

Research paper

Global poverty: A first estimation of its uncertainty[☆]Michail Moatsos^{a,*}, Achillefs Lazopoulos^b^a Department of Social and Political Sciences, Bocconi University, and Department of History and Art History, Utrecht University, Drift 6, 3512 BS Utrecht, Netherlands^b Institute for Theoretical Physics, ETH Zurich, Otto-Stern-Weg 1, 8093 Zurich, Switzerland

ARTICLE INFO

JEL Codes:
I32Keywords:
Global Poverty
MDG1
Cost of basic needs
dollar a day
total error
confidence interval

ABSTRACT

The dollar-a-day method, applied in monitoring the UN's development goals against poverty, provides no confidence interval for the official figures of global poverty reduction, a practice that does not allow statistical testing. Using Monte Carlo micro-simulations we construct confidence intervals that reflect the error introduced by the process of determining the International Poverty Line, as well as the uncertainty of the involved Purchasing Power Parity exchange rates. These estimates identify a reduction of 5.19% between 1990 and 2015 at 95% confidence level, in stark contrast with the remarkable 73% reduction of global poverty reported in the World Bank official statistics published on September 18, 2018. At the same time, MDG1 obtains with a 80% confidence level. The cost-of-basic-needs method paints a more promising picture identifying a 35.71% reduction at 95% confidence level, while the confidence level at which poverty in 2015 was half of 1990 stands at 46%. We conclude that the derivation method of the international poverty line introduces high levels of uncertainty in the estimates.

1. Introduction

Global Poverty reduction and eradication are the prime objectives of the global development agenda for the United Nations, while the reduction of poverty has been one of the most important indicators in development (Bowley, 1915; United Nations, 2015a). On a global scale concerns regarding the precision of the available global estimates have been recently raised:

“[T]he margin of uncertainty for the global poverty estimates is so large that there must be serious questions about whether they are worth doing in anything like their current form”, Advisory Board

member, *World Bank Commission on Global Poverty* (Atkinson, 2016, p.54).

“World Bank economists have often felt uneasy about the misleading precision with which our poverty estimates appear to become imbued in public debate, despite our best efforts to document in detail the very significant uncertainty involved in each of the various steps leading up to the final numbers.”, P. Romer, A.Revenga, & F. Ferreira, *A Cover Note to the Report of the Commission on Global Poverty* (The World Bank, 2016, p.5).

Both the Millennium Development Goal 1 (MDG1)¹, aiming at the

[☆] The authors wish to thank the editor and three anonymous reviewers of this journal for their comments and suggestions, Angus Deaton for providing us his estimates on PPP uncertainty, and Robert Allen for providing us his international poverty line preliminary calculations. We also benefited greatly from discussions with Francois Bourguignon, Martin Ravallion, Francisco Ferreira, Christoph Lakner, Prem Sangraula, Robert Inklaar, Daniel Oberski, Mark Sanders, Bram van Besouw, Tim van der Valk, Vincent Schippers, Svetlana Gerakaki, Sarah Carmichael, Auke Rijpma, and Jan Luiten van Zanden. We wish to thank for their remarks participants at the 2017 PEGNet Conference, 2018 PSE Summer School on Development, the 2018 University of Athens Open Seminar in Economic History, and the Utrecht School of Economics 2018 seminar series. We further wish to thank Jaap Oudesluijs for giving us access to necessary hardware, and Kees van Eijden for granting us access to the Dutch HPC (SURFSARA). All analysis has been conducted with R open source statistical computing software (R Core Team, 2018). All remaining errors are ours alone.

^{*} Corresponding author.

E-mail addresses: m.moatsos@uu.nl (M. Moatsos), lazopoulos@itp.phys.ethz.ch (A. Lazopoulos).

¹ MDG1: “Target 1.A: Halve, between 1990 and 2015, the proportion of people whose income is less than \$1.25 a day” from <http://www.un.org/millenniumgoals/poverty.shtml>, accessed on March 14, 2017. The United Nations inform us on the same page that “The target of reducing extreme poverty rates by half was met five years ahead of the 2015 deadline.”

reduction of global extreme poverty rate by half between 1990 and 2015, and the Sustainable Development Goal 1.1 (SDG1.1)² which aims at its eradication, rely on our ability to compare global poverty rate estimates. However, both goals are thus far evaluated on the basis of World Bank's point estimates which do not allow for any meaningful statistical test to be applied in order to support or reject their success at a desirable confidence level.³ Before setting a framework for tracking the evolution and the success of SDG1.1, one should take careful stock with respect to the results concerning MDG1, while considering—in particular—the magnitude of the uncertainties involved. According to the World Bank's official data MDG1 has been achieved 5 years in advance of its 2015 deadline,⁴ with the official estimate for 2015 being issued in September 2018.⁵ However, without the margins of error of the reported point estimates, one cannot know the confidence level at which this important claim is obtained. Naturally, the same will apply to SDG1.1, a decade or so down the road.

To that end, in October 2016 the *Commission on Global Poverty*, presided by the late Sir Tony Atkinson, published a set of recommendations to the World Bank (hereafter also referred to as the Bank) on the issue of measuring global poverty (Atkinson, 2016). Recommendation 5 stipulates that, the Bank should provide estimates of the errors involved in measuring poverty using a “total error” approach.⁶ (p.50) In doing so the Bank should evaluate the sources and magnitudes of error, “particularly nonsampling error and the error introduced by the process of determining the” dollar-a-day (hereafter DAD) poverty line (Atkinson, 2016, p.50).⁷

Poverty can be thought of as being related to the deprivation from economic resources that would cover one's (basic) needs (Rowntree, 1901; Townsend, 1979). Poverty is also related to the lack of some fundamental freedoms and lack of agency of the individuals, and the fulfillment of one's capabilities (Sen, 1985).⁸ Poverty is also being defined as a multidimensional phenomenon (Alkire & Foster, 2011). However, the definition used from the perspective of global poverty, and the aforementioned UN development goals, resembles the severe lack of the necessary economic resources, and it is usually dubbed as “extreme poverty”.

Global poverty is being measured since the early 90's using the DAD approach (Ravallion, Datt, & van de Walle, 1991). The DAD method consists of estimating a poverty line that is somehow representative of the national poverty lines (NPLs) used by a group of least economically developed countries. The assumption being that such a poverty line, dubbed the international poverty line (iPL), does capture well the level

of absolute poverty across the world. The additional assumption is that the conversion of the NPLs, and other statistics used, from local currencies to a common currency, and in this case the international PPP Dollars (more on this in the next section), does represent the relative living standard conditions of those living in poverty across the world. Once the iPL is established, then the global poverty rate can be estimated using the available consumption and income distributions across the world.

National Poverty Lines (NPLs) across most developing countries are estimated using the cost of basic needs method. This consists of specifying a basket of goods (e.g. food, housing and energy) and services (e.g. education and health) that are considered to be reflecting the basic human needs in their respective national contexts. Specifying the exact composition of the various items in the basket, and how that changes over time is a matter open to much debate. But, one way or another, the national authorities responsible for tracking poverty choose the exact configuration, often with the assistance of the World Bank, in order to produce the relevant national statistics. Those NPLs in turn feed the DAD approach in the process of defining the level of the iPL (details below).

In the present exercise we estimate the global aggregates on poverty, replicating the World Bank's DAD method, reproducing it within a framework of micro-simulations (using the Monte Carlo method) to estimate the margins of error. Consequently both the international Poverty Line (iPL) and the global poverty rates are hereafter considered as stochastic variables. In addition, we operationalize one more of the Commission on Global Poverty recommendations—number 15—that advises the Bank to utilize a cost of basic needs (CBN) method as an alternative global poverty indicator.⁹

We compare the two methods in terms of the size of their estimates' error terms, and in terms of the confidence level at which the global poverty rate in 2015 is half that of 1990, and, further, we investigate the percentage of poverty reduction that took place between 1990–2015 for each of the methods at a 95% confidence level. In doing so we come to the conclusion that none of the two methods provides evidence in support of MDG1's success at a 95% confidence level. The CBN method shows that the world has come over half-way in halving global poverty during the 1990–2015 period at 95% confidence level,¹⁰ while at the same confidence level DAD method identifies a reduction in global poverty at a rate of about 5%. As discussed in the sections below, this unexpected finding is attributed to the iPL's derivation method, and in particular the averaging of a set of national poverty lines.

2. Materials and methods

2.1. Dollar-a-day poverty Lines

This approach has been developed by the World Bank researchers over a long period since 1991 (Ravallion et al., 1991; Chen & Ravallion, 2001; Ravallion, Chen, & Sangraula, 2009; Ferreira et al., 2015), and its core conceptual origin can be traced further back to 1979 in Ahluwalia, Carter, and Chenery (1979) who seem to be the among the first to use

² SDG Target 1.1: “By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than \$1.25 a day” from <https://sustainabledevelopment.un.org/sdg1>, accessed on March 14, 2017.

³ The World Bank also functions as the UN custodian institution for monitoring global poverty.

⁴ http://www5.worldbank.org/mdgs/poverty_hunger.html, last accessed October 31, 2017.

⁵ On September 18, 2018 the World Bank published the global poverty estimates for 2015 (the latest year for which a global estimate is available).

⁶ Recommendation 5: “The World Bank poverty estimates should be based on a “total error” approach, evaluating the possible sources, and magnitude, of error, particularly nonsampling error and the error introduced by the process of determining the International Poverty Line” (Atkinson, 2016).

⁷ It should be noted that despite the long line of criticism on the Bank's global poverty measurement methodology (Deaton, 2010; Reddy & Pogge, 2010; Srinivasan, 2009; Subramanian, 2015; Moatsos, 2017), the point regarding the missing error terms has not been raised as much as its importance in monitoring the evolution of global poverty entails.

⁸ The Multidimensional Poverty indicator developed by Alkire and Foster (2011) has recently gained momentum, and has been adopted by the United Nations as one additional poverty measurement indicator. Here we focus on the traditional and more prominent version of global poverty measurement adopted by the World Bank and the United Nations for monitoring global poverty.

⁹ Recommendation 15: “The World Bank should develop a program of work, in conjunction with other international agencies, on a basic needs-based estimate of extreme poverty; these estimates would, when developed, form an alternative indicator to be included in the portfolio of Complementary Indicators, and serve to provide an interpretation of what the International Poverty Line would buy.” (Atkinson, 2016, p.xxi). The World Bank decided not to pursue with the implementation of this recommendation (The World Bank, 2016).

¹⁰ Do note that strictly speaking one cannot investigate the success of MDG1 using a CBN method, since MDG1 is formulated in DAD terms as it uses a dollarized poverty line—which also happens to be the iPL obtained via the DAD method. However, one can investigate whether some CBN based poverty rate has been halved during the MDG1 period. See Sections 2.4.2 and 3.2 for the details and the constraints of this investigation.

PPP exchange rates in international poverty estimates (Dhondge & Minoiu, 2011). Ravallion et al. (1991) replaced the Indian poverty line used by Ahluwalia et al. (1979), with the nearest round number of a small cluster of national poverty lines (NPLs) expressed in PPP dollars. That number being relatively close to one dollar (in 1985 prices) gave birth to the “dollar-a-day” method.

When the 2005 ICP PPP rates became available, (Ravallion et al. (2009), hereafter RCS) provided an update of the iPL estimate using the average of NPLs from the 15 countries with per capita consumption of less than \$60 per day in 2005 PPP dollars (the relevant plot is reproduced in Fig. 5 of the appendix). This group of 15 countries constitutes a core concept in the DAD methodology and it is dubbed as the *reference group*. The method for selecting this reference group builds on the relation between NPLs and consumption levels. Adding or removing countries from the reference group affects the resulting iPL.

