproduces full-survey “true” estimates with strong accuracy. For infrastructure indicators, the method records an average absolute error of roughly 1 percentage point when estimating the share of households—about 45 percent—lacking at least one basic service. For the overall multidimensional poverty measure, it achieves an average absolute error of about 0.5 percentage point in estimating the proportion of households—approximately 26 percent—experiencing multiple deprivations. Across all validation scenarios, mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE) remain low, while Pearson and Spearman correlations consistently exceed 0.9. Predictive accuracy is highest at urban/rural, and wealth-quintile levels and only modestly lower for subnational (ADM1) units, reflecting the greater variability inherent in small-area estimates. Sensitivity analyses further show that MAE is minimal when pseudo-data match the true mean and increases asymmetrically as sampling bias grows, with upward bias producing larger errors than comparable downward bias.

Our approach offers a practical alternative to more data-intensive survey imputation or full-scale simulation methods, making it well suited to contexts where high-quality microdata are scarce or fragmented. Aggregated deprivation rates and partial joint distributions can be drawn from multiple heterogeneous sources—or modeled directly—rather than relying on a single, comprehensive household survey. This flexibility is particularly valuable in low-capacity statistical environments, where indicators are often collected in separate instruments with differing sampling designs.

The simple probabilistic fusion framework enables the line-up of multidimensional poverty measures to a common reference year across countries, allowing analysts to estimate target-year metrics by combining gap-filled aggregate deprivation rates with conditional probabilities for a

subset of deprivations derived from existing microdata. This approach utilizes data from multiple surveys and administrative sources without imposing heavy modeling requirements. Two complementary strategies will guide this effort: an interpolation approach, used when population-level aggregates for some dimensions are available for the reporting year and the full joint distribution is known for both a prior and a subsequent year, and an extrapolation approach, applied when only a previous joint distribution is known and current-year margins must be reconciled with the earlier joint to project forward the dependence structure. Together, these methods will deliver timely, country-comparable multidimensional poverty metrics estimates even when data availability is incomplete or lags behind policy needs.

Despite its advantages, this approach has several limitations. First, it relies on the conditional independence assumption, which presumes that unobserved dependencies among indicators are negligible once observed margins are matched; violations of this assumption could bias reconstructed joint distributions. Second, the method assumes that overlapping indicators are consistently defined and measured across data sources, yet differences in survey design, question wording, or administrative definitions may introduce unobserved measurement error. Third, while probabilistic fusion uses observed margins to rebuild the joint distribution, it cannot fully capture structural changes in deprivation patterns when no overlapping indicators are available. Fourth, accuracy declines somewhat at finer geographic scales—such as ADM1—where small-sample variability is greater. Finally, the interpolation and extrapolation strategies depend on the availability of at least partial historical or future joint distributions; in countries with sparse or irregular survey schedules, these requirements may be difficult to satisfy. These constraints suggest that results should be interpreted with caution and underscore the importance of continued investment in high-quality, harmonized household data.

Methodological Extensions

Future work could enhance the current framework by relaxing the conditional independence (CI) assumptions that underpin the probabilistic fusion approach. Allowing selected pairwise or higher-order dependencies among indicators would better capture structural correlations—such as the common co-occurrence of sanitation and water deprivations—that strict CI may overlook. One avenue is to incorporate historical joint probability distributions from prior country-year data (size 2^k for k indicators) as informative priors. By treating these historical joints as reference distributions, analysts can estimate the current joint via a Kullback–Leibler (KL) divergence minimization, effectively finding the distribution closest to the historical prior while matching current marginal constraints. This would permit the fusion algorithm to borrow strength across time and space, improving robustness when contemporary data are sparse or partially observed.

Handling Multi-Granular Data

Another promising extension is to accommodate heterogeneous subgroup availability, where indicators are reported at different population partitions—for example, indicator X_1 observed only by urban/rural status and indicator X_2 by wealth quintile. This situation requires a reconciliation strategy that enforces consistency across overlapping subgroups while constructing a coherent national joint

distribution. Potential solutions include iterative proportional fitting (IPF) or Bayesian hierarchical models that jointly estimate cell probabilities while respecting all known margins. Such methods would enable the integration of diverse data releases—household surveys, administrative sources, and specialized modules—into a single, internally consistent multidimensional poverty measure without sacrificing the richness of subgroup information.

