Human Development Research Paper 2011/13

Inequality Decomposition without Income or Expenditure Data: Using an Asset Index to Simulate Household Income

Kenneth Harttgen
and Sebastian Vollmer

United Nations Development Programme
Human Development Reports
Research Paper

November 2011

Human Development
Research Paper
2011/13

Inequality Decomposition without Income or Expenditure Data: Using an Asset Index to Simulate Household Income

Kenneth Harttgen
and Sebastian Vollmer

United Nations Development Programme
Human Development Reports
Research Paper 2011/13
November 2011

Inequality Decomposition without Income or Expenditure Data: Using an Asset Index to Simulate Household Income

Kenneth Harttgen
and Sebastian Vollmer

Kenneth Harttgen is Researcher at the Swiss Federal Institute of Technology, Nadel, ETH Zurich.
E-mail: harttgen@nadel.ethz.ch.

Sebastian Vollmer is Instructor at the University of Hannover, Institute of Macroeconomics and Harvard School of Public Health, Department of Global Health and Population. E-mail: svollmer@hsph.harvard.edu.

Acknowledgments: We thank Stephan Klasen and Milorad Kovacevic for helpful comments and discussion
Comments should be addressed by email to the author(s).

Abstract

Commonly available survey data for developing countries often do not include income or expenditure data. This data limitation puts severe constraints on standard poverty and inequality analyses. We provide a simple approach to simulate household income based on publicly available Demographic and Health Surveys (DHS) and macroeconomic data. We illustrate our approach with DHS data for Bolivia, Burkina Faso, Indonesia and Zambia. We calculate standard inequality measures and decompose inequality by urban/rural, sex of the household head, household size and education of the household head.

Keywords: Inequality, Asset index, Income simulation

JEL classification: D31, I31, I32, O1

1 Introduction

Household income or expenditure data are often used to measure current and long-term welfare of households and within-country inequality (see for example Deaton, 1997). The availability of household survey data has increased the understanding of within-country inequality and its determinants. Large scale national representative household survey data has become more and more available in recent years. However, oftentimes commonly available survey data for developing countries – such as the Demographic Health Surveys (DHS) – do not include income or expenditure data.

Filmer and Pritchett (2001) and Sahn and Stifel (2001) have proposed a one-dimensional index based on household assets and other household characteristics as a proxy of long-term material welfare to overcome the problem of missing income and expenditure data. The so-called 'asset index' is often used in the empirical literature on poverty and inequality analysis as a proxy variable for household income. There is a large body of literature that uses an asset index to explain inequalities in educational outcomes (e.g. Ainsworth and Filmer, 2006; Bicego et al., 2003), health outcomes (e.g. Bollen et al., 2002; Schellenberg et al., 2003), child malnutrition (e.g. Sahn and Stifel, 2003; Tarozzi and Mahajan, 2005), or child mortality (e.g. Sastry, 2004) when data on income or expenditure is not available. In addition, asset indices are used to analyze changes and determinants of poverty (Harttgen and Misselhorn, 2007; Sahn and Stifel, 2000; Stifel and Christiaensen, 2007; World Bank, 2006). Although the asset index has some shortcomings, which we will address below, it has become a popular tool to overcome the problem of missing data on income or expenditure.

However, the asset index does not directly allow the calculation of standard measures of income inequality such as the Gini coefficient or the class of Atkinson inequality measures –

even if we assume that it is a good proxy for income or expenditure data. In this paper, we provide a simple and intuitive approach to simulate household income[1] based on an asset index and publicly available macroeconomic data. With the simulated income data we can perform any kind of standard poverty and inequality analysis within and across countries and over time and also by population subgroups and by socioeconomic household characteristics.

The paper is structured as follows: In Section 2, we explain our approach to simulate household income from commonly available DHS and macroeconomic data. In Section 3, we illustrate our approach with DHS data for Bolivia, Burkina Faso, Indonesia and Zambia. We calculate Gini coefficients and Atkinson inequality measures for every country. We also decompose inequality by urban/rural, sex of the household head, household size and education of the household head. In the fourth section we discuss some of the advantages and limitations of our approach. In the final section, we conclude.

2 Methodology

2.1 Simulating Household Income

Recently, a number of papers have estimated national income distributions solely from readily available from readily available macroeconomic data such as income quintiles (Sala-i-Martin, 2006) or Gini coefficients and GDP per capita (Vollmer et al., 2010). The basic idea of this paper is to link the DHS data to national income distributions in order to simulate household income.