Ferreira et al. (2015, hereafter FEA) update the iPL to the 2011 PPPs. However, FEA do not repeat the entire procedure that RCS specify, and decide instead to use the same 15 countries used in RCS. This makes the current iPL a hybrid PPP-based global poverty line.

In what follows we review the key points of the DAD method in more detail, and distinguish key steps that operate as uncertainty entry points. In that respect, our contribution to the DAD method is to introduce key elements of uncertainty.

We begin by examining the procedure that selects the reference group of countries that determine the iPL. To that end RCS fit the scatterplot of NPLs versus consumption per month for each country (see also Fig. 5), using the following equation:

$$Z_i = Z^* I_i + f(C_i)(1 - I_i) + \varepsilon_i \quad (1)$$

where Z_i is the poverty line in country i , Z^* is the average NPL of the reference group, I_i takes the value of 1 if country i is in the reference group and 0 otherwise, and $f(C_i)$ is a function of consumption per capita per month C_i . The function $f(C_i)$ is later defined and fitted as a linear function of C_i . The estimated version of Eq. 1, as it appears in RCS, is as follows (in parenthesis are provided t-ratios based on robust standard errors):

$$Z_i = 37.983I_i + (19.388 + 0.326C_i)(1 - I_i) + \hat{\varepsilon}_i(12.55)(2.99)(11.15)R^2 = 0.890, \quad n = 74. \quad (2)$$

The decision to split the group of countries, and attempt an elbow fitting to the data,¹¹ is made on the basis that the iPL capturing absolute poverty levels would be around the minimum NPL among the countries with low consumption per capita.¹² At the same time, they argue that simply taking only the poverty line of the country with the lowest consumption, and use it as the iPL, will make the method prone to measurement errors at the country level. Based on these observations they turn to an averaging of the NPLs from a group of least consumption affluent countries, aka the aforementioned reference group, which is

¹¹ Meaning fitting the data with two consecutive regression lines where a kink is created in their junction. In this particular case the first line is a horizontal one. See also Fig. 5, but do note that the figure is log-normal, thus the second straight line is shown as a curve. This procedure is also known as a piecewise fit.

¹² In more detail, they argue that for the countries with the lowest consumption per capita per month the estimated poverty line when using the entire sample is lower than the average poverty line for that group. Put differently, the fitted line of a regression that considers the entire dataset gives a lower estimate for poverty at the lowest level of consumption found in the sample. RCS found this identification as non-satisfactory.

Table 1

The yearly cost of basic needs consumption basket.

Item	Units	Basic Needs Basket	SD (%) ***
Energy Target	kcal/ day	MDER	****
Protein Target	kcal/ day	0.75 gr/kg	16.2
Minimization Meat or fish	- kg/ year	mean of 3 cheapest bundles 8 or 16*	none -
Butter or oil or ghee	kg/ year	8	-
Sugar	kg/ year	5	-
Main staple(s)	kg/ year	based on kcal/protein target (estimated as a residual of the above)	-
Clothing and Footwear	%	budget shares**	16
Water	%	budget shares**	65
Energy	%	budget shares**	20
Housing	%	budget shares**	22
Health	%	budget shares**	39
Education	%	budget shares**	33

*: 8 kg of meat or 16 kg of fish, whichever is cheaper.

**: The budget shares are available from World Bank's Global Consumption database. See Section 2.5 for more details.

***: For the discussion of the standard deviations (SD) of nutrient targets, food quantities and budget shares shown here see Section 2.5. Also note that for Brazil, India, and South Africa we use the country data for the uncertainty in the budget shares (shown in Section 2.4.2), while the aggregate of the three countries that is reported in this table is used as a substitute for the uncertainty in all other countries.

****: Depends on the country demographic profile and other parameters as discussed in Section 2.5.

achieved by constraining the first section of the elbow fit to have zero gradient. As RCS put it: “there are measurement errors and methodological differences between countries in how poverty lines are constructed, which can be interpreted as noise in the mapping from the

underlying welfare space into the income space.” Thus, averaging is an attempt to avoid country specific errors. Naturally this attempt, as any averaging, comes with an error in the estimate of the average that needs to be propagated into the global poverty rate estimates (Bailer-Jones, 2017, Section 2.8).

The consequent issue is how to decide where the reference group threshold should be. RCS set two criteria for this: (a) the continuity criterion that requires $Z^* = f(C^*)$ at the threshold, where the elbow's kink is also located, and (b) the consistency criterion that requires $C_i < C^*$ for all countries i in the reference group. Since $f(C)$ and Z^* are estimated from the data they can be used to determine the threshold level by enforcing the continuity criterion (a). Next RCS test if for the resulting C^* consistency criterion (b) holds as well. As a result, RCS choose the threshold to be at \$60 that gets 15 countries in the reference group, since, as they report, taking the 10 or 20 poorest countries produces “not consistent reference groups, unlike that defined by the poorest 15 countries”.

In addition to the uncertainty introduced by the estimation method of iPL (whose key operation is the splitting of the sample in two parts and averaging the NPLs of the reference group), we investigate the influence that the uncertainty in the PPP exchange rates introduces initially to the value of the NPLs and the consumption per capita, and

consequently to the identification of the threshold for the reference group. To do so we conduct a Monte Carlo pseudo-experiment in which the PPP exchange rates for both 2005 and 2011 are chosen from a lognormal distribution around the mean PPP rate for each country using their respective error terms (see Section 2.5 below for details on data sources). Arguably a more consistent application of the dollar-a-day methodology would be to redo the exercise in 2011 PPP rates altogether, both for the threshold selection and the averaging of NPLs. However, we are interested in examining the Dollar-a-Day poverty line methodology as it is applied by the World Bank for monitoring MDG1 and described in FEA.

Overall, there are two sources of uncertainty on the iPL considered in this work: one induced by the uncertainty on PPP exchange rates, and one induced by determining the iPL by averaging within the reference group. The later is also the greatest source of uncertainty in magnitude. The former influences the total error on the iPL in several intertwined ways, but their impact turns out to be relatively small.¹³

2.2. Cost of basic needs poverty lines

Following the aforementioned recommendation 15 of the World Bank's Commission on Global Poverty we employ the Cost of Basic Needs (CBN) method in evaluating global poverty for MDG1's reference years. The composition of the CBN poverty lines rests on two building blocks, as shown in Table 1. The upper block, consists of the food component, and the lower block of non-food expenses (separated by a double line). The food component builds on the bare bones basket concept introduced by Bob Allen (Allen, 2001; Allen, 2013) in estimating welfare ratios and real wages historically, as well as poverty lines contemporary (Allen, 2017). The quantities used on each food source are such that the global poverty rate for 2015 is comparable with the World Bank's point estimate according to the baseline scenario of the DAD method.¹⁴ This choice is motivated by the observation that higher DAD poverty lines show less appealing poverty reduction, since their downward trend is slower (Chen & Ravallion, 2010, p.1619, table VIII).¹⁵ We wish to avoid such a comparability problem that could arise if a CBN-based global poverty rate would be much higher or much lower than the DAD. Constrained by this methodological consideration our implementation of the aforementioned definitions becomes a frugal one, as shown in Table 1.¹⁶ Do note that we take the mean of 3 cheapest food bundles as a technical way to introduce some minimal dietary variety, and no uncertainty stemming from this step is considered.

The non-food component uses the budget shares of expenditure categories that are relevant to the aforementioned poverty definitions (as shown in Table 1). The cheapest bundles that satisfy the recipe shown are identified using linear programming, thus accounting for substitution effects by those living in conditions of poverty, and the solution in the main staple may contain more than one products. This means that the consumption basket is not held fixed in its composition, rather it is a goal oriented basket that achieves a certain level of welfare according to a fixed recipe. The total amount of basic nutrients (calories and proteins) that the food component should contain is calculated using the FAO (2001) methodology for Minimum Dietary Energy Requirement (MDER). Accordingly, the calories required per day are estimated via the

MDER, and the proteins as a linear function of the population weighted average body weight derived in the MDER calculations with a proportion of 0.75 gr of protein per kg of weight (World Health Organization, 1985, Section 8.2.2). Section 2.5 discusses the various uncertainties shown in Table 1.

2.3. Underlying uncertainties common to both methods

Here we discuss two important methodological choices that the dollar-a-day and the cost of basic needs approach have in common.

The first one concerns the use of sparse distributional data. In order to achieve a satisfactory population coverage, the World Bank re-uses the available consumption or income distributions outside of their year of origin. For comparability, we follow this approach of the Bank as detailed in (Chen & Ravallion, 2010; Chen & Ravallion, 2004 & Ferreira et al., 2015). To extrapolate the distribution to a different year a relevant variable from national account statistics (NAS) is used to estimate the mean of the distribution, while the shape of the distribution remains unchanged.¹⁷ Preference is given to real growth in household final consumption expenditure from the NAS, and, when that is not available, the fall-back option is to use real GDP per capita growth. The main empirical issue with this method is that there is a divergence among the "organic" growth rates between consecutive household surveys (HHS) and the growth from NAS (Deaton, 2005, p.2). To account for this discrepancy between HHS and NAS growth we follow the method applied by FEA. It uses an 87% multiplier as an adjustment factor between NAS and HHS growth rates to all countries, with only two exceptions: for India the correction factor is 51% and for China 72%. Although this adjustment is a welcomed correction towards the right direction, it is not in itself error less. However, we choose to ignore this source of error in the present treatment, as more straightforward addressable and important sources of error are considered. Further, since this source of error is common for the two methods that we compare, our choice bears no cost in this respect, but in all likelihood this translates to an underestimation of the "total error".

The second important methodological choice, that the two methods have in common, concerns the treatment of countries without sufficient data to estimate their poverty rates. In such situations the World Bank opts to impute poverty rates for the countries without data with the average regional poverty. The implicit assumption is that countries which have (or manage to somehow acquire) the means to monitor poverty are a good proxy for poverty levels in countries that don't. This is a rather strong assumption that one would like to account for. However, we will not be treating this aspect of the calculation in our main approach, and only as a robustness check we will simply add a normally distributed error term each time such an imputation is made.¹⁸

In addition, do note that both methods exclude the developed countries. Fortunately, for investigating the success of MDG1 this is appropriate, as only the developing world is the reference domain of MDG1.¹⁹ Finally, the measurement error and the incomplete measurement of consumption, or income, is roughly considered using a normally distributed error term. Again this is used only as a robustness check, and it is not part of our main treatment.²⁰

¹³ Observe the quadrupling of the confidence interval in the iPL when considering the averaging and when not, as shown in Fig. 3. At any rate, ignoring the uncertainty stemming from the averaging step in the method of derivation of the iPL results in MDG1's success at a 95% confidence level (see Section 3 for details).

¹⁴ Corresponding to row with ID 1 in Table 2

¹⁵ See also Fig. 6 in the appendix. Observe there that the derivative of the curve would be decreasing at values beyond the standard iPL.

¹⁶ Otherwise, other more demanding CBN configurations could have been operationalized here, more similar to those discussed in Allen (2017).

¹⁷ For years between two available distributions their (time) weighted average is used.

¹⁸ The exact values used are shown in the caption of Table 2 in the results.

¹⁹ For comparability we keep the same definition for developing countries with the World Bank. If the country was a developing one in 2005 it is included for both 1990 and 2015 regardless of its actual development status at those benchmark years. Note that the application domain of SDG1.1 is the entire world and not just the countries characterized as developing back in 2005.

²⁰ See Table 4 and next section for more details.

