Measure for Measure: Comparing Survey Based Estimates of Income and Consumption for Rural Households

Gero Carletto, Marco Tiberti, and Alberto Zezza

This paper uses a large database of surveys of household incomes to characterize income underreporting in household surveys in low- and middle-income countries. The objective is to document (a) the extent of this underreporting, and (b) whether and how it varies systematically with respondent, household, income, and survey design features. Drawing on rural household data from 20 developing and transition countries, and using consumption expenditure as a benchmark, results indicate that the observed income/consumption ratios are very small, being on average around 0.76. Results suggest that income underreporting is systematically associated with household and survey characteristics. In particular, the degree of underreporting is strongly associated with the income source, with agricultural income being the component suffering more than any other components from underreporting. The analysis also provides evidence supporting the well-established proposition that underreporting tends to increase with household welfare: richer households appear to underreport income more. Implications for survey design and for future research are drawn.

JEL Codes: C8, D1, I3, O12
Keywords: income, consumption expenditure, income specialization, measurement error, household surveys.

> *"The practical and conceptual difficulties of collecting good income data are severe enough to raise doubts about the value of trying"*
> A. Deaton (1997), p. 30.

Introduction: Investigating the Systematic Deviations between Income- and Consumption-Based Measures of Welfare

Measurement of household income in developing countries is notoriously fraught with problems, and a widely held view is that income is often heavily underreported

The World Bank Research Observer
© The Author(s) 2021. Published by Oxford University Press on behalf of the International Bank for Reconstruction and Development / THE WORLD BANK. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com
doi: 10.1093/wbro/lkab009 37:1–38

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

(Coulombe and McKay 1995; McKay 2000). These concerns notwithstanding, the are many reasons that make the quest for the collection of quality income data one worth pursuing.

First, while consumption expenditure is the most commonly used aggregate for poverty measurement, income is still preferred as a measure of welfare in some countries and regions. Traditionally, the majority of countries in Latin America base their official poverty figures on income (SEDLAC 2012). Countries in Eastern and Central Europe and the Balkans are also increasingly moving towards adopting income as a measure of welfare in order to meet some of the statistical requirements for full accession to the European Union (Atkinson and Marlier 2010; Krell et al. 2017). Second, measures of income are necessary to study the sectoral composition of the economy in microeconomic analyses, how households derive their livelihoods, and how productive different household assets and different economic activities are.

Consumption expenditure being the preferred metric for poverty measurement, the collection of consumption expenditure data has received, at least in developing countries, considerably more attention than the collection of income data. In some countries (Integrated) Household Budget Surveys (I-HBS), Living Standards Measurement Study (LSMS) surveys, and other similar surveys have collected very little, if any, income data. According to data from PovcalNet (the World Bank global poverty computational tool), in the Middle East and North Africa and in Sub-Saharan Africa, all countries use consumption to measure poverty. In South and East Asia and the Pacific 63 percent do, but in Europe and Central Asia the numbers are virtually the same (51 percent of countries use income) and in Latin America completely reversed (87 percent use income) (see supplementary online appendix fig. A1). In aggregate, of 1,171 surveys included in PovcalNet at the time of writing, 52 percent use consumption to measure poverty, which corresponds to 64 percent of the countries (some countries have more frequent surveys than others or feature more in this database for other reasons).

Against this backdrop, practical guidelines have been developed to assist researchers and analysts computing broadly comparable and theoretically consistent consumption aggregates and poverty measures from household surveys (Deaton and Zaidi 2002), but much less information is available *for low-income countries* in terms of looking at income data. The Luxembourg Income Study, the Canberra Group, and the Wye Group Handbook, the three major efforts in systematizing work on household income data, all share a bias towards working with high- and middle-income countries. The SEDLAC initiative covers some lower income countries, but is limited in geographic scope to Latin America and the Caribbean.

This paper is primarily concerned with the measurement of income in lower income countries,[1] and for that purpose it draws on the work of the Rural Income Generating Activities (RIGA) project, which has assembled a database of 45 household living standards surveys and created income aggregates in a methodologically

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

consistent manner, and with an explicit (albeit not exclusive) focus on low- and middle-income countries.

The objective of this paper is to provide a systematic assessment of the quality of survey-based income measures in low- and middle-income countries, and to perform a cross-country exploration on the key features of this deviation. Although the concern about income underreporting is not new, this paper contributes to the literature and practice of income data collection by considerably expanding the evidence base, and by linking income underreporting to the sources of income and to key survey design characteristics, in a way that has not to our knowledge been done systematically for low- and middle-income countries. To that end the paper uses survey-based measures of consumption as a benchmark, and looks systematically at how income and consumption measures differ based on (a) the composition of income; (b) household and individual characteristics; and (c) the basic characteristics of the survey instruments used in collecting both income and consumption data. In the process, implications for analysis and data collection are drawn.

The issue of the comparability of consumption and income aggregates is an exercise which carries operational relevance. The World Bank for instance faces the challenge of measuring "shared prosperity," one of its twin corporate goals (the other being the eradication of poverty). Shared prosperity can be defined as growth in the income of the bottom 40th percentile of the consumption/income distribution (Jolliffe et al. 2015). Indeed, the shared prosperity indicator for a given interval based on consumption is, in all likelihood, different to the one based on income. The data used in the first major report published by the World Bank on shared prosperity (covering at least two data points between 2003 and 2013), have similar features to those that characterize the PovcalNet database, with the same regional patterns observed: the near totality of Latin American countries uses income, many East Asian and African countries use consumption, and Europe and Central Asia lie somewhere in between. The same report recognizes that, in terms of empirical estimation of the shared prosperity indicator, "the difference between a shared prosperity measure based on consumption and one based on income is not trivial" (Jolliffe et al. 2015, 86). The conceptual and empirical issues of the comparability between countries, and in particular cases even within countries over time, of consumption and income-based poverty estimates provided by the World Bank are addressed in Ferreira et al. (2016). In particular, Ferreira et al. (2016) point out that, by definition, households can report zero income, while zero is not a feasible value for consumption, and this difference is reflected in income and consumption data distribution. One aspect that they lament but do not provide specific references for, is that lower income countries with a predominance of agricultural self-employment and little formal employment usually report poor quality data on income, while consumption data in richer countries are more prone to suffer bias from respondent fatigue. As a result of the fact that there are few observable patterns that are fixed in the relationship between poverty

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

and income, however, PovcalNet does not introduce any adjustment to the income or consumption distributions.

The structure of the paper is as follows. The next section offers a theoretical framework about the possible sources of bias leading to errors in consumption and income measurement in household surveys. That is followed by a review of key stylized facts about income measurement found in the literature, a discussion of the data and methods and a presentation of the results of the data analysis. The final section offers some concluding remarks and suggestions for future work.

Competing Sources of Income Measurement Error: Theoretical Considerations

Before delving into the data, it is useful to define ex ante expectations of the sources and features of income misreporting in surveys, based on theory and on evidence stemming from the empirical literature. While income measurement is a challenging enterprise regardless of the country or society in which it is carried out, the weight of the different challenges for data quality varies with income levels. In higher income countries (where much of the literature originates, see Moore et al. [2000] for a review; Brandolini [1999]), non-response is a big issue, as are the estimation of income from financial assets and estimates of pre- and after-tax income. In low-income countries, these issues are also relevant, but the challenges are much more related to survey populations working in largely informal sectors, who are often illiterate, do not keep accounting books or records of any kind, and are engaged in activities that are often seasonal, and for which markets are absent or very thin.

The first problem in assessing the quality of any measure is that of identifying a benchmark. Common options for studies of survey-based income data are the use national account estimates, or comparisons of income estimates across surveys or with administrative data, often tax records. In this study we use instead a measure of a different construct, consumption expenditure. This has the disadvantage of basing the comparison on two measures that we do not expect to be equal, but has the advantage of drawing comparisons between measures that were collected with the same data collection mode, by the same survey teams, and over the same units (households). It is also important to note that comparisons of income measures from different sources such as national accounts or different income surveys are also faced with differences in definitions and other problems of comparability (Atkinson and Micklewright 1983; Brandolini 1999; Moore et al. 2000).

We categorize four possible sources of discrepancy between the observed income and consumption survey measure: (a) Definitional discrepancy between income and consumption, the main difference between the two being savings (or dissavings);

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

(b) intentional or unintentional non-reporting or misreporting of specific survey items; (c) survey design features; and (d) issues with survey implementation.

Firstly, income and consumption are aggregates with different meanings and definitions. For any household at a single point in time income may be higher or lower than consumption as households may be either saving or dissaving. In aggregate however, and over a reasonably long period of time such as the 12 months most surveys take as a reference, the expectation is for savings to be positive, and hence income to be larger than consumption. In practice, the literature shows that especially in low-income countries, the ability to save can be very limited (De Magalhaes and Santaeulalia-Llopis 2018), although the evidence is mixed and context-dependent (see, for example, Gertler, Martinez, and Rubio-Codina 2012). Also, the marginal propensity to save is an increasing function of income so, in terms of definitions alone, we would expect the difference between income and consumption to be larger among better-off households.

Secondly, households may intentionally or unintentionally misreport income. Misreporting may take the form of households failing to report a source of income completely, or of households misreporting the income levels for those sources of income they do report. Both cases may be due to either unintentional or deliberate actions. Refusals represent a particular case of intentional misreporting, and are still relatively rare in low-income countries, particularly in rural areas. Even if refusal rates can be expected to be correlated with income, they would not affect the comparison of income and consumption within the same unit of analysis.

Thirdly, there may be survey design features that lead to a larger or smaller incidence of omissions or misreporting of income values. The way a question is worded, the length of a list of items, the greater or lesser reliance on proxy respondents, the length of the recall period, or the timing of a survey (e.g., closer to the harvest in an agrarian economy, or closer to the period of tax reporting) all affect the estimates of the same underlying constructs.

Finally, survey implementation is key. Identical survey instruments will yield very different results depending on the quality of implementation, depending on features such as the amount and quality of training, the amount and quality of supervision, or the emphasis on data quality control during fieldwork. Since our comparisons refer to measures collected within the same survey, we do not need to be concerned with the variability in implementation across surveys. There is however the possibility of variability within surveys if different parts of the survey receive different levels of attention in terms of training and quality control. We do not have information on survey implementation which we can systematically incorporate in the analysis, and this is therefore an area the paper does not venture into beyond this mention.

Of the four possible sources of discrepancy, we have already stated that based on the first source (definitions) one would expect income to be on average larger than consumption, if both were accurately measured. If that is true, measures of income

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

lower than consumption for the same population are a pretty strong indication of an underreporting of income, unless one believes that consumption is consistently overestimated. Of the fourth source (survey implementation), in the absence of a way to control for it, our working assumption will be that there is no a priori reason to expect that it would affect measures of income and consumption differently.

We therefore focus on the second and third sources, misreporting and survey design features. These are intimately tied, as survey practitioners ultimately make decisions on survey design aiming to reduce reporting errors to the extent possible via survey design choices.

Friedman et al. (2017) propose a classification of reporting errors in surveys and their discussion (which is focused on consumption) is relevant for measures of both income and consumption. They write of "recall error" as the tendency, increasing with the length of recall period and the low saliency of an event, for respondents to forget specific events. "Telescoping" is the opposite of recall error and reflects the tendency, particularly for short recall periods, for respondents to report events that in fact occurred outside of the recall period. They refer to the fact that respondents may report not based on actual events, but on heuristic response strategies such as multiplying some more or less regular transaction by a notion of the frequency of the event as "rule of thumb error" (what Moore et al. [2000] define as "error prone reconstruction strategies"). "Personal leave out error" occurs when an individual responding on behalf of other household members fails to report about an event that happened to another member simply because they may not be aware of that or just because it does not occur to them. Other types of errors listed by Friedman et al. (2017) include rounding, social desirability bias, and strategic responses. Social desirability has also been documented with respect to income sources when, for instance, program participants fail to report participation in public transfer programs (Moore et al. 2000). Strategic responses arise when respondents learn that a given response pattern will lead to more questions, and will therefore—to limit fatigue and interview time—start to intentionally misreport or omit specific events.