References

- Alkire, Sabina, and James Foster. "Counting and multidimensional poverty measurement." *Journal of public economics* 95, no. 7-8 (2011): 476-487.
- Alkire, Sabina, José Manuel Roche, Paola Ballon, James Foster, Maria Emma Santos, and Suman Seth. *Multidimensional poverty measurement and analysis*. Oxford University Press, USA, 2015.
- Mahler, Daniel Gerszon, R. Andrés Castañeda Aguilar, and David Newhouse. "Nowcasting global poverty." *The World Bank Economic Review* 36, no. 4 (2022): 835-856.
- Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. "Distributional national accounts: methods and estimates for the United States." *The Quarterly Journal of Economics* 133, no. 2 (2018): 553-609.
- Roy, Sutirtha Sinha, and Roy Van Der Weide. "Estimating poverty for India after 2011 using private-sector survey data." *Journal of Development Economics* 172 (2025): 103386.
- Sen, Amartya. "Capability and well-being." *The quality of life* 30, no. 1 (1993): 270-293.
- Stojkoski, Viktor, Luis F. Lopez-Calva, Kimberly Bolch, and Almudena Fernandez. "'Development acupuncture:' mapping the network structure of multidimensional poverty." *Journal of Economic Policy Reform* (2025): 1-20.
- Ummel, Kevin, Miguel Poblete-Cazenave, Karthik Akkiraju, Nick Graetz, Hero Ashman, Cora Kingdon, Steven Herrera Tenorio, Aaryaman Sunny Singhal, Daniel Aldana Cohen, and Narasimha D. Rao. "Multidimensional well-being of US households at a fine spatial scale using fused household surveys." *Scientific Data* 11, no. 1 (2024): 142.
- World Bank, 2024a. Global Monitoring Database (GMD) (version September 2024)
- World Bank, 2024b. Poverty and Inequality Platform (PIP) (data set). World Bank, Washington, DC. <https://pip.worldbank.org/home>.

Supplementary Information

Implementation of the fusion algorithm

The probabilistic fusion algorithm is implemented in R as a function which takes a list of dataframes containing probability distributions and sequentially fuses them using the chain rule. The core fusion logic is handled by a helper function *fuse_two_dfs*, which: (1) detects common variables for merging, (2) calculates conditional probabilities, (3) merges dataframes on common variables, and (4) multiplies marginal and conditional probabilities to compute joint probabilities.

Additional results for the infrastructure deprivation and multidimensional poverty metrics are available below.

Table A1 – Validation of infrastructure deprivation (at least two) metrics by scenario and population level of fused data

	MAE	MAPE	RMSE	r	ρ	N
e-w-s						
National	2.593	48.863	3.586	0.990	0.978	573
Quintile	1.963	48.580	2.967	0.994	0.982	573
RuralUrban	1.729	45.752	2.609	0.996	0.984	573
Subnational	1.961	47.353	2.816	0.994	0.986	566
e-ws						
National	1.821	90.311	2.904	0.995	0.987	573
Quintile	1.370	89.683	2.492	0.997	0.989	573
RuralUrban	1.331	88.792	2.330	0.997	0.990	573
Subnational	1.401	91.328	2.382	0.997	0.989	566
es-w						
National	1.507	34.208	2.446	0.995	0.984	573
Quintile	1.331	37.127	2.348	0.996	0.985	573
RuralUrban	1.099	33.547	1.906	0.997	0.989	573
Subnational	1.200	34.601	2.003	0.997	0.990	566
es-ws						
National	0.960	74.904	1.746	0.998	0.994	573
Quintile	0.867	76.211	1.855	0.998	0.993	573
RuralUrban	0.652	73.118	1.430	0.999	0.995	573
Subnational	0.701	76.123	1.376	0.999	0.994	566
ew-es						
National	1.022	19.663	1.876	0.996	0.984	573
Quintile	0.958	21.066	1.873	0.997	0.985	573
RuralUrban	0.800	16.281	1.479	0.998	0.990	573
Subnational	0.859	16.058	1.601	0.997	0.991	566
ew-s						
National	2.042	29.748	3.056	0.995	0.984	573
Quintile	1.563	29.036	2.472	0.996	0.985	573
RuralUrban	1.435	25.184	2.181	0.997	0.989	573
Subnational	1.606	24.996	2.401	0.996	0.990	566
ew-ws						
National	1.738	12.229	2.837	0.997	0.997	573
Quintile	1.255	14.369	2.138	0.998	0.995	573
RuralUrban	1.211	12.024	1.963	0.999	0.996	573
Subnational	1.353	11.185	2.167	0.998	0.997	566

Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Note: The table shows different validation metrics of every survey data for a country, then average over all countries. Aggregated numbers weigh each country by the inverse of its number of surveys, such that the total weight of each country equals one. Results from different population groups are aggregated to the national target level.