2.4. Treatment of uncertainty with Monte Carlo Micro Simulations

The main purpose of the paper is to assign an uncertainty to the global poverty rate. Depending on the method used to evaluate it, the global poverty rate depends on a number of external variables (e.g. the PPP exchange rates, monthly consumption data, NPLs, prices etc.) or choices (e.g. the choice within the DAD method to average the reference group NPLs to get an iPL). In traditional error propagation approaches, each external variable's value is used as a point estimate, and its corresponding uncertainty is propagated so that it impacts the total uncertainty, in a way that depends on the functional form of the derived quantity as a function of the external variables.

In this work, we consider a set of key external variables that comes with a central value and an uncertainty, as being stochastic. Instead of using the central value, we use a random value drawn from a distribution that follows the distribution of that external variable from its provider, when available (e.g. a lognormal for the PPP rates) or a normal distribution, in the absence of detailed information and when there are no good reasons to assume that the central limit theorem does not apply. Treating each external variable as a stochastic one, is, in fact, equivalent to simulating the fluctuations of the variable, due to its inherent uncertainty, away from its central value. This stochastic modeling builds volatility and variability into the simulation and therefore provides a better representation of the real life behaviour of the external variable. Similar methods are widely used for uncertainty estimation over many scientific fields, from subnuclear physics to engineering and risk management, under the generic rubric of Monte Carlo simulations: one artificially recreates a stochastic process, many times, and directly observes the variability of the results.²¹

Following this prescription for key external variables, we compute, deterministically, the global poverty according to the method of choice. Naturally, the resulting value can be anywhere within a range of possible values. Repeating the procedure a large number of times (typically 10,000) we get a distribution of global poverty rates. The mean of that distribution is expected, based on general ergodic arguments, to converge to the global poverty rate we would find if we computed it using only central values for the external variables. The distribution, however, provides us not only with this most likely value, but also with what ranges are reasonable. The standard deviation of the distribution, in particular, provides us with a natural candidate for the uncertainty on the global poverty rate, which is what we are after.

In all cases we treat measurement errors as classical, i.e. we assume that the error on an input variable is independent of the value of the input variable itself. We also treat errors on external variables as statistical, in the sense that we use them to define a distribution for the stochastic value of the variable we use. We do not attempt to distinguish between a systematic and a statistic nature of the error on these external variables, for example the PPP rates, as such information is hardly available. We also do not have the means to attempt to reduce any of these uncertainties, through validation sets or attempt to detect what kind of bias would the noise in the different components induce via e.g. a sensitivity analysis, along the lines described in Hausman, 2001 or more recently in Schennach (2020). The only information on the uncertainties of external variables that we employ is a reasonable choice for the distribution of the variable that they imply (or to follow the prescribed distribution of the error terms as in the case of the PPP rates).

At this point we would like to distinguish between uncertainties in global poverty induced by the input variables used to compute it, from uncertainties induced by specific choices a method employs in an attempt to mitigate errors. The latter is a type of systematic uncertainty inherent to the method, one that would persist even in the idealized situation when all input quantities are measured with infinite precision.

A characteristic example, central to the discussion in the rest of the paper, is the choice of the DAD method to create a reference group of low NPL and low consumption per capita rates, and take the average of NPLs within this reference group to define the iPL. Taking the average within a reference group is in the spirit of mitigating measurement errors in the NPLs of the poorest country (whether they are substantiated or not) while still estimating absolute, as opposed to relative, poverty.

One could take the stance that such choices are part of the definition of the method and therefore no corresponding uncertainty should be accounted for. We argue that this would go against the spirit of the method, and recommendation 5 of the World Bank's Commission on Global Poverty which explicitly stipulates the consideration of the error introduced by the method used to derive the iPL. Choices are made to achieve goals, and they have an impact on resulting estimates that should be included in uncertainties, especially when, as is the case with the DAD method, the results are very sensitive to these specific choices. Within the Monte Carlo method there is a very simple approach to account for the uncertainty induced by making a choice: make the choice at random. Within every simulation point, a different reference group is produced. Instead of taking the average of the corresponding NPLs, we sample one at random. On the aggregate this produces a distribution of iPLs that has an average very close to what it would have if we pick the average NPL at every point, but it also reflects the spread of the NPLs within the reference group, during each simulation. This procedure is identical to the consideration of any other variable, such as the prices of a particular product in the CBN method. There again, in each iteration, the prices of the commodities are drawn at random following the distribution of the available data.

In what follows, we clarify various salient points of the Monte Carlo simulation procedure that are dependent on the method used to determine the global poverty rate. Table 4 in the appendix expands the table of non-sampling errors from Atkinson (2016) and provides an overview of the sources treated here and how. Clearly we do not account for all the possible sources of error, but we do account for the sources that are characteristic to each of the method. The error sources that remain

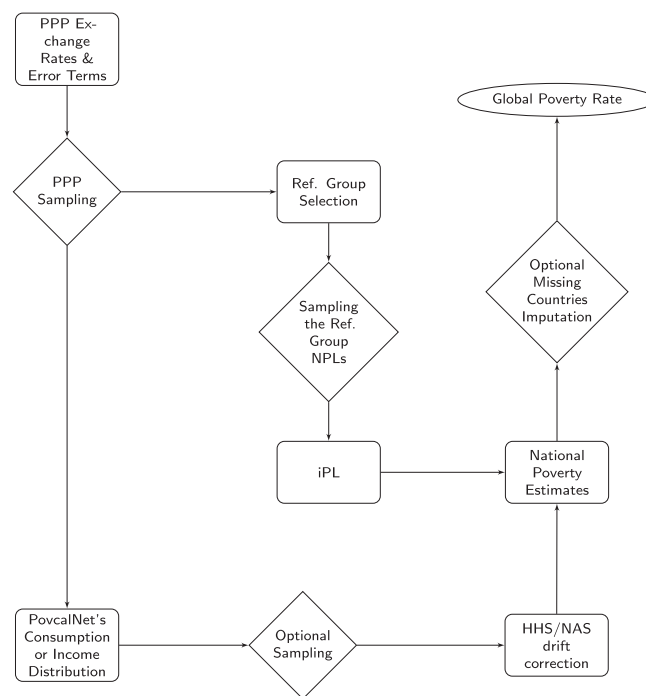


Fig. 1. Architecture of Data and Monte Carlo method for the DAD method. Diamond shaped objects mark a draw from a distribution in the Monte Carlo simulation (Authors' configuration). Entries marked as Optional are part of robustness checks only.

²¹ Monte Carlo methods in Econometrics, specifically, are reviewed in (Chib, 2001).

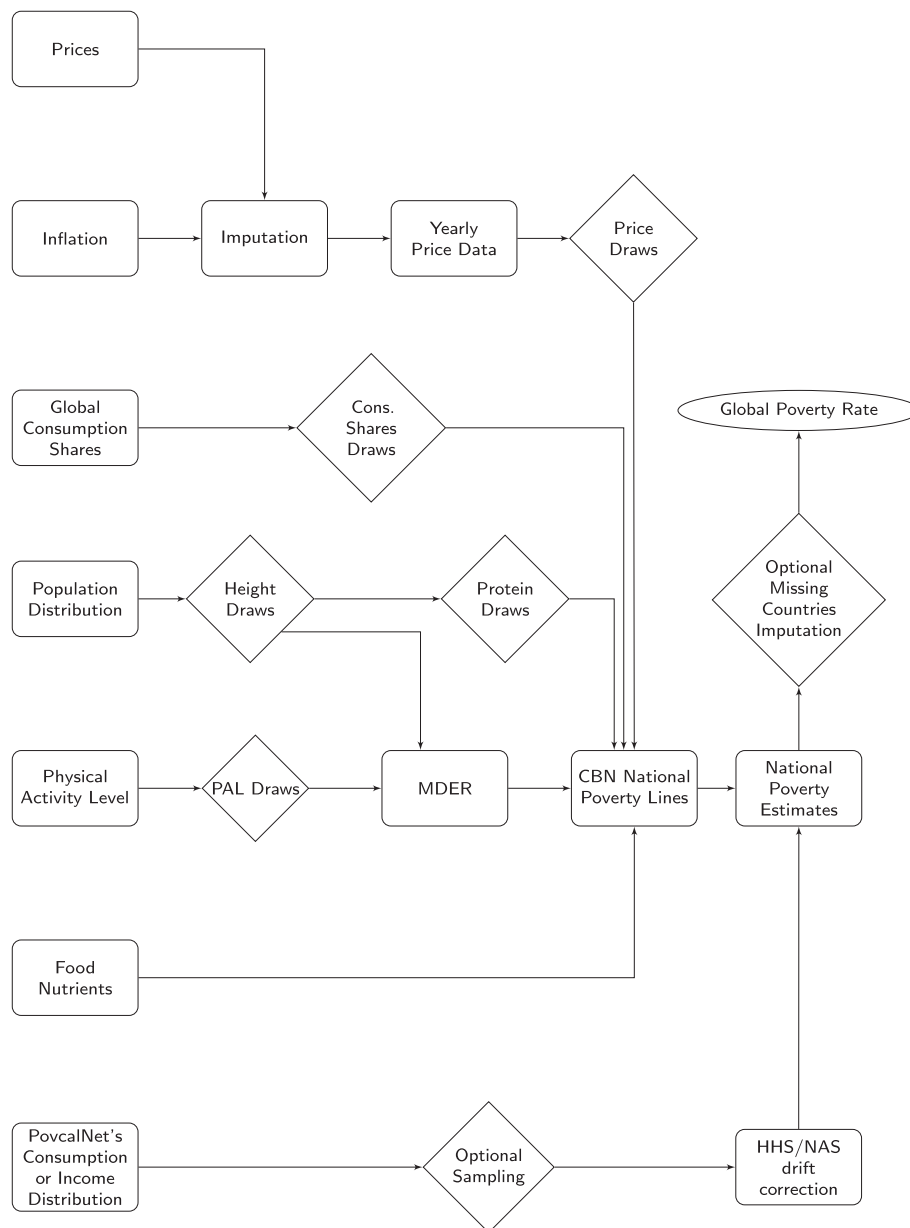


Fig. 2. Architecture of Data and Monte Carlo method for the CBN method. Diamond shaped objects mark a draw from a distribution in the Monte Carlo simulation (Authors' configuration adapting Moatsos, 2020a). Entries marked as Optional are part of robustness checks only.

unaccounted for would influence the two methods rather identically, and therefore in all likelihood this has no implication in their comparison.

2.4.1. DAD simulations

With respect to the DAD method, as discussed above, the main sources of error –which are specific to this method– that we account for are: (1) the averaging of NPLs, and (2) the underlying uncertainty in the PPP exchange rates. There are five main implications with respect to the role of PPP uncertainty are that: (a) the consumption per month threshold for selecting the countries in the reference group changes; (b) the relative and absolute consumption per month levels change; (c) the relative value of an NPL changes with respect to other countries changes; (d) the composition of the reference group selection procedure is different; (e) the relative value of the currencies within the reference group as opposed to those outside of that group also changes, making the resulting iPL relatively more or less expensive to the countries outside of the reference group.

Do note that these implications are intertwined and cannot be treated in isolation. Therefore at each PPP draw the entire set of steps of the DAD method needs to be repeated to estimate the PPP-draw specific iPL. In turn, this implies that there exists an unknown true value of the iPL that needs to be estimated separately for each of the two MDG1 reference years (1990 and 2015). Ideally if a new set of PPPs were calculated for each year, this would probably be what the World Bank would be doing as well, as implied by its decision to re-estimate the iPL after each new ICP round. The DAD method is expected to maintain welfare equivalence among PPP rounds, therefore any incompatibilities between the iPL of different PPP sets should not be seen as a problem in our estimation of uncertainty, but rather as an issue that the DAD method may have to attend to. By adding the PPP uncertainties, our approach –in addition to estimating the confidence interval– partially simulates the behavior of yearly updating the iPL.