Typically, the reasoning that one can articulate around reporting error and survey design features for measures of income and consumption are not dissimilar. Both sets of information are often solicited in a two-stage process (first, the prompting of income [consumption] sources, followed by the reporting of the amounts received [spent]). The magnitude and direction of the impact of such a two-stage process may vary according to several factors linked to the survey design features, to the respondent characteristics, and to the type of consumption items, as well as the source of income. Friedman et al. (2017) and Moore et al. (2000) report examples of both for consumption and income respectively.

Just as for consumption, income may be affected by issues related to the length of the recall period, particularly for seasonal, infrequent, or non-salient activities through a memory decay effect or via "error prone reconstruction strategies"

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

(Moore et al. 2000). Results from empirical studies about consumption measurement errors are quite unambiguous in finding that a lower reported consumption is associated with a longer recall period (see, for example, Scott and Amenuvegbe 1991). Similarly, some respondents may not have any knowledge of particular sources of income earned by other household members. This has been reported for income as well as for asset ownership in settings as diverse as Malawi (Kilic and Moylan 2016) and the United States (Moore et al. 2000). And similar cases of asymmetric information are well known for those consumption items that are often purchased and consumed away from home (Farfán, Genoni, and Vakis 2017). As suggested by Friedman et al. (2017, 97) "the degree of inaccuracy is likely to increase with the number of adult household members and with the diversity of their activities outside the home."

The bottom line is therefore that similar measurement error generating processes are likely to be at work for both income and consumption measures, and that these processes vary in complex ways with the components of income and consumption, as well as with household and individual characteristics so that it is not possible to formulate an a priori set of hypotheses and devise an empirical strategy (including a randomized experiment) to pin them down with the level of precision one might be aiming at for less complex constructs. The approach in this paper is therefore one of drawing on the literature to articulate some stylized facts of income measurement error, to then use a large pool of broadly comparable household surveys to draw an informative comparison that can be linked to actionable survey design features.

Literature Review: Some Stylized Facts About Income Underreporting

Overall, the literature on income and consumption measurement through household surveys supports the notion that income is underreported. Evidence of income underreporting is available in studies of both high-income (Atkinson and Micklewright 1983; Branch 1994; Weinberg 2004) and low-income countries (Berry 1985; Alderman 1993; Coulombe and McKay 1995; McKay 2000). One goal of this paper is to revisit the first, basic stylized fact that comes from this literature: ***income is underreported in low-income countries living standard and income and expenditures household surveys***.

McKay (2000) looks at this question using household surveys collected in the 1980s and 1990s in eight developing and transition economies. His results point to income measures being generally (though not always) lower than corresponding consumption figures, and the degree of correlation between per capita measures to vary anywhere between 0.1 and 0.8 (see supplementary online appendix table A1).

Individual characteristics that the literature shows to be related to biases in survey response (and specifically income reporting) include gender, age, and educational

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

attainment. A study conducted in Malawi presents evidence of differences in reporting of household income between husbands and wives. According to the study, in the majority of cases husbands do not accurately estimate their wives' income, underestimating it 66 percent of the time (Fisher et al. 2010). Important determining factors identified by the study include: The husband's education (the higher it is, the more accurate his estimation of total household income); the wife's education (the higher it is, the more likely the husband is to underestimate her income); and the composition of the household (the more working-age members in the household, the more likely the underestimate).

McKay (2000) identifies the lack of written records for household businesses and self-employment activities as making it difficult for the respondent to account for the costs incurred in the activity. Vijverberg and Mead (2000) review the literature on the collection of data on household enterprises and conclude that "negative profits appear to be lower in countries with a higher general educational level" (p. 108), and link that to the possibility of more accurate income reporting by more numerate respondents. The second stylized fact from the literature that we will aim to empirically validate as a general pattern in our database is that ***the extent of income underreporting varies systematically with respondent and household member individual characteristics. Specifically, we will test if income underreporting is higher for female and illiterate respondents.***

A third domain of investigation regarding the sources and direction of measurement bias entails the household characteristics. As mentioned earlier, "personal leave out error" source of bias is likely to be positively correlated with the number of adult household members and with the range of their activities away from home. In addition, as noted, the level of welfare of the household may affect the direction of the bias in income reporting: for several reasons, wealthier households may have incentives for underreporting their income. This last set of issues relates to sampling, nonresponse, and intentional underreporting prevalence among richer households. According to Fisher et al. (2010), the demographic composition of the household may also have an impact on the extent of income underreporting, with a greater number of working age adults resulting in more substantial underreporting.

Higher income households are thought to be under sampled due to difficulties in accessing them (Deaton 2000). Work at the Inter-American Development Bank that analyzed 18 Latin American countries finds that the top 10 percent of the income sample have income similar to that of a well-educated professional, but not that of households with the highest income, from which the authors infer that the full spectrum of income is not being captured (Székely and Hilgert 1999).

Evidence that high-income households may be more likely to underreport their income is provided by Grootaert (1993), Deaton (2000), Anand and Segal (2008) Korinek et al. (2006), and Gottschalk and Smeeding (1999), among others. Korinek et al. (2006), utilizing U.S. data, find a strong, significant, and negative

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

correlation between income and survey compliance. Additional analysis of U.S. data (Berry 1985) demonstrates the problem with properly surveying higher income households, as incomes from capital investments, i.e., financial assets, rental income, and own businesses, are difficult to capture and form a larger share of income among the highest income households. Data from India support these conclusions and suggest that between 20 percent and 40 percent of the difference between survey income and national accounts could be attributed to undercounting among the very rich (Banerjee and Piketty 2003). Hence, we will assess whether ***the extent of income underreporting varies systematically with household characteristics, in particular whether richer and larger households tend to underreport income more.***

In addition to the observable individual and household characteristics that impact income reporting, the source of income itself also plays a large role in the quality of its reporting. Both the incidence and value of misreporting are likely to vary across different income sources.

Wage employment is generally seen as the easiest income source to report accurately on. By matching Census and Social Security data in Micronesia, Akee (2007) is able to examine earnings volatility and self-reporting errors for wage employment. The results suggest that the reporting error is centered around zero and is "mean-reverting", indicating that over time the error will return to its average value.

As noted by Deaton (1997, 29) for agriculture and family business "incomings and outgoing are likely to be confused" and the only way to try and measure income is to impose a complex accounting framework on the data collection. With regards to self-employment, specific issues with data collection include the lack of financial record-keeping, the blurring of lines between the household and the household enterprise, ownership across multiple households, and inputs purchased in one period that may be sold in another period (McKay 2000; Joshi et al. 2009; Mel et al. 2009). Grootaert (1993), analyzing panel data from Côte d'Ivoire (1985–1988), finds that the majority of household enterprises reported negative net income from their activities. A more recent study of Sri Lankan microenterprises also finds that firms underreport their revenues by around 30 percent, and that the use of account diaries (introduced by the study) can have a significant impact on the reported revenues and expenses (Mel et al. 2009). Joshi et al. (2009) find that out of their sample of Indian informal enterprises, less than 4 percent keep any sort of book of accounts, which leads to reporting inaccuracies for households that do not. Hence, the fourth stylized fact that comes from the literature is that ***income from own-account agriculture and other self-employed activities tends to suffer from underestimation and reporting error more than wage income***.

Besides individual, household, and income source characteristics, questionnaire design plays a major role in determining the outcome of any data collection effort. Lanjouw and Ravallion (1996) identify a change in survey design as a possible culprit

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

of an unrealistic poverty decline between the Ecuador 1994 and 1995 surveys: in 1994 the survey contained 73 expenditure items, while the 1995 list was increased to 94 items, including more non-food items.

A survey experiment in El Salvador administered a long and short version of the consumption module to similar households in an attempt to identify systematic differences (Jolliffe 2001). Using tests of stochastic dominance, the author finds that for 95 percent of the sample the long survey consumption is greater than that of the short survey consumption, resulting in the short survey identifying one million more people as severely poor than the long survey (Jolliffe 2001). Similar results are found by Pradhan (2001) in his analysis of Indonesia's Susenas Survey. A field experiment conducted by Beegle et al. (2012) in Tanzania tests eight alternative methods of measuring household consumption, including modules with 58, 17, and 11 food items. They find that median food expenditure increases by 38 percent as the list is expanded from 11 to 58 food items, given a seven-day recall period.

Another survey design experiment implemented in Tanzania to gather labor data finds similar results, with more questions resulting in greater accuracy. Bardasi et al. (2011) test two elements of survey design; first, the level of screening (one versus three screening questions) and second, the respondent (individual reporting versus proxy reporting). Authors report that the differences that occur are a result of the interaction between survey design and individual characteristics. They find that the lack of proper labor screening questions results in lower labor force participation of women and lower wage employment rates for both sexes; while the use of a proxy respondent produces lower labor force participation and lower agricultural employment for men, along with lower hours of work conducted by women.

Survey design issues were also tested with informal enterprises in India utilizing the 56$^{\text{th}}$ round of the National Sample Survey in 2000 and 2001. On average, researchers find that the profits from single, direct questions are lower than the derived profits based on multiple questions (Joshi et al. 2009). de Mel, McKenzie, and Woodruff (2009), based on results from two panel surveys of microenterprises conducted in Sri Lanka, find that reported profits are significantly higher than reported net revenues, defined as the sum of reported revenues minus expenses, and the correlation between profits and net revenues is very low. This result is similar to findings in Vijverberg (1991), who concludes that net revenue may be the best single measure, despite the lack of a benchmark to compare this to. Based on this literature, the fifth and last stylized fact is generically formulated as follows: ***questionnaire design matters and can reduce the extent of income underreporting. In particular, more prompting will result in greater measurement accuracy.***

While the focus of the papers in on income measurement, survey design issues obviously also affect our benchmark measure, consumption expenditure. Approaches to measuring consumption expenditures in household surveys differ in many aspects, from the methods of data collection (diary or recall), the length of the reference

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

period, the respondent's selection, the number of field visits, the number of items included in the recall list, the way of collecting food away from home information (Zezza et al. [2017] provide an overview of survey characteristics impacting food consumption estimates). The survey design feature receiving the most attention in the literature is arguably that of the sensitivity of measures to changes in the recall period (Deaton and Grosh 1998; Deaton 2005; Beegle et al. 2012; Backiny-Yetna et al. 2017; Gibson 2003; Scott and Amenuvegbe 1991). Specifically, Beegle et al. (2012), Backiny-Yetna et al. (2017), and Battistin et al. (2020) provide experimental evidence that the design and implementation of survey instruments for collecting food consumption has substantial measurement impacts.

Smith et al. (2014), reviewing 100 Household Consumption and Expenditure Surveys, find that 56 surveys used exclusively interview methods, and 26 surveys used different recall periods depending on the source of acquisition or the frequency of the purchase. Thirty surveys only used one recall period, 13 have a seven-day recall period, four used a 14-day recall period, two used a one-month recall period, five used the "usual month" or "usual week" approach, and the rest used a different recall period. Only recently U.N. guidelines on food data collection for low- and middle-income countries (United Nations 2018) converged on seven-day recall as the recommended approach for food consumption.

Besides food consumption, Ferreira et al. (2016) highlight the heterogeneity in how consumption for specific non-food items is measured, due to questionnaire differences. In particular they raise the issue of estimating the rental value of housing and the use-value of durable goods. The heterogeneity in the definition of the welfare aggregate in such items typically reflects the heterogeneity in questionnaire design, making comparison over time and across countries challenging (see, e.g., Jolliffe 2001).

In the empirical section we will therefore also control for the length of the recall period in the different modules of relevance. That does not resolve the issues with using consumption as a benchmark, but to the extent that measurement error in consumption is not expected to be a function of the income composition, this is less of a problem for the main thrust of our analysis, which is related to the drivers of income underreporting, and particularly the difference in patterns across income sources.