Table A2 – Validation of infrastructure deprivation (all three) metrics by scenario and population level of fused data

	MAE	MAPE	RMSE	r	ρ	N
e-w-s						
National	2.836	65.821	4.616	0.964	0.948	573
Quintile	2.075	66.104	3.358	0.983	0.962	573
RuralUrban	1.653	59.328	2.578	0.989	0.964	573
Subnational	2.078	60.465	3.235	0.982	0.971	566
e-ws						
National	2.108	59.421	3.408	0.980	0.972	573
Quintile	1.540	61.129	2.544	0.990	0.976	573
RuralUrban	1.199	54.747	1.944	0.993	0.978	573
Subnational	1.483	53.258	2.333	0.990	0.983	566
es-w						
National	2.092	54.585	3.400	0.980	0.975	573
Quintile	1.587	54.544	2.601	0.989	0.978	573
RuralUrban	1.201	44.285	1.922	0.993	0.981	573
Subnational	1.500	46.386	2.334	0.991	0.984	566
es-ws						
National	1.104	35.617	2.004	0.993	0.988	573
Quintile	0.899	39.050	1.689	0.995	0.988	573
RuralUrban	0.654	29.247	1.233	0.997	0.991	573
Subnational	0.760	28.407	1.315	0.997	0.992	566
ew-es						
National	0.500	19.236	0.980	0.998	0.995	573
Quintile	0.475	34.908	0.906	0.998	0.993	573
RuralUrban	0.415	18.095	0.768	0.999	0.996	573
Subnational	0.424	17.154	0.776	0.999	0.997	566
ew-s						
National	1.654	51.657	2.711	0.985	0.979	573
Quintile	1.226	65.859	1.998	0.993	0.981	573
RuralUrban	1.016	46.114	1.623	0.995	0.983	573
Subnational	1.211	44.445	1.912	0.992	0.986	566
ew-ws						
National	0.555	28.849	1.165	0.997	0.992	573
Quintile	0.448	48.699	0.931	0.998	0.991	573
RuralUrban	0.396	30.945	0.827	0.999	0.992	573
Subnational	0.383	22.531	0.795	0.999	0.994	566

Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Note: The table shows different validation metrics of every survey data for a country, then average over all countries. Aggregated numbers weigh each country by the inverse of its number of surveys, such that the total weight of each country equals one. Results from different population groups are aggregated to the national target level.

Table A3 – Validation of “AF method (headcount * intensity)” metrics by scenario and population level of fused data

	MAE	MAPE	RMSE	r	ρ	N
p-c-r-e-w-s						
National	1.108	16.705	1.521	0.997	0.994	573
Quintile	0.685	11.790	0.976	0.999	0.996	573
RuralUrban	0.660	13.146	0.887	0.999	0.996	573
Subnational	0.759	13.078	1.018	0.999	0.995	566
p-cr-ews						
National	0.494	7.436	0.676	0.999	0.999	573
Quintile	0.222	4.139	0.348	1.000	1.000	573
RuralUrban	0.318	5.874	0.463	1.000	1.000	573
Subnational	0.361	5.755	0.506	1.000	0.999	566
pce-rews						
National	0.246	4.955	0.370	1.000	1.000	573
Quintile	0.099	3.134	0.183	1.000	1.000	573
RuralUrban	0.181	4.297	0.293	1.000	1.000	573
Subnational	0.201	4.226	0.307	1.000	1.000	566
pcew-rews						
National	0.192	3.640	0.308	1.000	1.000	573
Quintile	0.083	2.364	0.164	1.000	1.000	573
RuralUrban	0.153	3.376	0.265	1.000	1.000	573
Subnational	0.157	3.140	0.261	1.000	1.000	566
pcews-rews						
National	0.104	2.143	0.196	1.000	1.000	573
Quintile	0.052	1.598	0.132	1.000	1.000	573
RuralUrban	0.092	2.057	0.200	1.000	1.000	573
Subnational	0.087	1.904	0.175	1.000	1.000	566
pcr-ews						
National	0.448	6.112	0.636	1.000	0.999	573
Quintile	0.217	3.578	0.350	1.000	1.000	573
RuralUrban	0.283	4.710	0.440	1.000	1.000	573
Subnational	0.326	4.683	0.475	1.000	1.000	566
pcre-ws						
National	0.706	10.763	0.976	0.999	0.996	573
Quintile	0.454	7.963	0.670	1.000	0.997	573
RuralUrban	0.404	8.452	0.562	1.000	0.997	573
Subnational	0.490	8.319	0.660	1.000	0.997	566

Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Note: The table shows different validation metrics of every survey data for a country, then average over all countries. Aggregated numbers weigh each country by the inverse of its number of surveys, such that the total weight of each country equals one. Results from different population groups are aggregated to the national target level.