To investigate the effect that the aforementioned sources of error have on the size of the uncertainty of the global poverty estimates, all those points (along with the methodological choices that both methods

have in common) are accounted for through each iteration in the Monte Carlo procedure. The flow chart of Fig. 1 depicts how the flow of calculation and the points where draws are taken from the underlying distributions.

As a new set of PPP exchange rates is drawn on each iteration, and then the entire set of consumption per month, the countries in the reference group, the NPLs, and consequently the iPL, and the distributional data are re-calculated, or re-calibrated, using those new rates. In turn, they are used for the estimation of the global poverty rate corresponding to that PPP draw following the procedure as described above, and in FEA in greater detail.

Overall, changes in PPP conversion factors affect what the conversion of nominal terms to (internationally comparative) “real terms” means, so errors arising from the PPPs do pertain to the exercise of estimating the iPL for any year. However, one might object that although the diamond “PPP Error Terms Draw” of Fig. 1 is part and parcel of the uncertainty estimation, a subsequent re-selection of the reference group is not (shown as the “Ref. Group Selection” box in the same figure). Such reasoning assumes that once the set of reference countries has been decided it must be left untouched and constant in the estimation of the procedure’s uncertainty. Nonetheless, this reasoning would ignore the fact that PPP fluctuation recasts the relative values of the group of countries with low consumption thus re-ranking the countries and in addition the splitting of the groups into a reference and a non-reference group may not be consistent any more. As a consequence the whole DAD procedure needs to be repeated both for consistency, and in order to actually follow the entire RCS specification procedure that the DAD approach prescribes for iPL. Put differently, there is no \$60 (in 2011 PPP terms) threshold anymore as this has changed as a result of the PPPs fluctuations, therefore the whole procedure as depicted in Fig. 1 needs to be repeated from the beginning and in its entirety.

2.4.2. CBN simulation

For the CBN method, a range of key error sources are accounted for in our implementation as described in Table 4 of the appendix and shown in the flow chart of Fig. 2 below. For each error source, treated with the Monte Carlo technique, values are drawn at random from the distributions around the mean value of each variable.

More specifically, and with respect to the price data, we use the standard deviation of prices reported within the United Nations World Food Program price dataset. For these countries, we use the aggregate national empirical distribution of prices (among countries with more than 200 observations), while for the countries without underlying data we use the aggregate distribution from their corresponding region.

The error terms for the various consumption budget shares are derived from the data available in World Bank Global Consumption database. There are 3 countries that have sub-national data, namely India, Brazil and South Africa. For those three countries we draw from the empirical distributions of each budget category, while for all other countries we draw from the aggregate distribution of those three countries on each budget category (as shown in Table 1).²²

Height data are necessary to estimate the minimum caloric intake (MDER measured in kcal²³) and the amount of proteins in grams. Their uncertainty is taken from Zanden van et al. (2013) which use the

coefficient of variation (CV) of heights distribution, and cover a substantial share of the developing countries. For countries without available information on their CV the regional average is used.²⁴

For calculating the MDER and the protein intake we further need the physical activity level (PAL) of the population living in conditions of poverty. To this end we employ the method of PAL estimation introduced by Moatsos (2020a) that simulates a daily work/leisure schedule. The resulting distribution of PAL cannot be distinguished from a normal distribution, and has a standard deviation for this is 0.19 with a mean of 2.26. This range PAL values corresponds to the vigorous lifestyle 2.00–2.40 (FAO, 2001, p.38, table 5.3). In terms of PAL we therefore assume that the distribution, and the implied uncertainty for those living in poverty, is the same in all developing countries.

Further, table 16 in World Health Organization (1985) mentions that the pooled coefficient of variation of all listed studies in mean requirement of protein is 16.2%. We will use the same uncertainty for the 0.75 gr of protein per kg of weight that the report qualifies as appropriate, and we assume a normal distribution of this error term as well.

2.5. Data

The anthropometric data we use are the age and gender population distributions from the United Nations World Population Prospects dataset (United Nations, 2015b). The height data are taken from Baten and Blum (2015) which is the most complete dataset to date.²⁵ Price data are combined from the ILO’s October Inquiry data (Moatsos, 2020b), along with prices from the World Food Program (WFP) of the UN, and FAO data also from UN. Missing observations are filled in using the most appropriate available price index, which in order of preference are: a food CPI for the poor which is available from ILO for some countries, an average CPI, or finally a price index that follows price volatility of similar products in the same country.²⁶ The composition in nutrients of the various food items is taken from the USDA database,²⁷ and retention rates are applied for caloric values following Appleton, Emwanu, Kagugube, and Muwonge (1999). Consumption shares used to account for expenses in clothing and footwear, energy, housing, health and education are taken from World Bank’s Global Consumption Database.²⁸ The welfare distributions are those available at PovcalNet.²⁹ The average (population weighted) year of available household surveys (HHS) for 1990 is 1990.6 and for 2015 is 2013.9.

With respect to the estimation of the error terms in PPP exchange rates there are three available sources. Deaton and Dupriez (2011) provide PPP error estimates that reflect sub-national sampling variability of PPP rates. The issue with such estimates is that if we were to use those we would also ought to include information about sub-

²⁴ Also for countries without height data the regional average was used, and for countries with missing data the average growth from the region was used for data imputation.

²⁵ Dataset downloaded from the Clio Infra website, at <https://www.clio-infra.eu/>, last accessed April 15, 2018.

²⁶ For a more detailed discussion of these data sources see (Moatsos, 2017).

²⁷ <https://ndb.nal.usda.gov/ndb/>

²⁸ Data are available only for 2010. Last accessed on December 6th, 2016. The budget shares of the second consumption group marked as “Low” are used. Those are preferred over the “Lowest” consumption group since the food component of our basket is typically much lower than that in the “Lowest” consumption group. And since the budget shares of that group are lower than those of the “Low” group, using these budget shares would underestimate the threshold costs for those non-food components.

²⁹ Consumption based distributions when available, income based otherwise, as applied by the World Bank and FEA as well. Last accessed on October 3rd, 2018. The underlying distributions were retrieved from the World Bank using a slightly modified python script initially provided by Dykstra, Dykstra, and Sandefur (2014). The retrieved distributional data are available in Moatsos (2018) or directly at the Data Publication platform of Utrecht University.

²² The standard deviations (relative to the mean) of the three countries is respectively for India, Brazil, and South Africa: Clothing and footwear 12, 20, and 12%; energy 25, 18, and 13%; health 21, 49, and 36%; housing 11, 29, and 18%; education 32, 36, and 26%; and water 41, 85, and 29%.

²³ For further details about its calculation see (FAO, 2008; Moatsos, 2017) for its use in poverty measurement.

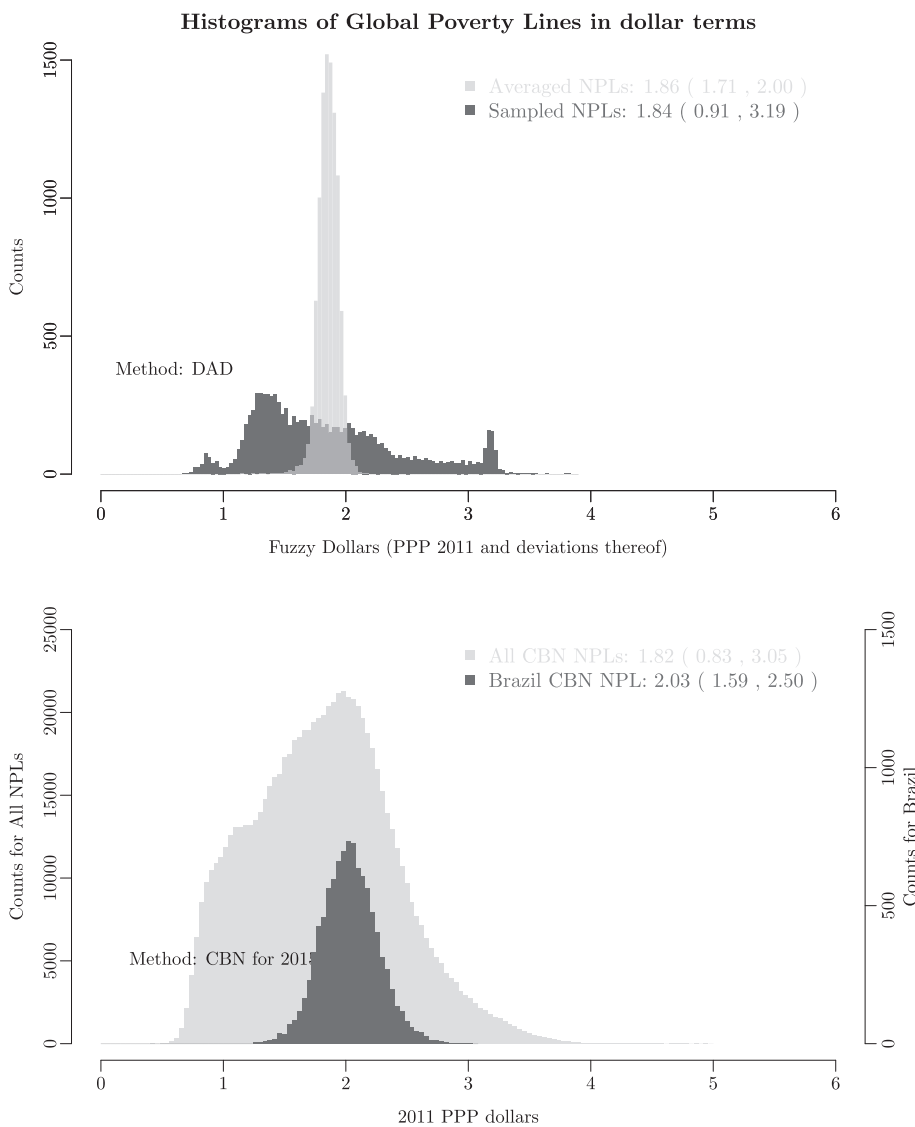


Fig. 3. Upper: includes all values of iPL, when PPP uncertainty is accounted for according to the baseline implementation of (ID 1 in dark gray), and contrasted with the estimated iPL distribution when using the average of NPLs at each iteration (shown in light gray). Lower: the CBN based NPLs for 2015 expressed in PPP dollars for ease in exposition (in light gray). For reference the NPLs calculated for Brazil are overlaid (in dark gray) using a separate y axis. The legend shows average values as well as the empirical 95% confidence interval.

national variation of income.³⁰ Fortunately the error estimates on PPP from a second source, Deaton (2012), do not reflect such sub-national information, but rather uncertainties due to variability in relative prices and expenditure patterns among countries. Deaton shows that the PPP rates of China and India have a relative standard error of 15%, with most of countries clustering at 15–17%. The potential issue with those estimates is that these estimates are work in progress, as part of a World Bank working paper, and they have not been duly refined to meet publication article standards.³¹ Another potential issue would be the fact that Deaton (2012) provides PPP standard error estimates for the GDP PPP exchange rates and not the household final consumption PPPs that are used in global poverty estimates by the DAD methodology.

A third source, and the one that we use here, is provided by Rao, Rambaldi, Huynh, Doran, and Ganegodage (2015).³² Those estimates cover 181 countries and are provided for both 2011 and 2005 ICP

rounds, while the two above sources only treat the 2005 ICP round. In addition they are available also for the PPP exchange rates related to the consumption GDP component, which is a concept closer to the required household final consumption PPP than those from GDP per se.³³ They are also less than half the errors estimated in Deaton (2012), with an average of 5.7% for 2005, making our analysis a conservative one with respect to the size of the PPP error terms. In addition, their error terms “are computed under the assumption of the lognormality” (Rao, Rambaldi, & Doran, 2010).³⁴

³⁰ This is so since nominal expenditures and nominal income correlate positively and spatially with higher price levels within countries. Therefore sub-national covariance assumptions regarding income and price levels would be required.