Data and Methods

The Dataset

The RIGA database is constructed from a pool of several dozen Living Standards Measurement Studies (LSMS) and other multi-purpose household surveys made available by the World Bank through a joint project with the Food and Agriculture Organization of the United Nations (FAO).[2] From this pool of possible surveys, the choice of

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

particular countries was guided by the desire to ensure geographic coverage across the three principal developing regions (Asia, Africa, and Latin America) and transition countries in Eastern Europe, as well as adequate quality and sufficient comparability in codification and nomenclatures. Furthermore, an effort was made to include a number of International Development Association (IDA)[3] countries as these represent developing countries with higher levels of poverty and are therefore of particular interest to the development and poverty reduction debate. Recently, the RIGA database has been gaining a disproportionate African focus due to the addition of several recent surveys from the LSMS-ISA project.[4] The list of 20 surveys included in the analysis is provided in table 1.

The construction of income aggregates that are comparable across countries is the principal output of the RIGA database. The database only includes surveys where income can be calculated based on extensive survey modules on the revenues and benefits from all the income sources, thus excluding surveys that prompt respondents directly about their incomes or that have a very limited set of questions on costs and revenues from a wide range of possible activities.

The definition of income applied in the RIGA methodology closely follows the definition given by the International Labour Organization (ILO) in the 2003 *Resolution concerning household income and expenditure statistics*. Income receipts are considered those receipts recurring regularly, contributing to *current* economic welfare, and not arising from a reduction in net worth. Operating costs are subtracted from revenues, so that income is net of costs.[5] In the RIGA database, income sources are grouped as follows: Wage income; self-employment (from all non-farm household enterprises); crop production; livestock; transfers; other sources (gross non-labor income from farm land rental, non-farm real estate rental, rental of owned assets, and other miscellaneous sources not specified in the questionnaire).

Of particular relevance to the analysis in this paper, is the categorization into *agricultural income* (from self-employment in crop and livestock activities), *wage activities* (comprising both agricultural and non-agricultural wage), and *non-agricultural self-employment*. This leaves transfer and other income as separate categories which will receive less attention in this paper.

Households are also classified according to their degree of specialization and diversification by using two different thresholds, and defining a household as specialized if it receives more than 50 or 75 percent of its income from a single source, and diversified if no single source is greater than that amount. These thresholds are arbitrary and other definitions of diversification and specialization are possible. The extent of diversification, clearly affected by the choice of the threshold, is around 10 percent or less in all cases when using the 50 percent definition and climbs to around 90 percent when using the 75 percent definition. The broad patterns, however, do not change with the choice of the threshold.[6]

The consumption aggregates included in our dataset are generally disseminated together with the raw data and are mainly computed by national statistical offices

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Table 1. List and Survey Characteristics of Analyzed Surveys

					Income		
Country	Name of Survey	Year of Survey	Survey	Consumption	Own-farm	Wage employment	Self-employment
Panama	Encuesta de Niveles de Vida	2002–2003	Panama 2003	Usual month	Last 12 months	Last month	Last 12 months
Bulgaria	Integrated Household Survey	2001	Bulgaria 2001	Last month	Last 12 months	Last month	Last month
Albania	Living Standards Measurement Survey	2005	Albania 2005	Diary	Other	months/weeks/ hours	Last month
Nigeria	Nigeria General Household Survey –Panel	2010–11	Nigeria 2010	7 days	2 visits	months/weeks/ hours	Last month
Guatemala	Encuesta de Condiciones de Vida	2000	Guatemala 2000	14 days	Last 12 months	Last month	Other
Bolivia	Encuesta de Hogares	2005	Bolivia 2005	Last month	Last 12 months	months/weeks/ hours	Other
Tanzania	Tanzania National Panel Survey	2012–13	Tanzania 2012	7 days	Last cropping season	months/weeks/ hours	Other
Nicaragua	Encuesta Nacional de Hogares Sobre Medición de Niveles de Vida	2001	Nicaragua 2001	14 days	Last 12 months	months/weeks/ hours	14 days
Pakistan	Integrated Household Survey	2001–2002	Pakistan 2001	14 days	2 visits	Last month	Last 12 months
Vietnam	Living Standards Survey	1997–1998	Vietnam 1998	Usual month*	Last 12 months	7 days	Last 12 months
Kenya	Integrated Household Budget Survey	2004–2005	Kenya 2006	7 days	Last 12 months	Last month	Last month
Cambodia	Household Socio-Economic Survey	2003–2004	Cambodia 2005	Diary	2 visits	Last month	Other
Tajikistan	Living Standards Survey	2003	Tajikistan 2003	7 days	Last 12 months	Last month	14 days
Ghana	Ghana Living Standards Survey	1998	Ghana 1998	Diary	Last 12 months	Last month	Last 12 months
Nepal	Living Standards Survey II	2003–2004	Nepal 2003	Usual month	Last cropping season	months/weeks/ hours	Last 12 months

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Table 1. Continued

					Income		
Country	Name of Survey	Year of Survey	Survey	Consumption	Own-farm	Wage employment	Self-employment
Bangladesh	Household Income-Expenditure Survey	2000	Bangladesh 2000	Diary	Last 12 months	months/weeks/ hours	Last 12 months
Madagascar	Enquête Permanente Auprès Des Ménages	2000–2001	Madagascar 2001	7 days	Last cropping season	months/weeks/ hours	Last month
Niger	National Survey on Household Living Conditions and Agriculture	2011	Niger 2011	7 days	2 visits	months/weeks/ hours	Last month
Malawi	Third Integrated Household Survey	2010–11	Malawi 2011	7 days	Last cropping season	months/weeks/ hours	Last month
Malawi	Integrated Household Survey 2	2004–2005	Malawi 2004	7 days	Last cropping season	months/weeks/ hours	Last month

Source: Own tabulation based on selected surveys included in the RIGA dataset.
Note: * Vietnam used a variant of the usual month approach with more prompting about the number of months during which purchases occurred and the frequency of purchase.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

or their advisors. While the broad methodology is common, based on guidelines expressed in Deaton and Zaidi (2002), they are most likely not as consistently comparable as the income aggregates, which were computed by a small team under strict common guidance. We therefore acknowledge that in drawing comparisons between income and consumption we cannot control for cross-country differences in consumption aggregate methodology.

Moreover, the sample consists of surveys using different methods (recall or diary) and different recall periods to collect information on both consumption and income. While asserting that the choice of the method and the recall period matter significantly for resulting estimates of consumption and income is stating the obvious,[7] it is very hard to form a priori expectations on the magnitude and impact of each choice. This is particularly true for the sample of surveys used in this paper, which present a variety of approaches to measuring both consumption and income (see table 1).[8]

It is not straightforward discerning a pattern or elaborating a typology of survey design approaches clearly affecting the consumption/income measurement discrepancy. Indeed, besides the length of the recall period, several other survey design characteristics may affect this discrepancy, such as the number of consumed items listed in the consumption module or the number of questions about agricultural inputs in the own-farm module, many of these factors being very difficult to control for in a cross-country analysis like this. However, much of our analysis in what follows is also organized around how differences between income and consumption vary with the income specialization categories described above. Errors in reporting surely affect both consumption and income, but we do not expect measurement error in consumption to be a function of the income specialization category.

Moreover, in the pooled multivariate analysis described below, we control for the length of the recall period/survey design (recall or diary) in the wage, self-employment, own-farm, and consumption modules. With these caveats, it is fair—in order to simplify the discussion—to use consumption as a benchmark against which to assess the extent of measurement error in the income aggregates.

Methodology

In the next section we use information from this dataset to gauge whether the data support the hypotheses developed earlier in the section on theoretical considerations. For each of the stylized facts we present a descriptive analysis by country largely based on cross-tabulations, which provides a prima facie overview of the association between underreporting and some key variables. More importantly, we also perform a meta-analysis type test across the 20 surveys in order to aggregate the results for the whole sample.[9] We then combine that with multivariate regression analysis to explore how robust the bivariate relationships are to the simultaneous introduction of a larger set of control variables.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Specifically, we estimate 20 individual country regressions of the following relationship:

$$D = \alpha + \beta_1 HH\_CHARS + \beta_2 ASSETS + \beta_3 SPECIAL + \varepsilon \quad (1)$$

The dependent variable D is the difference between income (Y) and consumption (C) expressed as a fraction of C. *HH_CHARS* is a vector of household characteristics that includes sex, age, education, and marital status of the household head, and the demographic composition of the household (number of working age adults and of children below age 15). *ASSETS* is, following Filmer and Pritchett (2001), a principal component index of household assets intended to proxy the level of household wealth, while being measured independently from income and consumption in the survey.[10] *SPECIAL* is a vector of mutually exclusive dummies related to whether the household is diversified, or specialized in agricultural, non-agricultural self-employment, or transfers (non-agricultural wage specialization being the reference category). The regressions also include an independently, identically distributed error term, ε. Household subscripts have been omitted from the notation for simplicity.

Moreover, we estimate the following pooled regression with microdata from the 20 surveys previously analyzed, by also including key survey characteristics among the explanatory variables:

$$\begin{aligned} \mathrm{D} = \alpha &+ \beta_1 \mathrm{HH\_CHARS} + \beta_2 \mathrm{ASSETS} + \beta_3 \mathrm{SPECIAL} + \beta_4 \mathrm{CONS} + \beta_5 \mathrm{WAGE} \\ &+ \beta_6 \mathrm{SELFEMPL} + \beta_7 \mathrm{OWNFARM} + \varepsilon \quad (2) \end{aligned}$$

where *CONS*, *WAGE*, *SELFEMPL*, *OWNFARM* are vectors of recall period for the consumption, wage, self-employment, and own-farm modules, respectively. In this specification, the standard errors are clustered at country level. We use this pooled regression mostly to ensure that the findings and generalizations coming from the descriptive and multivariate country regressions are robust to the introduction of survey design controls, but avoid reading too much in the sign of the coefficients on survey design features, as these are based on too few observations and may to some extent be capturing county fixed effects.

How Large is the Extent of Income Underreporting, and What Drives It?

Stylized Fact 1—Income is Generally Underreported in Developing Countries Living Standard and Income and Expenditures Household Surveys

Table 2 reports some basic statistics on how income and consumption measures relate in the surveys we analyze. In the vast majority of cases income appears to be lower than consumption, often by a large margin.[11] As observed by Deaton (1997, 30) the

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Table 2. Comparing Estimates of Income and Consumption Per Capita

Survey	Ratio of per capita Income to per capita Consumption	Difference between per capita Income and per capita Consumption normalized by Consumption		Mean per capita Monthly Savings (Income minus Consumption)*	Correlation between per capita Monthly Income and Consumption	Correlation between per capita Monthly Income and Consumption (Logs)
		Median	Mean			
Panama 2003	0.775	−0.335	−0.225	−73.518	0.618	0.617
Bulgaria 2001	0.839	−0.228	−0.161	−82.045	0.191	0.413
Albania 2005	1.114	−0.100	0.114	14.638	0.429	0.494
Nigeria 2010	0.609	−0.653	−0.391	−38.462	0.285	0.419
Guatemala 2000	0.725	−0.369	−0.275	−66.188	0.568	0.502
Bolivia 2005	1.083	−0.149	0.083	6.269	0.639	0.630
Tanzania 2012	0.690	−0.471	−0.310	−31.179	0.436	0.531
Nicaragua 2001	0.926	−0.214	−0.074	−20.840	0.306	0.527
Pakistan 2001	0.695	−0.261	−0.305	−28.144	0.347	0.392
Vietnam 1998	1.489	0.071	0.489	29.292	0.305	0.452
Kenya 2006	0.636	−0.557	−0.364	−36.516	0.497	0.486
Cambodia 2005	0.427	−0.720	−0.573	−61.973	0.242	0.361
Tajikistan 2003	0.369	−0.677	−0.631	−35.778	0.417	0.397
Ghana 1998	0.753	−0.392	−0.247	−51.045	0.283	0.472
Nepal 2003	0.894	−0.258	−0.106	−15.332	0.475	0.472
Bangladesh 2000	0.796	−0.206	−0.204	−14.456	0.482	0.526
Madagascar 2001	1.117	−0.141	0.117	−7.896	0.466	0.586
Niger 2011	0.234	−0.815	−0.766	−114.851	0.329	0.309
Malawi 2011	0.436	−0.646	−0.564	−30.455	0.536	0.534
Malawi 2004	0.526	−0.647	−0.474	−21.131	0.380	0.449
Mean	0.757	−0.388	−0.243	−33.981	0.412	0.478
Min	0.234	−0.815	−0.766	−114.851	0.191	0.309
Max	1.489	0.071	0.489	29.292	0.639	0.630

Source: Own calculations based on selected surveys included in the RIGA dataset.
Note:* Values converted to 2011 US$

large, systematic extent of household dissavings these figures imply is most likely an indication of the fact that income in household surveys is grossly underestimated, even more so as we also know consumption expenditure to be underestimated. Of the 20 surveys in the table only four (Albania, Bolivia, Vietnam, and Madagascar) have average income greater than consumption.[12] The data, it appears, overwhelmingly support the general validity of stylized fact 1.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

The breadth of our data also provides an opportunity to attempt some loose generalizations regarding the magnitude of income underreporting. When comparing survey sample averages, the measure of income is only about 76 percent of consumption, with a very large span of variation with income being anywhere from 23 percent (Niger) to 149 percent (Vietnam) of consumption (see table 2). Our preferred indicator for capturing the difference between income and consumption is however the median difference between the two per capita measures, expressed as a share of consumption per capita. The cross-country average for this indicator is −39 percent, with only Vietnam showing a positive value. The observed correlations between measures of income and consumption are also pretty low, with the average correlation coefficients being 0.41. The observed range of the correlations is between 0.19 (Bulgaria) and 0.64 (Panama) for per capita income and consumption.