First, we use principal component analyses to construct an asset index for each household. Second, we estimate the national income distribution from the Gini coefficient and GDP per

[1] Throughout this paper household income refers to household income per capita. This means that although we provide data on an individual basis, each household member within a household has the same household income. Any poverty and inequality analysis is then weighted by the household size.

capita. To link the asset index to the appropriate income, we make the assumption that the ranking of households within the asset distribution is the same as the ranking within the income distribution. The household at the p-quantile of the asset distribution gets the income level at the p-quantile of the national income distribution assigned.

If one is only interested in calculating other inequality measures than the Gini (or maybe some poverty measures), then the last step is not necessary. One can calculate any given inequality or poverty measure by drawing a large enough random sample from the fitted log-normal income distribution and calculating the measure for this sample. The big advantage of simulating income for the DHS data is that the DHS data include a large number of covariates that can for example be used to calculate inequality for certain sub-population or to calculate the income gradient of education or health.

2.2 Constructing an Asset Index

We follow the approach of Filmer and Pritchett (2001) and Sahn and Stifel (2003) to construct an asset index. The main idea of this approach is to construct an aggregated one-dimensional index over the range of different dichotomous variables of household assets capturing housing durables and information on the housing quality that indicate the material status (welfare) of the household:

$$A_i = b_1 a_{i1} + b_2 a_{i2} + \dots + b_k a_{ik} \quad (1)$$

$$a_{ik} = \beta_k c_i + u_{ik} \quad (2)$$

for i =1, ..., N households and k =1, ...K household assets. A_i is the asset index, the a_{ik} refer to the respective asset of the household i recorded as dichotomous variables in the DHS data sets, and the b_k are the weights for each asset that are used to aggregate the indicators to an one-dimensional index. In the model, the ownership of an asset k of household i, identified by a_{ik}, is a linear function of an unobserved factor, which in our case is material welfare c_i. The relationship between the asset k in c_i is given by β_k plus a noise component u_{ik}, where both terms have to be estimated (Sahn and Stifel 2000).[2]

For the estimation of the weights and for the aggregation of the index, we use a principal component analysis as proposed by Filmer and Pritchett (2001). The first principal component is our asset index.[3] Principal component analysis is a technique to identify those linear combinations from a set of variables that best capture the common information behind the variables. This means that we assume that household assets and housing characteristics explain the long-term wealth of a household measured by the maximum variance in the asset variables.

The principal component analysis is structured by a set of equation where the asset variable is related to a set of latent factors:

$$\begin{aligned} \tilde{a}_{1i} &= v_{11}A_{1i} + v_{12}A_{2i} + \cdots + v_{1k}A_{ki} \\ &\dots \\ \tilde{a}_{ki} &= v_{k1}A_{1i} + v_{k2}A_{2i} + \cdots + v_{kk}A_{ki} \end{aligned} \tag{3}$$

where the $\tilde{a}$ are the k asset indicators (the a's in equation 1) normalized by their mean and their standard deviations; A are the k principal components and v are the weights that relate the

[2] The model is based on the following assumptions: (i): households are distributed *iid*; (ii): $E(u_i|c_i)$=0; (iii): $V(u_i) = Diag\{\sigma_1^2, \dots, \sigma_K^2\}$.

[3] An alternative way to estimate the weights for the assets to derive the aggregated index is a factor analysis employed, for example, by Sahn and Stifel (2001). However, the two estimation methods show very similar results.

principal components to the ownership of the asset (Filmer and Scott, 2008). After the weights v have been estimated, the inversion of the equation system (3) yields the following set of equations:

$$
\begin{aligned}
A_{1i} &= b_{11}\tilde{a}_{1i} + b_{21}\tilde{a}_{2i} + \cdots + b_{k1}\tilde{a}_{ki} \\
&\dots \\
A_{ki} &= b_{1k}\tilde{a}_{1i} + b_{2k}\tilde{a}_{2i} + \cdots + b_{kk}\tilde{a}_{ki}
\end{aligned}
\tag{4}
$$

The equation for the first principal component is the equation with the highest variance. The weights that are used to aggregate the asset variables into a one-dimensional index are given by the set $(b_{11}, b_{21}, \dots, b_{k1})$. The asset index is calculated for each individual, weighted by the household size.

2.3 The Distribution of Income

We follow the approach of Vollmer et al. (2010) to estimate national income distributions. Thereby, we assume that national income distributions follow a log-normal distribution. Formally, the log-normal distribution LN (μ,σ) is defined as the distribution of the random variable $Y = \exp(X)$, where X has a normal distribution with mean μ and standard deviation σ.