³¹ As Prof. Deaton has warned us in a personal correspondence.

³² Available at: <http://uqicd.economics.uq.edu.au/index.php>, last accessed April 10, 2018.

³³ The PPP rates here are those of “individual consumption expenditure by households” used by the World Bank’s PovcalNet as well.

³⁴ For the very few countries that do not have an error estimate the average of the relative error terms of the developing countries is used.

Table 2
DAD Global Poverty estimates.

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	-	-	-	-
Impl. ID	Mean iPL	SD iPL	Mean 1990 (%)	SD 1990 (perc. points)	Mean 2015 (%)	SD 2015 (perc. points)	MDG1 Conf. Level (%)	Pov. Reduc. (%)				
1	1.84	0.59	38.18	14.32	10.93	7.32	80	5.19				
2	1.81	0.47	38.11	11.92	10.29	5.9	85	28.27				
3	1.78	0.37	37.81	9.93	9.73	4.6	91	41.6				
4	1.84	0.59	38.94	14.53	11.34	7.53	79	4.39				
5	1.84	0.59	38.87	14.52	11.35	7.53	79	4.29				
6	1.81	0.47	38.8	12.09	10.71	6.07	85	27.15				
7	1.78	0.37	38.49	10.08	10.14	4.75	90	40.45				
PCN	1.90	-	43.15*	-	11.62*	-	-	-				

*: World Bank PovcalNet estimates from <http://iresearch.worldbank.org/PovcalNet/povDuplicateWB.aspx>, last visited on January 4th, 2019.

Configuration IDs: (1) baseline specifications with PPP lognormal error terms; (2) previous specification and excluding the 1 lower and 1 higher NPLs from being randomly selected at the sampling procedure; (3) previous specification, but excluding the 2 lower and 2 higher NPLs; (4) baseline specifications with PPP lognormal error terms and a normally distributed error term of 10% and a 20% increase when imputing for countries with missing data; (5) previous specification, plus a 5% error in the mean of the income or consumption distribution; (6) previous specification and excluding the 1 lower and 1 higher NPLs from being randomly selected at the sampling procedure; (7) previous specification but excluding the 2 lower and 2 higher NPLs;

3. Results

3.1. Global poverty Lines

The upper sub-plot in Fig. 3 shows two partially overlapping distributions of iPLs for the dollar-a-day approach.³⁵ The light gray spiky distribution consists of the mean values of the NPLs from each reference group of the 10,000 iterations of the Monte Carlo procedure. Methodologically it is a replication of the original RCS method, as extended for 2011 PPPs by FEA with one additional considerations (as discussed in Section 2): at each iteration new randomly selected PPP rates for both 2005 and 2011 are used, based on their respective standard errors. This distribution of means shows how the iPL, computed as an average of NPLs of the reference group, varies with the variation of PPP rates. This approach merely propagates the uncertainty from PPP rates to the iPL estimate, and does not account for the uncertainties introduced by the definition of the international poverty line.

The second distribution in Fig. 3's upper sub-plot is composed from a set of NPLs that are sampled at random, one at each iteration from the each time occurring reference group. After 10,000 iterations the distribution of those sampled NPLs is formed. Thereby simulating the overall distribution of reference group NPLs. The mean of this distribution is therefore an iPL estimate. The width of the distribution probes the uncertainty to the iPL estimate due to –not only to the PPP uncertainty– but also to the method of derivation of the iPL (namely the averaging of the reference group's NPLs).

To elaborate on the rationale behind this point one needs to reconsider the reason for this averaging by RCS. The purpose of averaging in RCS is to cancel out measurement errors in NPLs. Independently of whether this actually works –and RCS do not discuss this point– it is a very natural way of selecting a typical NPL value from the reference group. However, when estimating the related uncertainty, we need to include information about the spread of the NPLs around their average. To be more explicit, imagine a situation where the NPLs of the reference group are almost perfectly aligned to each other, and therefore to their average value. This would certainly increase our confidence that the

average is indeed the value that should be uniquely identified as the iPL. In the opposite scenario, of a reference group with wildly fluctuating NPLs around the same average value as before, we would be quite uncertain which value to choose as an iPL.

However, if we only keep the average NPL for every PPP draw, the two scenarios are indistinguishable. We would like to retain this information – how spread are the NPLs inside the reference group at every draw. We therefore construct a simulated reference group by sampling one NPL for each PPP draw, and we consider this the distribution that describes the stochastic iPL. Without taking a random sample at each iteration one severely restrains the uncertainties involved in the determination of the iPL. This treatment is much similar to that of price data where the prices of each product for the CBN method are drawn from their distribution at each iteration. In that situation, the equivalent of keeping only averages of reference group's NPLs would be to draw 15 (as 15 are the initial reference group members) prices from the distribution of a particular commodity and then average them to get the price to use. This treatment would unwarrantedly reduce severely the uncertainty of the commodity prices. The same logic applies for the estimation of the iPL in our treatment.

Hence, it is the width of this second distribution which is the required input to propagate the underlying uncertainties of the DAD method. Observe further that the breadth of the 95% confidence interval of the sampled NPLs is more than 4 times larger than that of the averaged NPLs, indicating that the uncertainty introduced in the process of averaging is the decisive source of error behind the wide confidence intervals of the global poverty estimates.³⁶

In the lower sub-plot of Fig. 3 a similar graph this time for the CBN method is shown. Overall the comparison of the wider distributions' breadth would lead one to conclude that similar error terms for the CBN based global poverty rates are to be expected. The important difference though is that in this distribution all poverty lines calculated for all countries, with available data, are grouped together. Naturally those are not all applied on each and every country, as in the case of the DAD method. Rather, they are applied separately and only for the country for which they are constructed. To demonstrate the point we further show in the figure the distribution of the NPLs used specifically for Brazil (corresponding to the right axis).

The goal oriented tailoring of globally defined –yet locally estimated– NPLs, which is the characteristic of the CBN method, appears to make a decisive difference in the size of the error term for the global poverty rates comparisons between the two methods.

³⁵ There is one issue of representation with respect to the x axis in the upper sub-plot of Fig. 3. Each PPP draw during the Monte Carlo procedure constitutes an entirely different set of PPP rates, which is not directly comparable with any other PPP set, as each of these PPP sets represents a different constellation of relative currency values. As such the values on the x axis cannot strictly be 2011 PPP Dollars. The implication is that we have to use a convention and rely on the term “fuzzy” to characterize the dollars of the 2011 ICP round. This bears no consequence on the results.

³⁶ As reported in Table 2 in the next subsection.

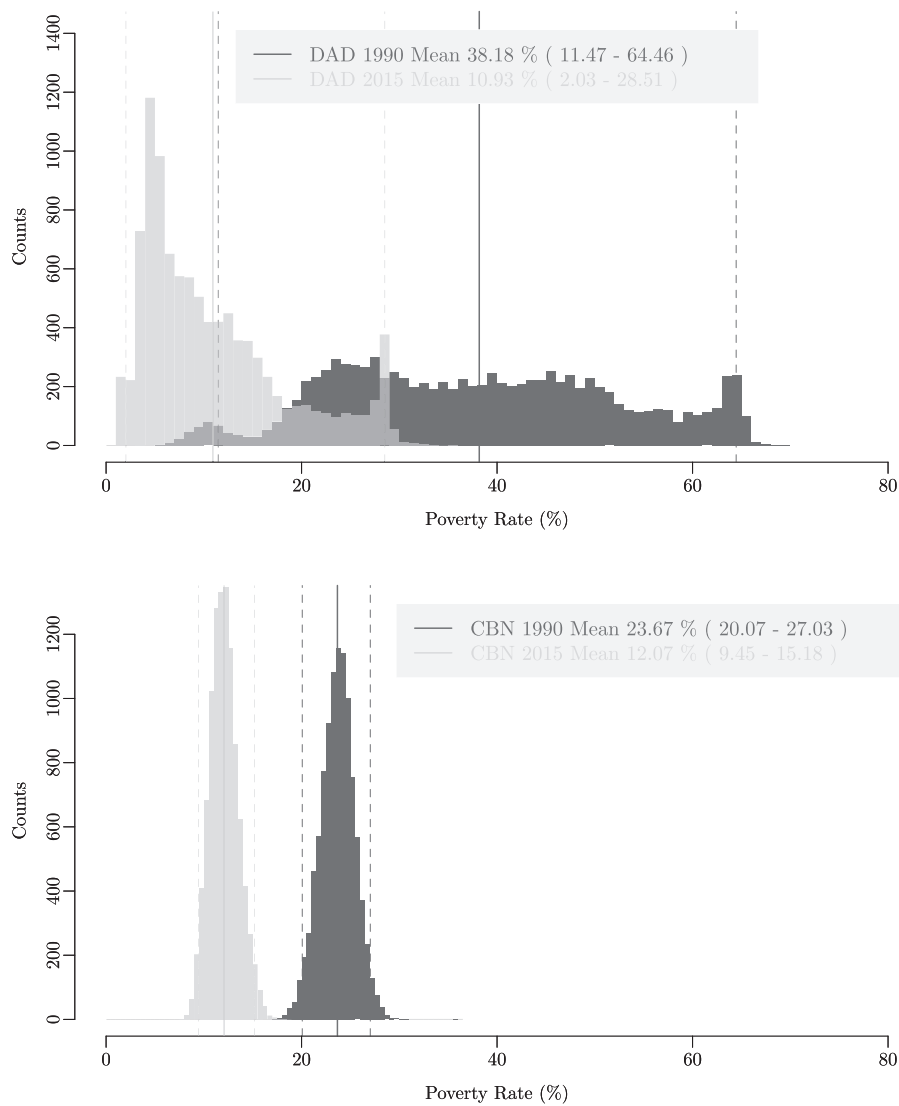


Fig. 4. Histograms of Global Poverty Rates of most complete implementations. DAD implementation 1 in the upper sub-plot, and CBN implementation 3 in the lower sub-plot. Each distribution is composed of 10000 points. The values of the 95% confidence interval of each distribution is shown in parenthesis, and its average value precedes it.

3.2. Testing MDG1

The various Monte Carlo implementations for the dollar-a-day method are shown on Table 2, and are numbered with IDs 1 to 7 in the first column. Columns 2–7 show the means and standard deviations for the iPL, and the global poverty rates estimates for 1990 and 2015. Column 8 shows the confidence level at which the MDG1 has been fulfilled, and column 9 provides the poverty reduction achieved between the benchmark years, both at 95% confidence level.³⁷ At the end of the table the official values of World Bank's PovcalNet are shown for reference (at the row labeled PCN in first column).

Implementation 1 corresponds to the configuration described in the sections above, and includes the uncertainty for the PPP exchange rates and the uncertainty from the definition of the iPL. This implementation identifies a poverty rate in 1990 of 38.18% with a standard deviation of

14.32 percentage points, and in 2015 10.93% with 7.32 percentage points of standard deviation. The confidence level for achieving MDG1 stands at 80%, and it is lower than the typical 95% or 90% confidence levels, which are typically used in social sciences. At the same time the poverty reduction between 1990 and 2015 at 95% confidence level stands at 5.12%. This implementation is the one shown in the upper sub-plot of Fig. 4 below, and the wide overlap between the 1990 and 2015 distributions of global poverty rates is evident.

The remaining implementations add various restrains and error considerations upon this basic configuration.