Table A4 – Validation of “Deprived in at least 1 (any)” metrics by scenario and population level of fused data

	MAE	MAPE	RMSE	r	ρ	N
p-c-r-e-w-s						
National	6.358	14.242	8.194	0.986	0.994	573
Quintile	2.974	7.955	3.873	0.997	0.997	573
RuralUrban	3.960	9.885	4.987	0.996	0.997	573
Subnational	4.524	10.775	5.788	0.992	0.996	566
p-cr-ews						
National	3.836	7.376	5.106	0.995	0.998	573
Quintile	0.965	2.413	1.335	1.000	1.000	573
RuralUrban	2.349	4.913	3.088	0.998	0.999	573
Subnational	2.688	5.387	3.568	0.997	0.999	566
pce-rews						
National	1.427	3.219	2.045	0.999	0.999	573
Quintile	0.404	1.247	0.624	1.000	1.000	573
RuralUrban	1.039	2.490	1.452	0.999	1.000	573
Subnational	1.142	2.624	1.609	0.999	1.000	566
pcew-rews						
National	1.102	2.365	1.639	0.999	0.999	573
Quintile	0.320	0.930	0.515	1.000	1.000	573
RuralUrban	0.859	1.972	1.243	0.999	1.000	573
Subnational	0.903	1.990	1.330	0.999	1.000	566
pcews-rews						
National	0.313	0.691	0.641	1.000	1.000	573
Quintile	0.086	0.282	0.166	1.000	1.000	573
RuralUrban	0.273	0.687	0.526	1.000	1.000	573
Subnational	0.257	0.607	0.502	1.000	1.000	566
pcr-ews						
National	3.217	6.096	4.261	0.997	0.999	573
Quintile	0.912	2.155	1.286	1.000	1.000	573
RuralUrban	1.937	3.993	2.533	0.999	0.999	573
Subnational	2.275	4.465	3.010	0.998	0.999	566
pcre-ws						
National	3.585	7.459	4.720	0.996	0.998	573
Quintile	1.788	4.204	2.399	0.999	0.999	573
RuralUrban	2.208	5.119	2.861	0.999	0.999	573
Subnational	2.601	5.686	3.395	0.998	0.999	566

Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Note: The table shows different validation metrics of every survey data for a country, then average over all countries. Aggregated numbers weigh each country by the inverse of its number of surveys, such that the total weight of each country equals one. Results from different population groups are aggregated to the national target level.

Table A5 – Validation of “Deprived in all 6 indicators (all)” metrics by scenario and population level of fused data

	MAE	MAPE	RMSE	r	ρ	N
p-c-r-e-w-s						
National	0.434	85.023	1.028	0.872	0.827	573
Quintile	0.230	59.526	0.533	0.954	0.851	573
RuralUrban	0.313	72.438	0.714	0.941	0.842	573
Subnational	0.365	72.378	0.850	0.939	0.866	566
p-cr-ews						
National	0.339	73.744	0.770	0.951	0.862	573
Quintile	0.126	41.725	0.290	0.987	0.879	573
RuralUrban	0.227	59.622	0.513	0.972	0.869	573
Subnational	0.274	60.885	0.617	0.982	0.891	566
pce-rews						
National	0.206	47.144	0.474	0.982	0.902	573
Quintile	0.119	37.873	0.273	0.988	0.906	573
RuralUrban	0.180	45.208	0.408	0.984	0.903	573
Subnational	0.178	41.614	0.393	0.991	0.916	566
pcew-rews						
National	0.164	41.810	0.375	0.983	0.914	573
Quintile	0.095	33.929	0.232	0.989	0.919	573
RuralUrban	0.153	41.310	0.347	0.986	0.915	573
Subnational	0.135	36.914	0.296	0.992	0.931	566
pcews-rews						
National	0.141	38.191	0.327	0.986	0.925	573
Quintile	0.079	30.878	0.206	0.991	0.929	573
RuralUrban	0.136	37.751	0.316	0.987	0.926	573
Subnational	0.117	33.242	0.260	0.992	0.939	566
pcr-ews						
National	0.239	59.117	0.550	0.971	0.887	573
Quintile	0.123	39.499	0.285	0.988	0.893	573
RuralUrban	0.159	46.941	0.360	0.983	0.891	573
Subnational	0.183	48.469	0.410	0.990	0.910	566
pcre-ws						
National	0.225	55.379	0.529	0.970	0.919	573
Quintile	0.126	38.191	0.309	0.982	0.923	573
RuralUrban	0.151	43.132	0.347	0.982	0.923	573
Subnational	0.175	44.893	0.409	0.986	0.934	566

Source: World Bank calculations using survey data accessed through the Global Monitoring Database.

Note: The table shows different validation metrics of every survey data for a country, then average over all countries. Aggregated numbers weigh each country by the inverse of its number of surveys, such that the total weight of each country equals one. Results from different population groups are aggregated to the national target level.