If we assume that on aggregate savings are likely to be positive (and we have argued that this is a safe assumption), these can be interpreted as lower bound estimates of the degree of income underreporting in our sample of countries. The magnitude of the observed differences could hardly be more striking, and point to the need to better understand the patterns of underreporting in order to devise ways to collect better income data in future surveys. In order to do that, we will now go beyond these average figures to look into how the extent of underestimation may vary systematically with individual and household characteristics (including welfare) as well as with sources of income. We reiterate here the caveats made earlier on the presence of measurement error also on the consumption side, and on the fact that treating consumption as a benchmark of income is an expedient to simplify our discussion, but one that we do not expect to impinge on the key messages coming out of the analysis.

Stylized Fact 2—The Extent of Underreporting Varies Systematically with Respondent and Household Member Individual Characteristics

In particular, we want to test whether the extent of underreporting depends on the education and gender of the respondent. One caveat to this part of our analysis is that the income questions in our surveys are often answered by different respondents, with different rules applied in different surveys. Also, in most of the datasets we do not have the respondents' information for some or all sections of the surveys. For these reasons, we conduct our analysis on the characteristics of the household head, making the implicit heroic assumption that the household head is the most likely respondent. The results on this particular point are therefore to be interpreted with special care.

Literacy. The literacy or numeracy of respondents are thought to have an impact on the accuracy with which income is reported (Vijverberg and Mead 2000). Both descriptive statistics (table 3) and multivariate analysis (tables 4 and 5) give inconclusive results. The literacy coefficient is not significantly different from zero in the pooled

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Table 3. Statistical Differences in Means

	Household head			Household head			Asset quintiles			Specialization					
(Y-C)/C	Male	Female	t-test	Literate	Illiterate	t-test	Q1	Q5	t-test	Own-Farm (1)	Wage (2)	Self-employed (3)	t-test (1)-(2)	t-test (1)-(3)	t-test (2)-(3)
Panama 2003	−0.159	−0.294	***	−0.188	−0.176		−0.162	−0.210		−0.508	0.050	−0.295	***	***	***
Bulgaria 2001	−0.178	−0.105		−0.047	−0.165		−0.078	−0.236	*	−0.293	0.023	−1.018	***	***	***
Albania 2005	0.083	−0.010		0.084	−0.019		−0.022	0.235	***	−0.254	0.387	1.421	***	***	***
Nigeria 2010	−0.371	−0.528	**	−0.354	−0.437	*	−0.541	−0.128	***	−0.620	0.257	−0.208	***	***	***
Guatemala 2000	−0.268	−0.310	*	−0.270	−0.299		−0.271	−0.327	**	−0.486	−0.155	−0.067	***	***	***
Bolivia 2005	0.120	0.055		0.142	0.106		0.134	0.166		−0.080	0.221	0.115	***	***	*
Tanzania 2011	−0.282	−0.359	**	−0.264	−0.384	***	−0.314	−0.226	**	−0.410	−0.097	0.032	***	***	**
Nicaragua 2001	−0.082	−0.060		−0.122	−0.026	***	−0.012	−0.142	**	−0.308	0.063	0.095	***	***	
Pakistan 2001	−0.300	−0.197	***	−0.251	−0.315	***	−0.337	−0.244	***	−0.647	−0.167	0.052	***	***	***
Vietnam 1998	0.448	0.616	***	0.498	0.374		0.285	0.623	***	−0.055	0.164	2.178	***	***	***
Kenya 2006	−0.363	−0.380		−0.334	−0.410	***	−0.392	−0.296	**	−0.616	−0.030	−0.180	***	***	**
Cambodia 2005	−0.583	−0.540	**	−0.568	−0.585		−0.625	−0.500	***	−0.724	−0.108	−0.497	***	***	***
Tajikistan 2003	−0.629	−0.630		−0.629	−0.635		−0.592	−0.679	***	−0.613	−0.674	−0.580	***		
Ghana 1998	−0.160	−0.355	***	−0.235	−0.211		−0.077	−0.249	**	−0.212	−0.166	−0.196			
Nepal 2003	−0.114	−0.059	*	−0.149	−0.075	***	−0.041	−0.185	***	−0.390	0.134	−0.139	***	***	***
Bangladesh 2000	−0.217	−0.073	***	−0.214	−0.199		−0.163	−0.236	**	−0.439	−0.191	−0.135	***	***	**
Madagascar 2001	0.137	0.263	*	0.008	0.166		0.151	−0.013	***	0.121	0.070	0.523		***	***
Niger 2011	−0.743	−0.783	***	−0.745	−0.757		−0.720	−0.754		−0.760	−0.616	−0.713	***	***	***
Malawi 2011	−0.528	−0.586	***	−0.530	−0.564	***	−0.557	−0.535		−0.588	−0.392	−0.435	***	***	*
Malawi 2004	−0.463	−0.564	***	−0.475	−0.507	**	−0.511	−0.459	***	−0.527	−0.340	−0.297	***	***	
Overall sample	−0.044	−0.359	***	−0.312	−0.033	***	−0.025	−0.283	***	−0.420	−0.078	-0.017	***	***	***

Source: Own calculations based on selected surveys included in the RIGA dataset.
Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Table 4. Regression Results By Country

(Y-C)/C	Panama 2003	Bulgaria 2001	Albania 2005	Nigeria 2010	Guatemala 2000	Bolivia 2005	Tanzania 2012	Nicaragua 2001	Pakistan 2001	Vietnam 1998
female headed hh	−0.129***	0.013	−0.067	−0.033	−0.051	−0.026	−0.043	−0.167*	0.150***	0.124*
	(0.042)	(0.076)	(0.106)	(0.088)	(0.035)	(0.081)	(0.035)	(0.085)	(0.022)	(0.063)
single head of hh	0.085**	−0.132*	−0.031	−0.148*	0.025	−0.023	−0.022	0.165**	−0.067***	0.005
	(0.036)	(0.076)	(0.087)	(0.077)	(0.029)	(0.075)	(0.043)	(0.072)	(0.019)	(0.088)
age head of hh	−0.001	0.010	0.045***	0.004	−0.002	0.025***	0.007*	0.012**	−0.011***	0.006
	(0.005)	(0.009)	(0.009)	(0.010)	(0.002)	(0.008)	(0.004)	(0.006)	(0.002)	(0.013)
agehead^2	0.000	−0.000	−0.000***	−0.000	0.000**	−0.000***	−0.000**	−0.000**	0.000***	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
literacy	−0.044	0.092	−0.042	−0.170***	−0.028	0.094	0.028	−0.092***	0.023	−0.063
	(0.046)	(0.131)	(0.076)	(0.047)	(0.019)	(0.103)	(0.025)	(0.031)	(0.014)	(0.069)
Number of people age >14 and <60	0.050***	−0.085***	−0.035**	0.020	0.051***	−0.012	−0.021***	0.010	0.011***	−0.022
	(0.011)	(0.018)	(0.017)	(0.014)	(0.010)	(0.023)	(0.008)	(0.011)	(0.004)	(0.017)
Number of people age < 14	−0.015*	−0.026	−0.076***	−0.007	−0.003	−0.017	−0.009**	−0.031***	−0.004	−0.072***
	(0.008)	(0.017)	(0.015)	(0.012)	(0.004)	(0.014)	(0.004)	(0.009)	(0.003)	(0.018)
wealth index	−0.067***	−0.106***	−0.083***	0.071	−0.069***	−0.009	−0.009	−0.099***	−0.018**	−0.109***
	(0.014)	(0.032)	(0.020)	(0.044)	(0.021)	(0.026)	(0.012)	(0.022)	(0.007)	(0.027)
Own−farm	−0.600***	−0.437***	−0.703***	−0.837***	−0.361***	−0.393***	−0.311***	−0.419***	−0.513***	−0.213***
	(0.034)	(0.091)	(0.053)	(0.086)	(0.025)	(0.077)	(0.039)	(0.038)	(0.021)	(0.037)
Diversified	−0.370***	−0.297***	−0.342***	−0.323***	−0.119***	−0.116	−0.355***	−0.162***	−0.026	0.101*
	(0.038)	(0.092)	(0.065)	(0.120)	(0.035)	(0.118)	(0.054)	(0.043)	(0.025)	(0.052)
Self-employment	−0.343***	−1.032***	1.097***	−0.428***	0.107**	−0.199***	0.132**	0.061	0.191***	2.085***
	(0.035)	(0.331)	(0.189)	(0.104)	(0.047)	(0.074)	(0.064)	(0.099)	(0.043)	(0.125)
Transfers	−0.433***	−0.515***	−0.604***	−0.564***	−0.222***	−0.295***	−0.645***	−0.210**	−0.186***	−0.407***
	(0.050)	(0.059)	(0.062)	(0.134)	(0.027)	(0.089)	(0.038)	(0.102)	(0.019)	(0.059)
Observations	2,910	1799	1,636	3,109	3,830	1,739	3,148	1,823	9,838	4,241
R-squared	0.126	0.143	0.287	0.068	0.127	0.033	0.087	0.091	0.155	0.294