Implementation 2 takes a conservative step and excludes the selection of the two most extreme values (the minimum and the maximum values) at each iteration from the process of random sampling of the NPLs. This choice reduces the SD for iPL and the poverty rates for both years. The confidence level of achieving MDG1 stands at 82%, while the identified poverty reduction at 95% confidence level rises at 18.47%. Implementation 3 extends the exclusion of extreme values to the two minima and maxima of each iteration, and this pushes down the uncertainty in the estimates even further. The confidence level of achieving MDG1 rises to 86% and the identified poverty reduction increases to 31.17%.

³⁷ The achieved poverty reduction is calculated by estimating its maximum at 95% confidence level. The confidence level of MDG1 fulfillment is calculated by estimating the probability that a point from the 2015 distribution is lower than half of another point from the 1990 distribution, as required by the definition of MDG1.

Table 3
CBN Global Poverty estimates

Implem. ID	Mean 1990 (%)	SD 1990 (perc.points)	Mean 2015 (%)	SD 2015 (perc.points)	MDG1 Conf. Level (%)	Pov. Reduc. (%)
1	25.47	1.5	13.3	1.4	36	36.85
2	24.08	1.64	12.31	1.43	45	36.39
3	23.67	1.77	12.07	1.45	46	35.71
4	23.69	1.78	12.07	1.46	46	35.75
5	23.71	2.18	12.09	1.82	47	32.15
6	23.67	2.33	12.08	1.72	47	32.31
7	23.66	2.31	12.12	1.85	46	31.38
8	23.66	1.9	12.09	1.57	46	34.53
9	23.62	2.8	12.12	2.3	48	26.43
PCN	43.15*	-	11.62*	-	-	-

*: World Bank PovcalNet estimates from <http://iresearch.worldbank.org/PovcalNet/povDuplicateWB.aspx>, last visited on January 4th, 2019.

Notes on implementation IDs: (1) only MDER calculation uncertainty included; (2) previous specification and the uncertainty of commodity prices is added; (3 - which is the baseline specification) previous specification and the aforementioned standard deviations for the consumption share rates are used; (4) previous specification, and a normally distributed 10% error term is added when imputing for countries with missing data; (5) previous specification, but with a 20% error term when imputing for countries with missing data; (6) specification as in '3', plus a 5% normally distributed error term in the mean of the income or consumption distribution; (7) previous specification, but with a 10% error term for distributions; (8) specification as in '6', and a normally distributed 10% error term is added when imputing for countries with missing data; (9) previous specification, but with a 20% error term when imputing for countries with missing data, and a 10% normally distributed error term in the mean of the income or consumption distribution.

Implementation 4 is based on implementation 1 and a normally distributed 10% standard error is added to account for the uncertainty in the imputation of poverty for countries without data,³⁸ it is also assumed that on average the poverty rate in countries without data will be somehow higher than those with data (in the same geographical region) since they evidently show lack of capacity to monitor local poverty development. To simply account for this observation the regional average when imputing the poverty rate is increased by an arbitrary 20%. The configuration of ID4 has a small impact on the results, by slightly increasing the poverty rates for both years and their respective standard deviation relative to our reference results of ID1. ID5 adds a 5% normally distributed standard deviation to the mean of income or consumption distributions, having a marginal impact overall.

IDs 6 and 7 take the same conservative steps relative to ID4 as implementations 2 and 3 relative to ID1, by excluding the selection of the two and four, respectively, most extreme values. As expected, their results follow the same patterns relative to ID4, as do IDs 2 and 3 relative to ID1. The confidence level of obtaining MDG1 increase to 85 and 90%, while the poverty reduction at a 95% confidence level increase to about 27% and 40% respectively. The means of the poverty estimates are slightly reduced, while a more substantial impact is observed on the standard deviations of those estimates as shown on Table 2.

Finally, Table 3, shows global poverty estimates and probabilities of success for an alternative MDG1 appropriately formulated for the CBN approach. In all cases, the success of the goal of poverty reduction by 50% in the 1990–2015 period at 95% cannot be concluded. The mean values of the CBN based estimates appear rather steady among the various implementations, remaining around 23.5~25.5% for 1990 and 12~13.3% for 2015. In comparison to the DAD, the CBN approach shows smaller relative standard deviations slightly short of 10% on average and 20% at maximum, while DAD relative standard deviations are around 25~65%. Comparatively, this can be explained by the lack of uncertainty from averaging NPLs to produce an iPL, and no PPP uncertainty, since no PPP exchange rates are necessary in CBN. It is clear from these figures that the CBN approach delivers a more precise set of global poverty estimates in all cases investigated (see notes on Table 3

for the implementation details, while the structure of the table –beyond omitting the iPL's mean and SD– is identical to that of Table 2).³⁹ This point is made evident also from the comparison of the lower sub-plot in Fig. 4, which depicts the baseline CBN implementation (ID 3), to the DAD sub-plot, and the clearly narrower breadth of the CBN global poverty estimates relative to those of the DAD method.

4. Conclusions

We have hereby shown that, as advocated by the Commission on Global Poverty report, the uncertainties of the DAD global poverty estimates should find their way in the official reports, as the findings we presented above suggest that they are of considerable magnitude. Using our findings for the DAD method, the number of people living in conditions of absolute poverty in the world's developing countries in 1990 stands between 502 and 2823 million people, while in 2015 the interval is between 128 and 1788 million.⁴⁰ We further show that MDG1 was achieved at a confidence level of 80%, considerably lower than the typical benchmark of 95%. Moreover, we find that at 95% confidence level a 5.19% reduction in global poverty is identified by the DAD method (baseline scenario). However, given the considerable poverty reduction identified by the CBN method in the same period, we conclude that the inability of DAD to identify substantial poverty reduction in the MDG1 period says more about the uncertainties of the method per se, rather than about the actual evolution of global poverty.

We acknowledge that perhaps at the level of implementing legislation towards poverty reduction, lower confidence levels could be acceptable. At any rate, it is debatable how low the confidence level in testing MDG1 success should be, given the importance of how our world view may be influenced by receiving a positive or a negative message regarding the world's capacity to succeed in halving global poverty within 25 years. However, the discussion about which would be the desired or acceptable confidence level, has not taken place for the purpose of monitoring global poverty reduction goals. We believe that such dialogue is long overdue, also with respect to SDG1.1. Given the central role of poverty reduction –among others– in the development goals

³⁸ The World Bank uses the simple average of the regional poverty rate to impute the poverty rate for countries without enough data to estimate the poverty rate.

³⁹ The population coverage for DAD method is 88.6% and 82.6% for CBN in 1990, while for 2015 coverage rates are 88.3% and 81.7% respectively. Given the relative size of uncertainties the approx. 6.3% difference in population coverage between the two methods should not be viewed as worrisome.

⁴⁰ Based on the 95% confidence level of poverty rates in the results from implementation 1 of Table 2, and the population in developing countries for 1990 and 2015.

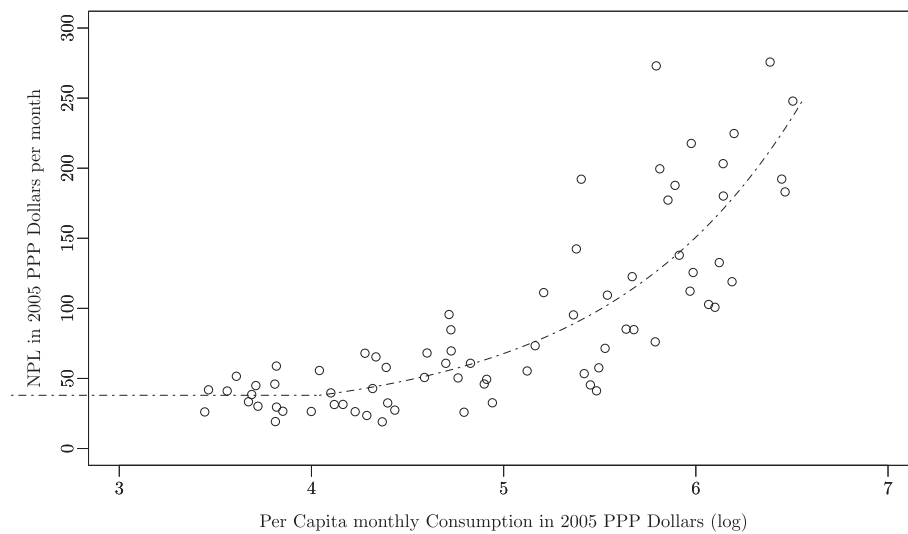


Fig. 5. The elbow fitting of the National Poverty Lines as a function of log consumption per capita for the 74 countries in the RCS data set, using the PovcalNet PPP exchange rates.

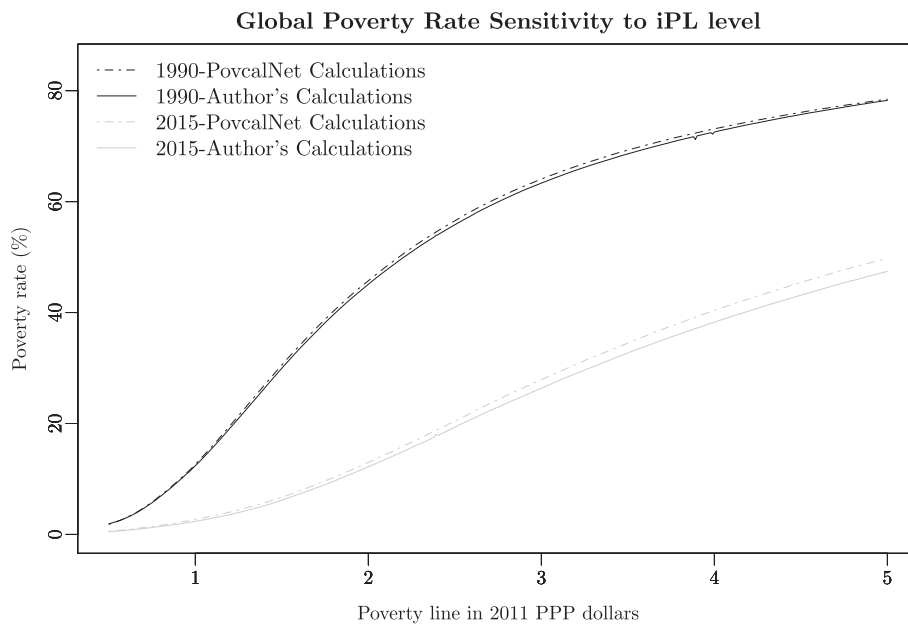


Fig. 6. Sensitivity of global poverty estimates to iPL values.

agenda, one could argue that not only each global poverty estimate should come with an uncertainty estimate, but also that long term goals like SDG1.1 should be stated with a concrete measure of success in terms of confidence levels.

The officially reported DAD point estimates show a remarkable 73% reduction of global absolute poverty between 1990 and 2015. This stands in stark contrast with our findings and in our view this shows that not only the comparison of point estimates of DAD results gives an overly optimistic picture of the fight against poverty, but also when the same method is used, and the underlying uncertainties are included, one is led to believe that poverty reduction is rather minimal at the typical level of statistical confidence. Given that the critical step producing the bulk of the uncertainty for the DAD method comes from the averaging of NPLs, it appears that updating the DAD method, and substituting the averaging with another procedure, may produce confidence levels much higher than the 80% we have identified here (e.g. [Klasen et al., 2016](#) investigates alternative approaches in estimating an iPL). Alternatively,

decomposing the NPLs in the [Ferreira et al. \(2015\)](#) dataset into absolute and relative components, as the theory behind [Ravallion et al. \(1991\)](#) stipulates, may give more accurate estimation of the implied iPL even if an averaging method would still be used. We find our results to be in line with the suggestion of [Allen \(2017\)](#) to use a CBN configuration as the basis for global poverty measurement.

Finally, the main policy implication of our conclusion is that the profile of the global poor and the distribution of poverty around the world might be misleading given the large uncertainty in the DAD global poverty estimates.⁴¹ It would be worth implementing the above method on a regional or national level to demonstrate this remark more directly.

⁴¹ We wish to thank an anonymous reviewer for this remark.

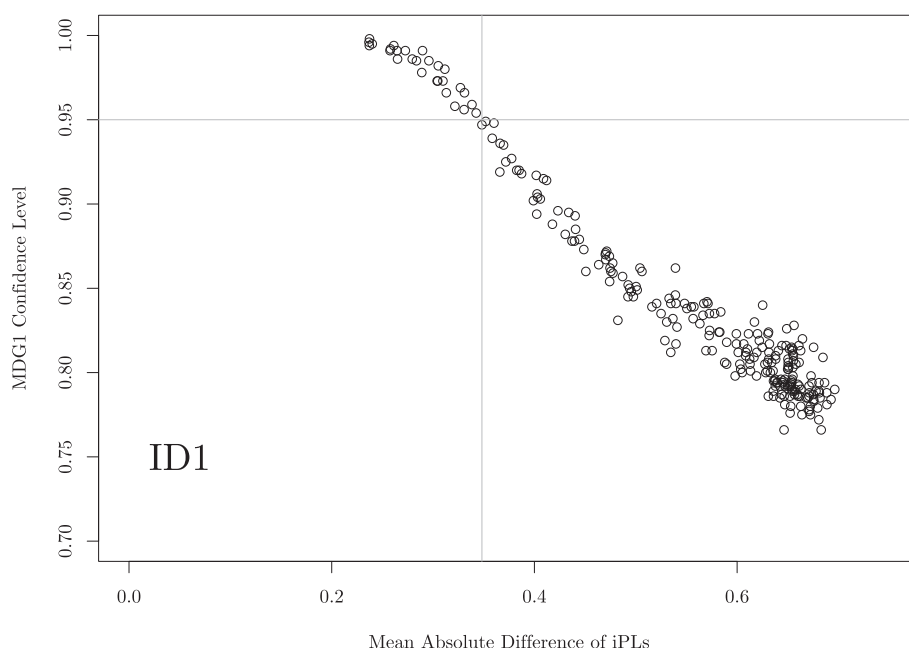


Fig. 7. Critical MAD values for the baseline (ID1) DAD scenario. Vertical gray lines at the critical value of 0.35.

CRedit authorship contribution statement

Michail Moatsos: Conceptualization, Data curation, Methodology, Formal analysis, Software, Validation, Writing - original draft. **Achillefs Lazopoulos:** Methodology, Formal analysis, Validation, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

A.1. Sensitivity of Global Poverty Estimates to iPL changes

Fig. 6 serves as a benchmark of how successful the replication of the PovcalNet methodology is in our implementation. Fig. 6 contains four cases. The upper two refer to the year 1990 for which we compare our own calculations with those taken directly from probing the PovcalNet service with a sequence of iPL values.⁴² The visible discrepancy between the two curves has an average deviation of 1.1% in relative, or 0.52 percentage points in absolute terms. Specifically at the standard 1.9\$/a-day iPL value the difference is 0.34 percentage points. This, rather small, discrepancy is most likely the result of two differences between our implementation and PovcalNet. For countries that do not have a survey during a benchmark year PovcalNet only uses a survey if it is available within a 2 years interval for global poverty estimates after 2010, and in a 3 years interval estimates for the 1981–2008 period (Ferreira et al., 2015, p.28). In our implementation we use a 4 years interval throughout, which allows us to include the 1993.5 household survey from India (in addition to the 1987.5 one), and be able to provide

increased coverage for the 2015 estimate due to lack of surveys around that year.⁴³

The other possible discrepancy is the difference in selecting the appropriate NAS growth statistic to be used in shifting the HHS mean to the benchmark year. Since PovcalNet does not report the exact NAS used at each individual case we cannot be certain in providing an exact replication of those choices. In light, however, of the uncertainties in the global poverty lines and rates presented in the results section we hold the position that the small deviations presented here do not undermine our results and conclusions, and that can be safely ignored.

The two lower lines, showing the comparison for 2015, are also very close for all iPL values, and at the 1.9\$ a day the difference is 0.72 percentage points and should be similarly attributed to the aforementioned causes (1.31 percentage points on average).⁴⁴ This gap might be worrisome and even puzzling, however since this means that our point estimates between 1990 and 2015 are further apart than the estimates of the World Bank we are tentatively introducing a bias that works against our findings. We therefore remain confident with respect to the validity of our implementations and our conclusions.

A.2. Comparing the 1990 and 2015 global poverty distributions

In estimating the confidence level of global poverty reduction between the estimates of global poverty for 1990 and 2015 one needs to take random samples from the common iPL distribution independently for the two years, and compare them to identify if MDG1 has obtained. We use this approach for our results in accordance with the stochastic nature of the quantities being compared. Our approach therefore implies that there exists a true, unknown value of the iPL, for which we don't know precisely how it evolves in time and therefore it needs to be estimated at each and every point in time effectively as a distribution due to the uncertainties involved in its estimation process. This

⁴² This is done with an updated version of the script initially provided by Dykstra et al. (2014). The values we used are from 0.5 up to 5 PPP dollars with a 0.01 step.

⁴³ This increased coverage comes at a price in terms of HHS extrapolations; see Section 2 for the methodological details.

⁴⁴ At the edges of the 95% confidence interval of the iPL, shown in Table 2 at row ID 4 and in the upper sub-plot of Fig. 3, the difference is 0.29 and 1.68 percentage points, with a mean of 0.91 percentage points within the interval.

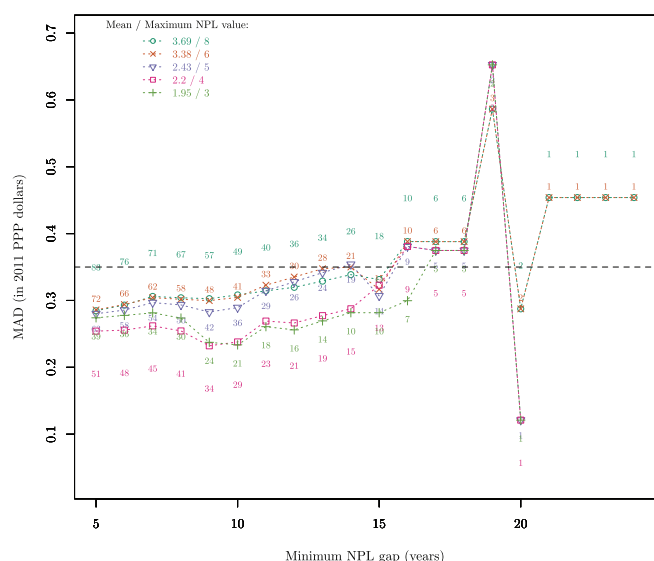


Fig. 8. MAD evolution of estimated NPLs as a function of the minimum gap of years among same country NPLs. Numbers near the lines represent the number of observations N at each point. Horizontal dashed lines at critical values 0.35 (ID1).

subsection discusses tentative alternatives.

It is perceivable that the findings presented in Table 2 may seem misleading because of the idea stating that there should be only one PPP draw underlying each poverty rates pair comparison between 1990 and 2015. This rationale could probably be explained by the frequency that the ICP project produces PPP rounds: about once every 6–10 years, and the subsequent repetition of the entire series of estimates by the World Bank following each new ICP round. Nevertheless, ICP rounds should have been yearly, but budgetary and other constraints –irrelevant to the point made here– simply do not allow for yearly ICP rounds, and extrapolation is used around the particular ICP benchmark year.⁴⁵ In this –more appropriate– case, a methodologically and data wise ideal 1990 round should be used for 1990 and likewise the 2015 ICP round for 2015. There is no reason why the PPP rates should be fixed at each ICP round, and our random sampling approach also simulates this lack of PPP rates with yearly frequency as they ought to be delivered. This point suggesting multiple PPP-benchmarks is corroborated by the updating procedure followed by the Maddison project, (Bolt, Inklaar, de Jong, & Zanden van, 2018b, p.11), and by the recommendations of Aten and Heston (2010, p.176) in averaging different PPP estimates between benchmark years.

Even more rigidly, some may find the approach of fully independent draws for the two benchmark years hard to agree with, as they might expect that even if one does not know the exact level of the iPL in 1990, whichever that level might be it must be exactly the same also for 2015 when the reduction rate between those years is estimated. If one accepts such a strict reading, such as the one discussed in the previous subsection, then the comparison between the two benchmark years would be as if error-less.

Although we consider that this entire notion of fully inter-yearly dependent iPLs builds on the idea that PPPs are error-less, we discuss here an alternative approach that can be used which avoids strict independence, and instead opts to produce critical values regarding how the iPLs from 1990 and 2015 are paired in the sampling procedure. For this purpose, we use the mean absolute difference (MAD) between the

Table 4

“Illustrative Checklist for Nonsampling Errors”, details added upon the original (Atkinson, 2016, p.52). Entries marked with the term “Robustness” are not part of the main treatment.

	Source of error	Accounted for in	Approach/Comment
1	Incomplete country coverage	Robustness	Add a 5% or 10% normally distributed error term on those regional averages.
2	Incomplete measurement of consumption/ measurement error	Robustness	Add a 10% or 20% normally distributed error term to all distributions.
3	Use of income in place of consumption	None	Consumption is arguably a better indicator of welfare than income, but not all countries have consumption HHS.
4	Population missing from sampling frame	None	E.g. those incarcerated.
5	Survey differential nonresponse	None	Under-reporting by the rich and under-representation of the poor.
6	Inaccurate or out-of-date population totals	None	-
7	Errors in the determination of the poverty line	Both	Various uncertainties (PAL, MDER, NPL averaging, PPP, etc, see text for details).
8	Standard error of PPP indexes to calculate baseline local currency poverty line	DAD	Estimated PPP standard errors and its impact on iPL determination. This is not applicable to CBN.
9	Surveys not comparable over time	None	E.g. due to methodological changes.
10	Extrapolation of out-of-date survey data	Both	A correction is applied based on the history of HHS and NAS growth rates following FEA.
11	Bias in domestic CPI to update local currency poverty line	CBN	This is encapsulated in the CBN approach.
12	Differential inflation for the poor	CBN	The consumption baskets used in CBN are CPIs for the extreme poor.
13	Rural/Urban and other geographical differences	Both	Rural/Urban split for the countries with available data, namely: China, India and Indonesia.
14	Use of equivalence scale in place of per capita calculation	None	Access to the HHS microdata is required for addressing this. On the required scale only World Bank researchers have the required access.

randomly selected iPLs for 1990 and 2015. This MAD is the condition that the 1990/2015 iPL pairs must satisfy in order to guarantee that the values of those lines are closely linked. Our baseline scenario (ID 1 in Table 2) requires a 0.35 MAD value (in 2011 PPP dollars) so that it supports the success of MDG1 at 95% confidence level (see Fig. 7 for the full curves).

Now we need to find empirical evidence against which to compare this value. This will allow us to indicate if this critical MAD value is too restrictive or too loose. If it is too restrictive (loose) it will mean that the actual MAD should be higher (lower), rendering the MDG1 unsuccessful (successful) at 95% confidence level.