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Table 4. Continued

	Kenya 2006	Cambodia 2005	Tajikistan 2003	Ghana 1998	Nepal 2003	Bangladesh 2000	Madagascar 2001	Niger 2011	Malawi 2011	Malawi 2004
female headed hh	0.033	−0.009	0.002	−0.168***	0.007	0.046	0.073	−0.024	−0.007	−0.085***
	(0.032)	(0.018)	(0.028)	(0.056)	(0.047)	(0.036)	(0.108)	(0.020)	(0.017)	(0.025)
single head of hh	−0.058*	−0.000	−0.010	−0.068	0.035	0.004	−0.031	−0.072***	−0.045**	−0.007
	(0.032)	(0.039)	(0.026)	(0.045)	(0.055)	(0.029)	(0.094)	(0.024)	(0.019)	(0.024)
age head of hh	0.003	0.006*	0.000	0.006	0.003	0.001	−0.012	0.002	0.008***	0.001
	(0.006)	(0.003)	(0.002)	(0.007)	(0.005)	(0.003)	(0.013)	(0.002)	(0.002)	(0.002)
agehead^2	0.000	−0.000	0.000	−0.000	−0.000	−0.000	0.000	−0.000	−0.000***	−0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
literacy	0.057*	−0.018	0.008	−0.102*	−0.031	0.007	−0.211*	−0.015	0.014	−0.008
	(0.029)	(0.013)	(0.019)	(0.053)	(0.030)	(0.014)	(0.115)	(0.011)	(0.012)	(0.016)
Number of people age > 14 and < 60	0.003	0.004	−0.006**	−0.023	−0.007	−0.009	0.015	0.047	−0.004	0.001
	(0.015)	(0.005)	(0.003)	(0.014)	(0.008)	(0.006)	(0.020)	(0.029)	(0.004)	(0.006)
Number of people age < 14	0.006	−0.023***	−0.005*	−0.005	−0.027***	−0.025***	−0.062***	−0.140***	−0.008***	−0.021***
	(0.006)	(0.004)	(0.003)	(0.010)	(0.008)	(0.004)	(0.017)	(0.020)	(0.003)	(0.004)
wealth index	−0.018	−0.086***	−0.032***	−0.033	−0.063***	−0.023*	−0.140***	−0.053***	−0.012**	−0.013***
	(0.020)	(0.021)	(0.007)	(0.028)	(0.012)	(0.013)	(0.030)	(0.014)	(0.005)	(0.004)
Own-farm	−0.667***	−0.556***	0.049***	−0.162***	−0.540***	−0.207***	−0.099	−0.152***	−0.199***	−0.178***
	(0.057)	(0.030)	(0.017)	(0.056)	(0.037)	(0.021)	(0.095)	(0.031)	(0.015)	(0.021)
Diversified	−0.420***	−0.351***	0.044*	−0.159**	−0.206***	0.062***	0.214	−0.145***	−0.227***	−0.141***
	(0.073)	(0.041)	(0.023)	(0.079)	(0.045)	(0.016)	(0.195)	(0.033)	(0.019)	(0.033)
Self-employment	−0.214***	−0.302***	0.107	−0.034	−0.241***	0.108***	0.377***	−0.099***	−0.040	0.045
	(0.068)	(0.034)	(0.161)	(0.102)	(0.052)	(0.024)	(0.134)	(0.029)	(0.036)	(0.036)
Transfers	−0.259***	−0.578***	−0.050**	−0.403***	−0.108*	0.218***	−0.443***	−0.298***	−0.421***	−0.412***
	(0.053)	(0.071)	(0.023)	(0.080)	(0.060)	(0.027)	(0.161)	(0.031)	(0.018)	(0.044)
Observations	8,216	9,268	2,612	3,212	2,694	5,024	1,961	2,410	9,816	9,700
R-squared	0.075	0.139	0.031	0.018	0.123	0.075	0.037	0.099	0.056	0.029

Source: Own calculations based on selected surveys included in the RIGA dataset.

Note: Robust standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Table 5. Regression Results – Pooled Sample

(Y-C)/C	Coef	Std. Err.
Female headed household	−0.002	(0.027)
Single head of household	−0.017	(0.020)
Age of household head	0.003	(0.002)
Age of household head, squared	−0.000	(0.000)
Literacy	−0.050	(0.039)
Number of people age >14 and <60	0.013**	(0.006)
Number of people age <14	−0.009*	(0.004)
Wealth index	−0.018**	(0.007)
Own-farm	−0.364***	(0.061)
Diversified	−0.195***	(0.045)
Self-employment	0.116	(0.157)
Transfers	−0.207***	(0.054)
Consumption recall = 14 days	0.132	(0.105)
Consumption recall = Usual month	0.058	(0.094)
Consumption recall = Last month	0.319***	(0.094)
Consumption recall = Diary	−0.011	(0.092)
Reference = 7 days		
Own-farm recall = Last cropping season	0.186	(0.128)
Own-farm recall = Last 12 months	0.196***	(0.052)
Own-farm recall = Other	0.650***	(0.105)
Reference = 2 visits		
Wage recall = Last month	−0.733***	(0.064)
Wage recall = Months/weeks/hours per day	−0.695***	(0.096)
Reference = 7 days		
Self-employment recall = Last month	0.057	(0.104)
Self-employment recall = Last 12 months	0.162	(0.119)
Self-employment recall = Other	0.078	(0.116)
Reference = 14 days		
Observations	87,987	
R-squared	0.132	

Source: Own calculations based on selected surveys included in the RIGA dataset.
Note: Standard Errors clustered at country/survey level
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

regression (table 5), and a statistically significant negative association is found in only four of 20 cases in the individual country regressions (the coefficient is positive and significant in one case, table 4).

This result may not be conclusive for at least two reasons. The first, which we have already mentioned, is the possibility that the household head's education may not be a good enough predictor of the respondents.[13] The second is that given the nature of our indicator, what we are really testing is whether income underreporting is larger than consumption underreporting for less literate households/individuals.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Gender. Similar to the literacy issue above, our results about the gender of the respondent (or household head) are very tentative as we cannot really control for the gender of the respondent, but only for the gender of the household head. Also, our measures of income and consumption are aggregated at the household level and cannot be directly linked to individual information.

It is not straightforward to argue a priori whether male or female heads of households should be expected to be able to report more accurate household income information. If female headship was associated with fewer income earners (for example because of the death or exit from the household of a husband) we would then expect this variable to be associated with more precise income reporting. If on the other hand most of the income is earned by men, even in female headed households, one would expect female headship to be associated with less accurate income figures.

Both the by-country descriptive analysis in table 3 and the coefficients of multivariate regressions on the female headship variable in table 4 confirm that there is no clear association between female headship and income underreporting. In eight surveys the degree of underreporting of income is larger and statistically significant for female headed households, while in three surveys the degree of underreporting of income is larger and statistically significant for male headed households. In two cases (Vietnam and Madagascar) income is overreported by female headed households and the difference is statistically significant. These inconclusive results are reflected by the regression analysis by country (table 4), where in four cases the coefficient on the female headship variable is negative against two in which it is positive (in the other 14, no relationship is found). The t-test over the pooled sample shows that female headed households underreport income statistically more than male headed households. The pooled regression results confirm the weak relationship among the sex of the household head and the level of underreporting (table 5).

Stylized Fact 3—The Extent of Income Underreporting Varies Systematically with Household Characteristics

Income underreporting and household wealth. A widely accepted tenet of the literature is that the rich tend to underreport income more. Quantifying the extent of the underreporting in different "welfare groups" is complicated by the presence of measurement error in both income and consumption, and by the fact that households overreporting their income or consumption will tend to be classified "by construction" in a higher welfare group when income or consumption are used as welfare measures. In addition, though in our analysis it is not possible to control for shocks, it is worth mentioning that wealthier households are likely to have a greater ability to smooth consumption in the presence of income shocks, and this could partly explain why wealthier households may underreport income more.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

While descriptive statistics by country give inconclusive results (table 3), the sign of the coefficients on the wealth index variable in the regression model supports the proposition that income underreporting be larger for richer households in this particular sample of countries (table 4). Once we control for other factors, wealthier households do appear to underestimate income more in 15 of the 20 survey regressions. This is in fact one of the most consistent results across all the regressors in our model and it is also confirmed by the result from the pooled regression (table 5), as well as the t-test for the pooled sample in the descriptive analysis.

Income underreporting and household demographic structure. If income data are collected by asking one or a limited number of respondents, it is likely that the quality of the data reported decreases as the number of breadwinners in the household increases. Proxy respondents will likely have less precise information on the earnings of the other household members. On the other hand, if households with more children have a higher propensity to consume one would expect the difference between income and consumption measures to be smaller in these households—but this time reflecting the lower amount of savings, not measurement error. Another way in which the number of children may be playing out in this relationship is through the possible correlation between certain types of expenditure and household demographics. If for instance specific components of consumption expenditure that tend to receive particular attention in these surveys (e.g., food) are also more important in households with children, we would expect measures of consumption expenditure to be higher, the higher the number of children, other things being equal. In this case, we would expect our measure of income underreporting to be larger, but this time because of measurement error in the consumption variable.[14]

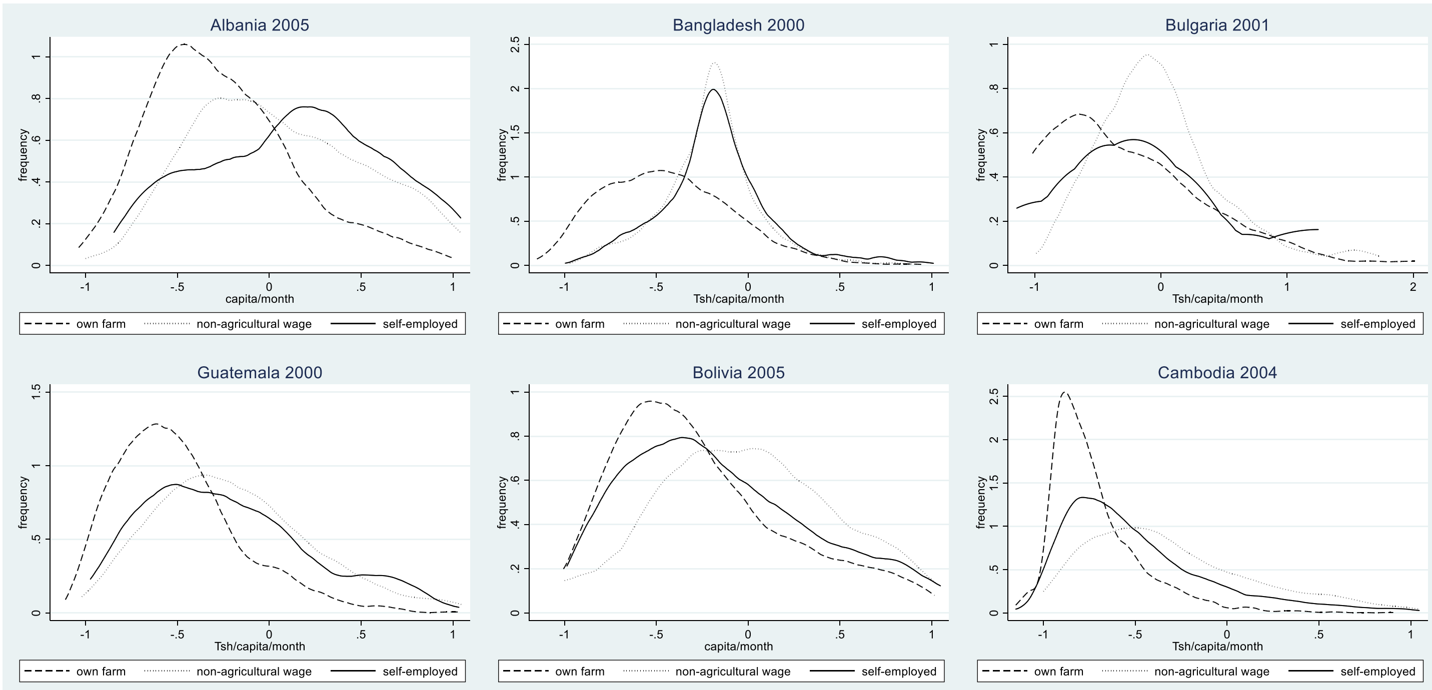
The results in table 4 are mixed for the coefficient on the number of working-age adults in the household (positive in three cases and negative in another four), but the coefficient from the pooled regression shows a positive effect on the normalized difference. On the other hand, the evidence appears to be firmly in support of the hypothesis we just outlined for the number of children. Out of the 20 survey regressions, 13 times the estimated coefficient turns out to be negative and statistically significant, as well as in the pooled sample regression, while it is not significantly different from zero in the other seven cases.