We took stock of all the poverty rates published by WB based on NPLs, and then inverted them using the PovcalNet data to get the implicit NPL value (this is the method applied by World Bank Poverty and Inequality Team researchers Jolliffe & Prydz, 2016). Fig. 8 shows the results of this investigation. The NPLs we calculate this way are paired on a per country basis. The x axis indicates that NPLs at point t on x axis have at least t years of difference between them. Each line on the graph is produced by a different split of the data, depending on the maximum

⁴⁵ Extrapolation is the term used by Bolt, Inklaar, de Jong, and Zanden van (2018a) for this method of keeping the PPP terms fixed and extrapolating them locally in each country with the use of CPIs.

PPP dollar value allowed in the set as marked in the figure's legend.

The lower the maximum value allowed the lower the number of NPLs in the group. For the time span of 25 years that we are interested in, there are simply not enough data available. However, all the trajectories in the graph imply that we would get an MAD somewhere in the area between 0.35 and 0.45, close but above the critical MAD value. Therefore this supplementary approach does not seem to directly challenge our findings, but rather generally indicates a trajectory that would support them.

Comparing the critical MAD values with some reference values can also be addressed using the NPLs calculated here based on the cost of basic needs approach. Those poverty lines do express the same goal (nearly barebone subsistence with some frugal additional expenses), and their volatility with respect to their PPP dollar value is indicative of the consistency of the PPP assumption for the consumption needs of those living in conditions of extreme deprivation. The MAD for the CBN is 0.47 for the 25 years span of MDG1 (in line with the trajectories in Fig. 8), which would make the 0.35 value appear rather restrictive.⁴⁶

Regardless, since we don't have a set of actual official NPL values, but instead we estimate them indirectly, an alternative interpretation presents itself. Such an interpretation would state that the MAD as a function of years between NPLs (from the same country) is increasing in time because of differences in the underlying methodologies of the household surveys. But this should find satisfactory response in the fact that it is only the mean absolute difference that increases, while the mean difference remains very close to zero.

A.3. Omitted error sources

As it can be seen from Table 4 we are not treating all possible sources of non-sampling error identified by Atkinson (2016). Also, in some cases the treatment we operationalize is a crude and horizontal one, such as for HHS consumption (or income) measurement error, while in other cases we only provide treatment for one of the two methods that we compare (e.g. item 12 on the domestic CPI for the poor).

The important question, however, is whether the items that are not treated threaten the validity of our results and conclusions. If in the future a reliable method to estimate the missing non-sampling uncertainties emerges –as the Commission on Global Poverty strongly advocates that it should– the total non-sampling uncertainty of global poverty estimates can be obtained by adding these uncertainties to the one we arrive at, in quadrature, assuming that the missing uncertainties are reasonably uncorrelated with the ones provided in this work. This will certainly increase the total uncertainty, so our results here can be seen as a lower bound to the unknown total uncertainty. In case there are missing effects that are strongly correlated, their uncertainty should be added linearly to the present one, hence the total would increase further. In order for the total uncertainty to shrink, one has to assume that some of the missing effects are not only as significant, in magnitude, as the ones included, but also strongly anti-correlated to the latter, and that on a global scale. In fact, in order to invalidate the main result of the paper, one needs spectacular decrease in the uncertainty estimates, since as shown in the results section, the uncertainty on the success of DAD poverty rates is rather large, and the confidence level where MDG1 obtains is within a considerable distance from the more typical value of 95%.

⁴⁶ Do note that since the CBN NPLs adjust for changes in the underlying demographics of each country, (a small) part of this larger MAD should be attributed to this aspect of the CBN NPLs. The analysis found in Fig. 1 in Moatsos (2017) indicates that the median of the between benchmark years change that should be attributed to the change in the underlying demographics is 5.6%.

References

- Ahlwalia, M. S., Carter, N. G., & Chenery, H. B. (1979). Growth and poverty in developing countries. *Journal of Development Economics*, 6(3), 299–341.
- Alkire, S., & Foster, J. (2011). Counting and multidimensional poverty measurement. *Journal of Public Economics*, 95(7–8), 476–487.
- Allen, R. C. (2001). The Great Divergence in European Wages and Prices from the Middle Ages to the First World War. *Explorations in Economic History*, 38, 411–447.
- Allen, R.C. (2013). Poverty lines in history, theory and current international practice. University of Oxford, Department of Economics, Discussion Paper Series, (No. 685).
- Allen, R. C. (2017). Absolute poverty: When necessity displaces desire. *American Economic Review*, 107(12), 3690–3721.
- Appleton, S., Emwanu, T., Kagugube, J., & Muwonge, J. (1999). Changes in poverty in Uganda, 1992–1997. *Centre for the Study of African Economies, University of Oxford*, 99(22).
- Aten, B., & Heston, A. (2010). Use of Country Purchasing Power Parities for International Comparisons of Poverty Levels: Potential and Limitations. In S. Anand, P. Segal, & J. E. Stiglitz (Eds.), *Debates on the Measurement of Global Poverty*, chapter 6 (pp. 154–186). Oxford: Oxford University Press.
- Atkinson, A. B. (2016). *Monitoring Global Poverty: Report of the Commission on Global Poverty*. Washington, DC: The World Bank.
- Bailer-Jones, C. A. L. (2017). *Practical Bayesian inference: A primer for physical scientists*. Baten, J. & Blum, M. (2015). Clio Infra Global Height dataset.
- Bolt, J., Inklaar, R., de Jong, H., & Zanden van, J.L. (2018a). Maddison Project Database, version 2018. Maddison Project Working paper 10.
- Bolt, J., Inklaar, R., de Jong, H., & Zanden van, J.L. (2018b). Rebasings 'Maddison': new income comparisons and the shape of long-run economic development. Working Paper The Maddison Project, (January):67.
- Bowley, A.L. (1915). The nature and purpose of the measurement of social phenomena. P.S. KING and SON, LTD, London, UK.
- Chen, S., & Ravallion, M. (2001). How did the world's poorest fare in the 1990s? *Review of Income and Wealth*, 47(3), 283–300.
- Chen, S., & Ravallion, M. (2004). How Have the World's Poorest Fared since the Early 1980s? *The World Bank Research Observer*, 19(2), 141–169.
- Chen, S., & Ravallion, M. (2010). The Developing World is Poorer than We Thought, But No Less Successful in the Fight Against Poverty. *The Quarterly Journal of Economics*, 125(10), 1577–1625.
- Chib, S. (2001). Chapter 57 - Markov Chain Monte Carlo Methods: Computation and Inference. volume 5 of *Handbook of Econometrics*, pages 3569–3649. Elsevier.
- Deaton, A. (2005). Measuring poverty in a growing world (or measuring growth in poor world). *The Review of Economics and Statistics*, LXXXVII(Febuary):1–19.
- Deaton, A. (2010). Measuring Poverty in a Growing World (or Measuring Growth in a Poor World). In S. Anand, P. Segal, & J. E. Stiglitz (Eds.), *Debates on the Measurement of Global Poverty* (pp. 187–222). Oxford University Press.
- Deaton, A., & Dupriez, O. (2011). *Spatial price differences within large countries*. (July).
- Deaton, A.S. (2012). Calibrating measurement uncertainty in purchasing power parity exchange rates. Princeton University Working Paper, (April).
- Dhondge, S. and Minoiu, C. (2011). Global Poverty Estimates: A Sensitivity Analysis.
- Dykstra, S., Dykstra, B., and Sandefur, J. (2014). We Just Ran Twenty- Three Million Queries of the World Bank's Website.
- FAO (2001). Human energy requirements. Technical report, Food and Agriculture Organization of the United Nations, Rome.
- FAO (2001). Human energy requirements. Technical report, Food and Agriculture Organization of the United Nations, Rome.
- Ferreira, F.H.G., Chen, S., Dabalen, A., Jolliffe, D., Narayan, A., Prydz, E.B., Revenga, A., Sangraula, P., & Yoshida, N. (2015). A global count of the extreme poor in 2012: data issues, methodology and initial results. Policy Research Working Paper 7432. Washington, DC: The World Bank.
- Hausman, J. (2001). Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left. *The Journal of Economic Perspectives*, 15(4), 57–67.
- Jolliffe, D., & Prydz, E. B. (2016). Estimating international poverty lines from comparable national thresholds. *Journal of Economic Inequality*, 14(2), 185–198.
- Klasen, S., Krivobokova, T., Greb, F., Lahoti, R., Pasaribu, S. H., & Wiesenfarth, M. (2016). International income poverty measurement: Which way now? *Journal of Economic Inequality*, 14(2), 199–225.
- Moatsos, M. (2017). Global Absolute Poverty: Behind the Veil of Dollars. *The Journal of Globalization and Development*, 7(2).
- Moatsos, M. (2018). Scrapping PovcalNet's Distributions for Poverty Research. Munich Personal RePEc Archive MPRA, (89872).
- Moatsos, M. (2020a). Long run trails of poverty, 1925–2010. *Empirical Economics*.
- Moatsos, Michail (2020b). Global extreme poverty: Present and past since 1820. In Jan Luiten van Zanden, Marco Mira d'Ercole, Mikolaj Malinowski, & Auke Rijpma (Eds.), *How Was Life? Volume II*. Paris: OECD. In press https://www.oecd-ilibrary.org/sites/3d96efc5-en/1/3/9/index.html?itemId=/content/publication/3d96efc5-en&cs_p=2c2e680562193998e9d20ed6a45a9242&itemIdGO=oecd&itemContentType=book.
- R Core Team (2018). R: A Language and Environment for Statistical Computing.
- Rao, D.P., Rambaldi, A.N., Huynh, L.T., Doran, H.E., & Ganegodage, K.R. (2015). UQ International Comparisons Database: UQICD Version 2.1.2.
- Rao, D. S. P., Rambaldi, A., & Doran, H. (2010). Extrapolation of purchasing power parities using multiple benchmarks and auxiliary information: a new approach. *The Review of Income and Wealth*, 56(Special Issue 1).
- Ravallion, M., Chen, S., & Sangraula, P. (2009). Dollar a Day Revisited. *The World Bank Economic Review*, 23(2), 163–184.
- Ravallion, M., Datt, G., & van de Walle, D. (1991). Quantifying absolute poverty in the developing world. *Review of Income and Wealth*, 37(4), 345–361.

- Reddy, S. G., & Pogge, T. (2010). How not to count the poor. In S. Anand, P. Segal, & J. E. Stiglitz (Eds.), *Debates on the Measurement of Global Poverty* (pp. 42–51). Oxford University Press.
- Rowntree, B. S. (1901). *Poverty: A study of town life*. Macmillan.
- Schennach, S. M. (2020). Chapter 6 - Mismeasured and unobserved variables. In S. N. Durlauf, L. P. Hansen, J. J. Heckman, & R. L. Matzkin (Eds.), *Handbook of Econometrics, Volume 7A, volume 7 of Handbook of Econometrics* (pp. 487–565). Elsevier.
- Sen, A. (1985). *Commodities and Capabilities*. North-Holland, Amsterdam.
- Srinivasan, T. N. (2009). Irrelevance of the \$1-a-Day Poverty Line. In *Initiative for Policy Dialogue Working Paper Series*.
- Subramanian, S. (2015). Once more unto the breach. *Economic & Political Weekly*, L (45):35–40.
- The World Bank (2016). A cover note to the Report of the Commission on Global Poverty.
- Townsend, P. (1979). *Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living*.
- United Nations (2015a). *The Millennium Development Goals Report*. United Nations.
- United Nations (2015b). *World Population Prospects: The 2015 Revision, Key Findings and Advance Tables*. Working Paper No. ESA/P/WP.241.
- World Health Organization (1985). *Energy and protein requirements*. World Health Organization Technical Report Series, (724).
- Zanden van, J. L., Baten, J., Foldvari, P., & van Leeuwen, B. (2013). *The Changing Shape of Global Inequality 1820–2000; Exploring a New Dataset*. *Review of Income and Wealth*, 1–19.