Stylized Fact 4—The Extent of Underreporting Varies Systematically With the Sources of Income, In Particular Income From Own-Account Agriculture and Other Self-Employed Activities Tends To Suffer From Underestimation More Than Wage Income

Figure 1 goes more into depth than table 2 into the observed differences between income and consumption measures in the RIGA surveys. Each graph includes

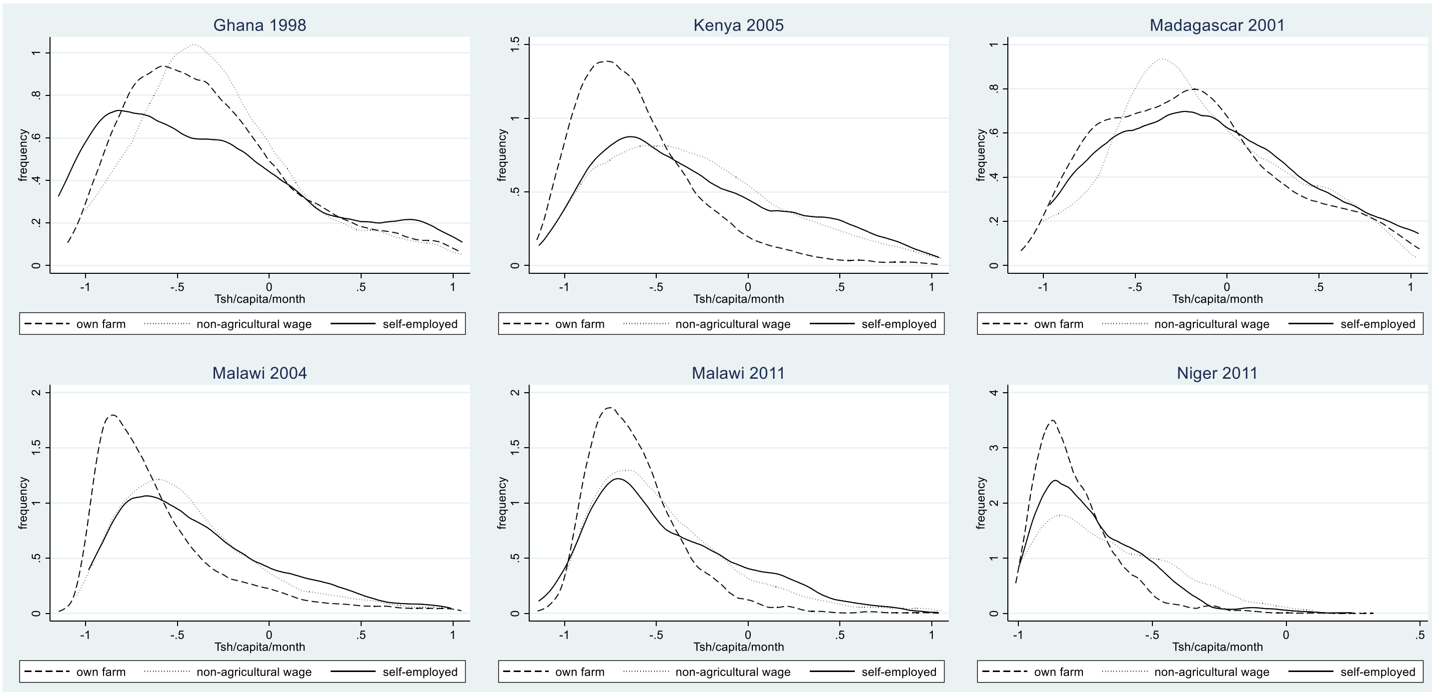
Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Figure 1. Density Distribution of the Normalized Income/Consumption Difference by Specialization



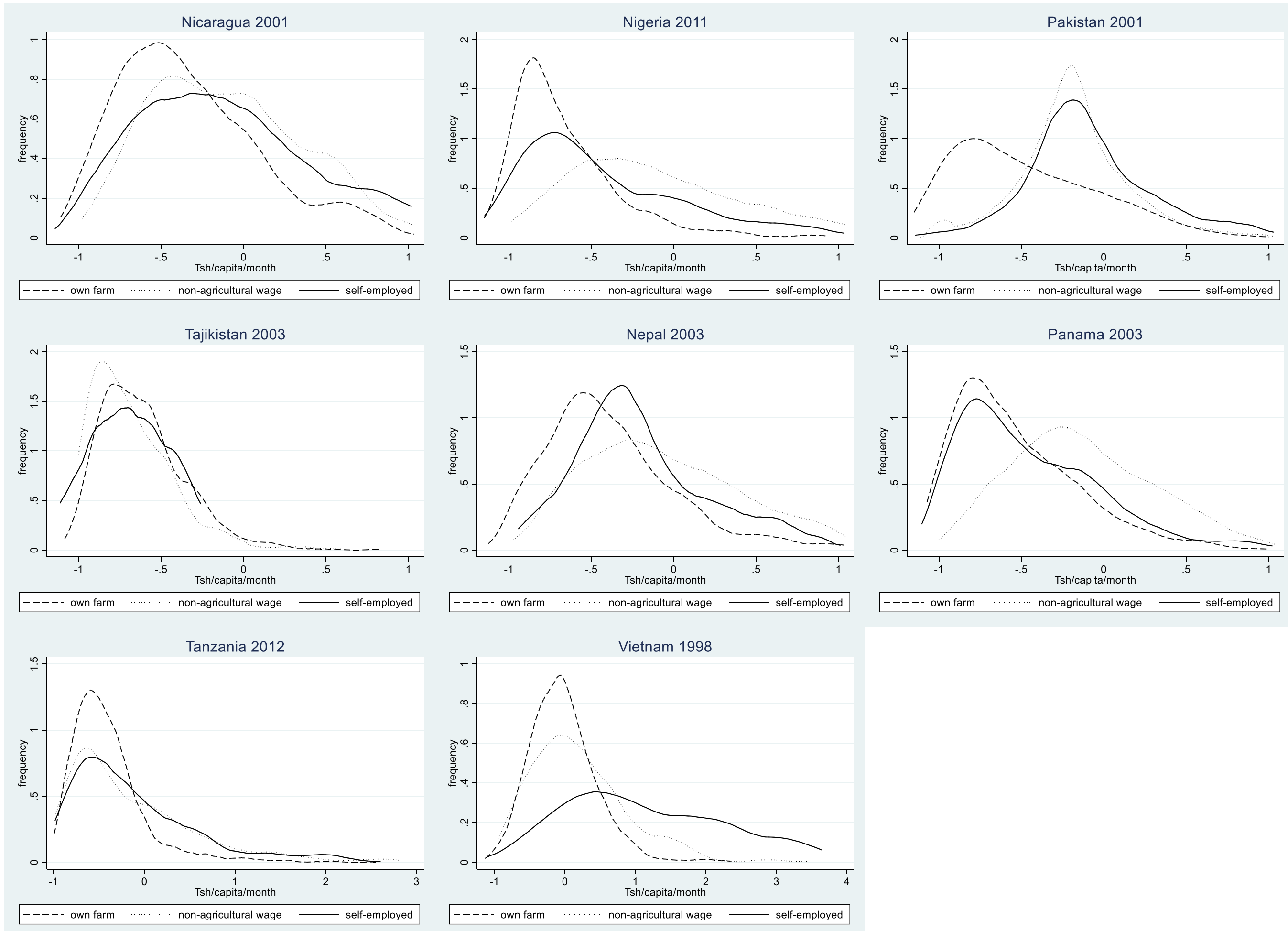
Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Figure 1. Continued



Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Figure 1. Continued



Source: Own calculations based on selected surveys included in the RIGA dataset.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

three density distribution lines of the normalized difference between income and consumption: One for households specialized in agricultural activities, one for households specialized in wage activities, and one for households specialized in non-agricultural self-employment activities. Diversified households and households specialized in transfer income are not reported.

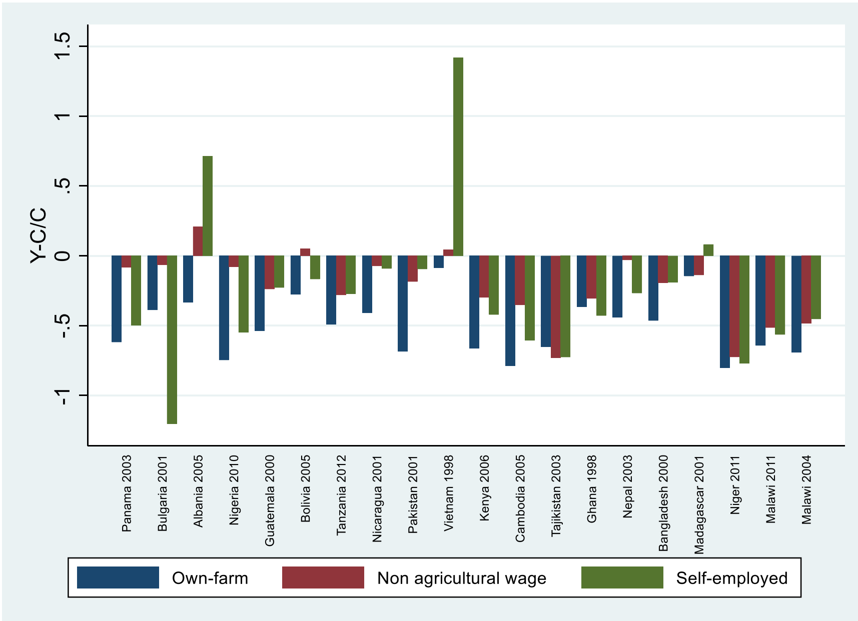
The density distribution of the difference is mostly in the negative part of the horizontal axis, indicating that for a majority of households income as measured by these surveys is lower than consumption expenditure. The mode of the difference distributions tends to be further to the left for agricultural specialized, followed in most cases by non-agricultural self-employed. For wage specialized the line is often centered around zero, indicating that in a majority of cases the normalized difference between income and consumption within that group of households is only marginally negative. The density distribution tends to decline quite rapidly for positive values, which shows that—particularly for agricultural specialized—relatively few households in all countries report a level of income substantially higher than their consumption expenditure.

The observed differences in distributions among the three groups clearly point to the existence of systematic differences in measurement errors in the income variable with the income source. Concerning a comparison of the magnitude of the observed differences across specialization types (and by implication by sources of income), it is somewhat difficult to make generalizations as the variation within each group is quite wide, the highest being between −1.20 for on-farm specialized and 1.42 for self-employment specialized. The median values for the three categories do however convey once again the idea of both the magnitude and the relative size of the underreporting of income by source (fig. 2). Farming is by far the highest with a median value of −0.49 and most of the countries well below −0.30, followed by self-employment (median of −0.24 and 10 out of 20 surveys below the −0.30 threshold) and wage specialized (median of −0.20 and only six surveys below −0.30). However, it is worth noting that self-employment specialized show by far the largest magnitude of overestimation of income over consumption, namely in Vietnam and Albania. In addition, even the t-tests of the difference in means in table 3 clearly show that the level of underreporting is statistically different across the three specializations for almost all countries at 1 percent, the own-farm specialized households being those underreporting the most.

The results of the regression analysis (table 4) strongly support the hypothesis that farm income is the component of income that is most underreported, alongside income from transfers. The coefficient on the dummy for farm specialized is negative and significant in 18 cases, and positive and significant in only one case. The coefficient on the transfers specialized is negative and significant in 19 cases, positive and significant in one. Of the negative coefficients, both those on farm specialized and on transfers specialized are usually the largest in absolute value. The

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Figure 2. Median Normalized Income/Consumption Difference by Specialization



Source: Own calculations based on selected surveys included in the RIGA dataset.

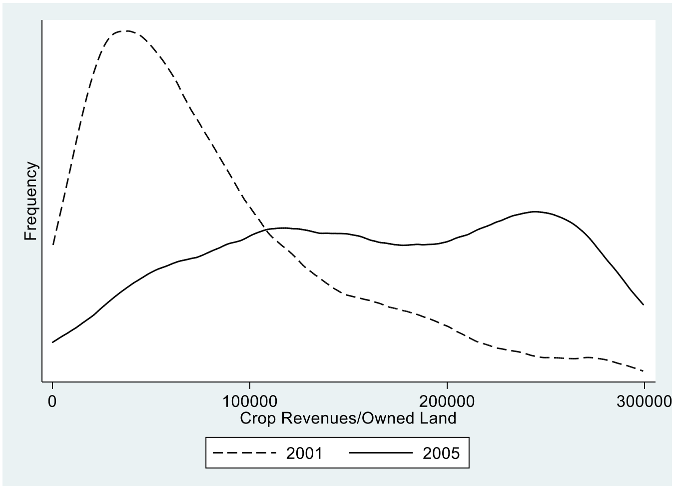
results for self-employment contradict the a priori expectation that self-employment is generally associated with larger underreporting (and hence a negative coefficient) when compared to wage income: in seven surveys the coefficient is positive and significant, in eight negative and significant. The pooled regression (table 5) confirms the by-country regressions, with the own-farm coefficient being negative and the largest in absolute value, while the self-employment is positive but not significant. Diversified households are also more likely to underreport income (14 negative significant coefficients out of 20 country regressions, negative and significant coefficient in the pooled regression), which is not surprising since the relative majority of the income of diversified households comes from farming.

Stylized Fact 5—Questionnaire Design Matters and Can Reduce the Extent of Underreporting

In order to reduce the extent of income underreporting in future surveys it is important to understand how questionnaire design can contribute to more accurate measurement. To test the extent to which income is underreported due to features of

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Figure 3. Kernel Distribution of Crop Revenues Over Land Owned for Albania 2001 and 2005



Source: Own calculations based on selected surveys included in the RIGA dataset.

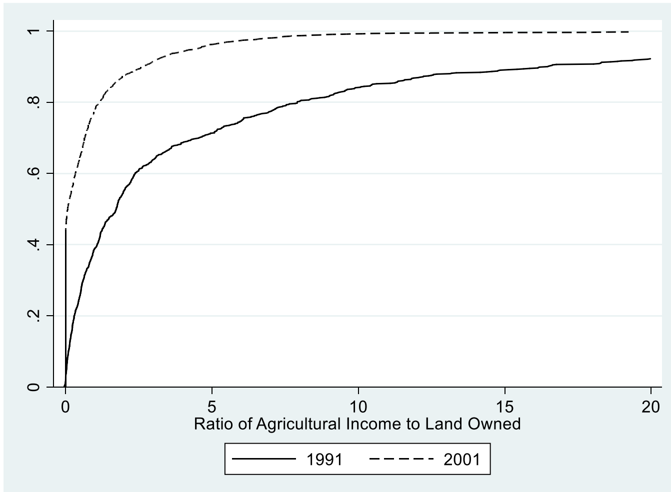
questionnaire design, we rely on additional, earlier rounds of the surveys in Albania and Pakistan. We picked these two countries as these are cases in which we could observe year-on-year differences in questionnaire design that are relevant to our analysis.

In particular, we focus here on differences in the collection of agricultural revenue data, as these are prime suspects as far as the underreporting of agricultural income is concerned. Figure 3 graphs the density distributions of agricultural revenues over land owned for the two survey years in Albania. When comparing the 2001 and 2005 survey results it is immediately apparent how the 2005 lies to the right of the 2001 distribution. The two surveys are identical in the way agricultural revenue data were collected, except for the fact that the 2005 survey prompted the respondents specifically about 44 crops as opposed to 33 in 2001. This confirms, not surprisingly, that asking more detailed questions on agricultural revenues can be an effective way to improve information on farm income.

A similar case can be illustrated with data from Pakistan where, on the contrary, the number of agricultural revenue questions was reduced between the successive survey years. In Pakistan the number of crops fell from 39 in 1991 to just seven in 2001, and the questions were greatly simplified by dropping questions on seed and feed use and own consumption of crops. As a result, the distribution of the agricultural gross income over owned land moved decisively to the left. Figure 4 clearly shows

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Figure 4. Stochastic Dominance of Agricultural Income Over Land Owned for Pakistan 1991 and 2001



Source: Own calculations based on selected surveys included in the RIGA dataset

how the 1991 distribution is always to the right of that of 2001, regardless of the point at which the difference between the two is drawn.

Results from the pooled multivariate analysis confirm that survey design features are associated with the magnitude of the income-consumption difference. In particular, a 30-day recall period in consumption modules is associated with a larger value of the Y-C difference, confirming results from methodological studies suggesting that a longer recall period results in an underestimate of consumption expenditure (see, for example, Scott and Amenuvegbe 1991). Twelve-month recall periods in own-farm modules are associated with a higher income reporting in comparison to two-visits surveys. For wage employment income, a last month recall period is associated with a higher income underreporting compared to a seven-day recall period, which is compatible with the finding in Bardasi et al. (2011), who investigate survey characteristics leading to an underestimation of the employment rate. However, as already noted, we do not want to read too much into these pooled results as the variability in survey design features may also be capturing other dimensions including country fixed effects and other survey design characteristics that we cannot control for.

Summary and Conclusions

This paper uses an innovative household level database to ask two basic questions related to the well-known issue of income underreporting in household surveys in

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

developing countries: (a) What is the extent and range of this underreporting in practice? And (b) does income underreporting vary systematically with respondent, household, income, and survey design features. These are important questions both for the analysis of income information from household surveys, as well having implications for the improvement of the income component of future data collection efforts.

Drawing on data from 20 surveys of developing and transition countries, our results indicate that the observed differences between income and consumption are extremely large, the income/consumption ratio being on average around 0.76 and in four cases below 0.50 (the lowest value being 0.25).

We also find evidence of the underreporting being systematically associated with some household and survey characteristics. Above all the degree of underreporting appears to be strongly associated with the income source. This emerged clearly in the empirical analysis, and findings very clearly identify agricultural income to be the component suffering more than any other component from underreporting.

In all countries in the sample, households that receive a majority of their income from agriculture are those for which the degree of underreporting is largest. They are also the most likely to be reporting negative income. Income from non-agricultural self-employment is also affected, although with patterns that are somewhat less consistent across countries than is the case for agricultural income. Wage specialized, on the other hand, tend to report income figures that are reasonably close to consumption expenditure figures. The analysis also provides evidence supporting the well-established proposition that the extent of underreporting tends to increase with household-level welfare: richer households (with wealth measured in terms of asset ownership) appear to underreport income more.

Taken together these results point to the fact that any analysis of income composition and of the association between level of welfare and sources of income based on household survey data is necessarily going to be fraught with problems stemming from the biases in measurement error we just described.[15] Estimates of the share of agricultural and self-employment income in highly informal economies, for instance, are likely to be affected to a degree that is difficult to capture with any level of accuracy. The observed negative association between agricultural income and poverty is probably robust enough to issues of measurement error, but the fact that biases in measurement by welfare level and income source intersect in ways that we are not able to quantify with accuracy, poses a problem for analyses that look at issues such as the contribution of different sources of income to poverty reduction. Analyses of returns to sector-specific assets are also going to be badly affected by measurement error in these domains. Thus, interpreting and drawing lessons from income data needs particular caution.

Despite all the measurement error, and to paraphrase the Angus Deaton's quote at the beginning of the paper, we emphasize "the value of trying" to collect income

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

data, and of trying harder to improve data quality. Even when using consumption to measure standards of living, income-based measures may be more effective in determining the chronic and transient poor (Chaudhuri and Ravallion 1994), and understanding the sectoral composition of income is likely an important part of the explanation for the study of poverty dynamics. Analyses of intrahousehold distribution of resources, clearly of great interest in well-being and distributional analyses, would also benefit from more precisely and accurately measured income, and these are just a few examples of why abandoning this measurement agenda does not seem to be a viable option. Hence, while we acknowledge both the conceptual and empirical issues in comparing consumption and income, and in measuring welfare through income data in low-income countries, we emphasize the fact that analyses of livelihoods and productivity will still require income, so improving its measurement is essential in order to serve those goals. Finally, as GDP per capita levels increase and the formal sector becomes dominant, income is more likely to become easier to recall for respondents, at the same time as other measurement issues for income will kick in (non-response, capturing non-labor income, and all those other issues OECD countries are familiar with and on which much of the literature on income measurement focuses) and consumption will be seen as a less satisfactory measure of wellbeing. A transition to income for poverty measurement is likely to happen along this income gradient.

The analysis in this paper points to two parallel agendas for action. The first is in the domain of data collection, and relates to intensifying the efforts to harmonize and collect better income data, particularly on agricultural income. The paper has clearly shown how questionnaire design does have important implications for the quality of the data generated by the surveys. The fact that age and literacy of the respondents affect the outcome also points to the need to ensure particular care in survey design and fieldwork to minimize the measurement error when working with illiterate survey respondents. Additional work is needed to explore in more depth how different components of income and different aspects of survey methodology affect how income is measured. In the last decade, modern technologies have brought about improvements in data quality in many areas of survey design and implementation in low-income countries, and income can be one of the next areas of focus. Mobile communication technology can for instance be applied to reduce the costs of administering diaries over extended periods of time at a fraction of the cost. Examples exist for the use of mobile phones to aid collection of diary data on agricultural labor, and on extended harvest crops, but the technology has not to our knowledge been applied to the collection of income data at scale.

The second agenda is an agenda for future research. While we have started putting some numbers to the extent of income underreporting and to its "sectoral composition," much of the variance in income underreporting remains unexplained. In particular, we would need to gauge a better understanding of how the sectoral source of income underreporting relates to questionnaire design. Survey methods

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

experiments can help shed more light on that, and would allow the extracting of more pointed guidance for survey design practitioners. Also, the variability in the observed magnitude in underreporting is such that it is difficult at this stage to think of an "adjustment factor" one may use to correct observed income data in future surveys. More analysis is needed to carefully assess how different issues we have identified in this paper have implications for measurement error, and to try and quantify at least some plausible ranges of adjustment factors that could be applied with confidence to income data, or to develop methodologies to consistently estimate the degree of income underreporting in household surveys.

Notes

Gero Carletto (gcarletto@worldbank.org), Marco Tiberti (mtiberti@worldbank.org), and Alberto Zezza, (azezza@worldbank.org; corresponding author) are respectively Manager, Economist and Program Manager with the Data Production and Methods unit in the Development Data (DECDG) Department of the World Bank, 1818 H St NW, Washington DC, USA.

The views expressed in the paper are the authors' only, and should not be attributed to the organizations they are affiliated with. The authors are grateful to Carlo Azzarri, Kathleen Beegle, Katia Covarrubias, Ana Paula de la O Campos, Carly Petracco, Kinnon Scott, and three anonymous reviewers for inputs and comments at various stages in the production of this paper.

1. The paper uses the expression low-income countries loosely with no direct reference to official classifications.
2. Up-to-date information on the RIGA database can be found on the FAO's website.
3. IDA is the part of the World Bank Group that deals with concessional loans with more favorable terms for borrowers. IDA eligibility is determined by the level of GNI per capita and the lack of creditworthiness to borrow on market terms.
4. Information on the LSMS-ISA program and the datasets and documentation it produces is available on the World Bank's website.
5. A detailed description of the construction of the income aggregates can be found in Carletto et al. (2017) and in Covarrubias et al. (2009).
6. The paper uses a 50 percent threshold to define households as being specialized in one activity. Tests were performed using a 75 percent threshold and the results were found to be robust to the change in definition. The use of the 50 percent thresholds is preferable because it yields larger sample sizes in each group.
7. On the consumption side see Deaton and Grosh (1998), Beegle et al. (2012), Backiny-Yetna et al. (2017), Zezza et al. (2017); on income Marquis and Moore (1990) and Sudman and Bradburn (1982).
8. On the consumption side, eight surveys use a seven-day recall module, three have a 14-day recall period, two a one-month recall period, three surveys collect consumption for a "usual month" (implying a 12-month recall period), four surveys use a diary survey. On the income side, survey design choices vary with the type of employment and income sources. For on-farm income, 10 surveys have a last 12-month recall period, four surveys a two-season (typically dry/rainy seasons) recall period, five surveys use a last cropping season recall period, and one a multiple recall period. For wage employment, eight surveys report income for a last-month recall period, 11 surveys have a multiple recall period (Months/weeks/hours per day), one survey has a one-week recall period. For non-farm business income,

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

eight surveys report income for a one-month recall period, six surveys have a 12-month recall period, four have a multiple recall period, and two have a 14-day recall period.

9. In order to provide a global t-test for each stylized fact over the whole sample, we estimate a Weighted Least Square regression, controlling for heteroskedasticity.

10. The choice of assets incorporated varies by country but generally it includes household durables (TV, VCR, stove, refrigerator, etc.) as well as household infrastructure (running water, brick walls, etc.). "Agricultural" assets (land, livestock, machinery) are not included as this variable is meant to capture overall wealth rather than participation in a specific productive activity. The values of the indices are not comparable across countries, though the method of construction is comparable and in all cases the values go in the same direction: More is better. Thus, while for the econometric analysis the sign of the parameter is comparable across countries, the magnitude of the effect is not.

11. For comparison with our data, table A1 in the supplementary online appendix reports income and consumption levels from the dataset used in McKay (2000).

12. Both income and consumption aggregates were deflated by a spatial price deflator, as is standard practice in cross-sectional survey analysis to account for the fact that "people who live in different parts of the country pay different prices for comparable goods" (Deaton and Zaidi 2002).

13. The result in the multivariate analysis does not change if average household education is used instead of household head literacy.

14. The use of the per capita measure (instead of adult equivalents) was tested and did not affect the results. When the same analysis is performed using total household income and consumption, the results are qualitatively the same.

15. Analyses of welfare and inequality issues that try to reconcile income and consumption-based measures will also be affected, but this is an area on which there have been substantial contributions from the literature (Meyer and Sullivan 2003; Meyer and Sullivan 2012; Joliffe et al. 2015) and that is not touched upon in this paper.

References

Akee, R. K. Q. 2007. " Errors in Self-Reported Earnings: The Role of Previous Earnings Volatility." Institute for the Study of Labor Discussion Paper No. 3263. Bonn, Germany: IZA.

Alderman, H. 1993. "Obtaining Useful Data on Household Incomes from Surveys." In *Data Needs for Food Policy in Developing Countries: New Directions for Household Surveys*, edited by J. von Braun and D. Puetz, 193–201. Washington, DC: International Food Policy Research Institute.

Anand, S., and P. Segal. 2008. "What Do We Know About Global Income Inequality?" *Journal of Economic Literature* 46 (1): 57–94.

Atkinson, A. B., and E. Marlier. 2010. *Income and living conditions in Europe*. Luxembourg: European Union.

Atkinson, A. B., and J. Micklewright. 1983. "On the Reliability of Income Data in The Family Expenditure Survey 1970–1977." *Journal of the Royal Statistical Society* 146 (1): 33–53.

Backiny-Yetna, P., D. E. Steele, and I. Yacoubou Djima. 2017. "The Impact of Household Food Consumption Data Collection Methods on Poverty and Inequality Measures in Niger." *Food Policy*, 72: 7–19.

Banerjee, A. V., and T. Piketty. 2003. *Top Indian Incomes, 1956–2000*. Cambridge, MA: Massachusetts Institute of Technology Department of Economics Working Paper Series 03–32.

Bardasi, E., K. Beegle, A. Dillon, and P. Serneels. 2011. "Do Labor Statistics Depend on How and to Whom the Questions Are Asked: Results from a Survey Experiment in Tanzania." *World Bank Economic Review* 25 (3): 418–47.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Battistin, E., De Nadai M., and Krishnan N.. 2020. "The insights and illusions of consumption measurement: evidence from a large-scale randomization." Policy Research Working Paper 9255. World Bank, Washington DC.

Beegle, K., J. De Weerdt, J. Friedman, and J. Gibson. 2012. "Methods of Household Consumption Measurement Through Surveys: Experimental Results from Tanzania." *Journal of Development Economics* 98 (1): 3–18.

Berry, A. 1985. "On Trends in the Gap Between Rich and Poor in Less Developed Countries: Why We Know so Little." *Review of Income and Wealth* 31 (4): 337–54.

Branch, E. R. 1994. "The Consumer Expenditure Survey: A Comparative Analysis." *Monthly Labor Review* 117 (12): 47–55.

Brandolini, A. 1999. "The Distribution of Personal Income in Post-War Italy: Source Description, Data Quality, and the Time Pattern of Income Inequality." *Giornale degli Economisti* 58 (2): 183–239.

Carletto, C., K. Covarrubias, B. Davis, M. Krausova, and P. Winters. 2017. Rural Income Generating Activities Study: Methodological note on the construction of income aggregates. FAO: Rome.

Covarrubias, K., A.P. de la O Campos, and A. Zezza. 2009. Accounting for the diversity of rural income sources in developing countries: The experience of the Rural Income Generating Activities Project. FAO: Rome.

Chaudhuri, S., and M. Ravallion. 1994. "How well do static indicators identify the chronically poor?." *Journal of Public Economics* 53 (3): 367–94.

Coulombe, H., and A. McKay. 1995. "An Assessment of Trends in Poverty in Ghana, 1988–92." PSP Discussion Paper No. 81, Poverty and Social Policy Department, World Bank, Washington DC.

Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Baltimore, Maryland: Johns Hopkins University Press.

———. 2000. "Counting the World's Poor: Problems and Possible Solutions." *World Bank Research Observer* 16 (2): 125–47.

———. 2005. "Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)." *Review of Economics and Statistics* 87 (1): 1–19.

Deaton, A., and M. Grosh. 1998. "Chapter 17: Consumption." In *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Survey*, edited by M. Grosh and P. Glewwe, 105–37, vol. 2. Washington, DC: World Bank.

Deaton, A., and S. Zaidi. 2002. *Guidelines for Constructing Consumption Aggregates for Welfare Analysis*. Washington, DC: World Bank.

De Magalhaes, L., and R. Santaeulalia-Llopis. 2018. "The Consumption, Income and Wealth of the Poorest: An Empirical Analysis of Economic Inequality in Rural and Urban Sub-Saharan Africa for Macroeconomists." *Journal of Development Economics* 134: 350–71.

de Mel, S., D. J. McKenzie, and C. Woodruff. 2009. "Measuring Microenterprise Profits: Must We Ask How the Sausage is Made?" *Journal of Development Economics* 88: 19–31.

Ferreira, F. H. G., S. Chen, and A. Dabalen et al. 2016. "A Global Count of the Extreme Poor in 2012: Data Issues, Methodology and Initial Results." *J Econ Inequal* 14: 141–72.

Farfán, Gabriela, Genoni María Eugenia, and Vakis Renos 2017. "You are what (and where) you eat: Capturing food away from home in welfare measures." *Food Policy* 72(C): 146–56.

Filmer, D., and Pritchett LH. 2001. "Estimating wealth effects without expenditure data–or tears: an application to educational enrollments in states of India" *Demography* 38 (1): 115–32.

Fisher, M., J. J. Reimer, and E. R. Carr. 2010. "Who should be Interviewed in Surveys of Household Income?" *World Development* 38 (7): 966–73.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

Friedman, J., K. Beegle, J. De Weerdt, and J. Gibson. 2017. "Decomposing Response Error in Food Consumption Measurement: Implications for Survey Design from a Randomized Survey Experiment in Tanzania." *Food Policy* 72: 94–111.

Gertler, Paul J., Sebastian W. Martinez, and Marta Rubio-Codina. 2012. "Investing Cash Transfers to Raise Long-Term Living Standards." *American Economic Journal: Applied Economics* 4(1): 164–92.

Gibson, J., Huang J., and Rozelle S.. 2003. "Improving Estimates of Inequality and Poverty from Urban China's Household Income and Expenditure Survey." *Review of Income and Wealth* 49: 53–68.

Gottschalk, P., and M. Smeeding. 1999. *Empirical Evidence on Income Inequality in Industrial Countries*. Luxembourg Income Study Working Paper No. 154. Luxembourg Income Study, Luxembourg.

Grootaert, C. 1993. "How Useful Are Integrated Household Survey Data for Policy-Oriented Analyses of Poverty?" Policy Research Working Paper Series 1079, Washington, DC: World Bank.

Jolliffe, D. R. 2001. "Measuring Absolute and Relative Poverty: The Sensitivity of Estimated Household Consumption to Survey Design." *Journal of Economic and Social Measurement* 27 (1–2): 1–23.

Jolliffe, D., P. Lanjouw, S. Chen, A. Kraay, C. Meyer, M. Negre, E. Prydz, R. Vakis, and K. Wethli. 2015. *A Measured Approach to Ending Poverty and Boosting Shared Prosperity: Concepts, Data and the Twin Goals*. Washington DC: World Bank.

Joshi, K., R. Hasan, and G. Amoranto. 2009. "Surveys of Informal Sector Enterprises – Some Measurement Issues." Asian Development Bank Economics Working Paper Series No. 183. Asian Development Bank, Manila, Philippines.

Kilic, T., and H. Moylan. 2016. *Methodological experiment on measuring asset ownership from a gender perspective (MEXA): technical report*. Washington, DC: World Bank.

Korinek, A., J. A. Mistiaen, and M. Ravallion. 2006. "Survey Nonresponse and the Distribution of Income." *Journal of Economic Inequality* 4 (1): 33–55.

Krell, K., J. R. Frick, and M. M. Grabka. 2017. "Measuring the Consistency of Cross-Sectional and Longitudinal Income Information in EU-SILC." *Review of Income and Wealth* 63: 30–52.

Lanjouw, P., and M. Ravallion. 1996. "How Should We Assess Poverty Using Data from Different Surveys?" Poverty Lines, No. 3. A joint publication by the Policy Research and the Poverty and Social Policy Departments of the World Bank. Washington DC: World Bank.

Marquis, K., and J. Moore. 1990. "Measurement Errors in SIPP Program Reports." *Proceedings of the Sixth Annual Research Conference*, Washington, DC: U.S. Bureau of the Census, March, 721–45.

McKay, A. 2000. "Should the Survey Measure Total Household Income?" In *Designing Household Survey Questionnaires for Developing Countries: Lessons from Fifteen Years of Living Standard Measurement Study*, edited by M. Grosh and P. Glewwe, 83–104. Washington, DC: World Bank.

Mel, S. de, D. J. McKenzie, and C. Woodruff. 2009. "Measuring Microenterprise Profits: Must we Ask How the Sausage is Made?" *Journal of Development Economics* 88: 19–31.

Meyer, B. D., and J. X. Sullivan. 2003. "Measuring the Well-being of the Poor Using Income and Consumption." *Journal of Human Resources* 38 (Special Issue on Income Volatility and Implications for Food Assistance Programs): 1180–220.

Meyer, B. D., and J. X. Sullivan. 2012. "Consumption and Income Poverty in the U.S.." In *The Oxford Handbook of the Economics of Poverty*, edited by P. N. Jefferson, 49–74. Oxford and New York: Oxford University Press.

Moore, J. C., L. L. Stinson, and E. J. Welniak. 2000. "Income Measurement Error in Surveys: A Review." *Journal of Official Statistics-Stockholm* 16 (4): 331–62.

Pradhan, M. 2001. *Welfare Analysis with a Proxy Consumption Analysis*, Tinbergen Institute Discussion Paper No. 092/2. Tinbergen Institute, Amsterdam and Rotterdam.

Scott, C., and B. Amenuvegbe. 1991. "Recall Loss and Recall Duration: An Experimental Study in Ghana." *Inter-Stat* 4 (1): 31–55.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023

SEDLAC. 2012. *A Guide to the SEDLAC Socio-Economic Database for Latin America and the Caribbean*. CEDLAS and the World Bank.

Smith, L.C., Dupriez O., and Troubat N. 2014. Assessment of the reliability and relevance of the food data collected in national household consumption and expenditure surveys. IHSN Working Paper No. 8.

Sudman, S., and N. Bradburn. 1982. *Asking Questions*. San Francisco: Jossey-Bass.

Székely, M., and M. Hilgert. 1999. "What's Behind the Inequality We Measure? An Investigation Using Latin American Data." *Research Department Working Paper No. 409*, Inter-American Development Bank, Washington, DC.

United Nations. 2018. Statistical Commission Report on the forty-ninth session (March 6–9, 2018). Economic and Social Council Official Records, 2018 Supplement No. 4. E/2018/24-E/CN.3/2018/37.

Vijverberg, W. P. M. 1991. *Measuring Income from Family Enterprises with Household Surveys*. LSMS Working Paper No. 84. Washington, DC: World Bank.

Vijverberg, W. P. M., and D. C. Mead. 2000. "Household Enterprises." In *Designing Household Survey Questionnaires for Developing Countries: Lessons from 15 years of the Living Standards Measurement Survey*, edited by M. Grosh and Glewwe, vol. 2. Washington DC: World Bank.

Weinberg, D. H. 2004. "Income Data Quality Issues in the Annual Social and Economic Supplement to the Current Population Survey." Prepared for American Enterprise Institute-University of Maryland Seminar on Poverty Measurement, October 12, 2004.

Zezza, A., C. Carletto, J. L. Fiedler, P. Gennari, and D. Jolliffe. 2017. "Food counts. Measuring food consumption and expenditures in household consumption and expenditure surveys (HCES). Introduction to the special issue." *Food Policy* 72: 1–6.

Downloaded from https://academic.oup.com/wbro/article/37/1/1/6298662 by World Bank user on 27 July 2023