The Gini coefficient G of LN (μ,σ) is given by $G = 2\Phi(\sigma/\sqrt{2}) - 1$, where Φ is the distribution function of the standard normal distribution. Therefore, the parameters μ and σ of LN (μ,σ) can be determined from the average income E(Y) and the Gini coefficient G as follows.

$$\sigma = \sqrt{2\phi^{-1}}\left(\frac{G+1}{2}\right), \qquad \mu = \log\big(E(Y)\big) - \sigma^2/2.$$

Our crucial assumption for mapping the asset index to household income is that the ranks within both distributions are equal, i.e. the household at the p-quantile of the asset distribution will also be at the p-quantile of the income distribution. The corresponding household income is then given by the p-quantile x(p) of the log-normal distribution of income:

$$x(p) = e^{\mu + u(p) \cdot \sigma},$$

where u(p) is the p-quantile of the standard normal distribution and μ and σ are the parameters of the log-normal distribution of income.

3 Empirical Analysis

3.1 Data

To illustrate our approach we use Demographic and Health Survey (DHS) data for four countries: Bolivia, Burkina Faso, Indonesia and Zambia (see Table 1). The DHS are undertaken by Macro International Inc., Calverton, Maryland (usually in cooperation with local authorities and funded by USAID) and started in 1984. They provide detailed information on child mortality, health, and fertility. To date, DHS data is available for 84 developing countries for several years – resulting in more than 240 large scale household surveys. The data are self-weighted national survey of women aged between 15 and 49. The average sample size is about 5,000 to 6,000 women, some are surveys are even larger than that.

The DHS include a household member module and an individual recode for women of reproductive age. The household member recode lists all member of the household. At the household level, the DHS provide information on basic demographics, education and on the possession of household assets. Although the DHS are not completely standardized across time and countries, the design and coding of variables (especially on assets and dwelling characteristics) are generally comparable.

We use the following variables to construct an asset index: radio, TV, refrigerator, bike, motorized transport, capturing household durables and type of floor material, type of wall material, type of toilet, and type drinking water capturing the housing quality and we calculate the asset indices separately for each country and period. The results from the principal components analysis are shown in Table 2. The macro data, i.e. the Gini coefficient and income per capita, are taken from PovCalNet. The values are shown in Table 3.

3.2 Inequality Decomposition

In Tables 4-7 we illustrate our method for Bolivia, Burkina Faso, Indonesia and Zambia. We estimate income per capita and inequality for the total population and for the following subgroups: rural/urban, sex of the household head, household size and education of the household head.

We report the Gini coefficient and three Atkinson inequality measures. The Gini coefficient can be interpreted as the expected income difference between two individuals randomly selected from the population scaled by the average income. The Gini coefficient is probably the most widely used and commonly known inequality measure, but it lacks economic theory. The Atkinson inequality measures are directly derived from welfare theory. Individual incomes enter the social welfare function with diminishing marginal utility. The equally distributed equivalent income then represents the level of income per capita which, if equally shared, would generate the same level of social welfare as the observed distribution. Based on this notion, the Atkinson inequality measure is defined as the percentage of average income that is lost due to inequality (in relation to the equally distributed equivalent income). The exact form of the Atkinson inequality measure depends on the parameter ε of inequality aversion in the social welfare function. Commonly used values for ε are 0.5, 1 and 2. One additional advantage of the Atkinson measure over the Gini coefficient is that it can be decomposed into inequality within and between groups.

We find that inequality is the highest in Bolivia. At the highest level of inequality aversion (ε=2) we find that 75 percent of average income is lost for social welfare due to inequality. Even at the lowest level of inequality aversion (ε=0.5), this number is still sizeable with 30 percent. Bolivia also has the strongest gradient between rural and urban areas as well as for education. The average income in urban areas is about five times higher than in rural areas. The average income of individuals with more than secondary education is about seven times higher than the average income of individuals with less than primary education. Both for rural/urban and for education, between group inequality adds a sizeable share to overall inequality. For sex of the household head and household size there is hardly any between group inequality.

Inequality is smallest in Indonesia. If inequality aversion is high (ε=2), 27 percent of average income is lost for social welfare due to inequality. If inequality aversion is low (ε=0.5), only 7.5 percent of average income is lost for social welfare due to inequality. Also in Indonesia, there is an income gradient between rural and urban areas as well as for education. In both cases, between group inequality has a sizeable contribution to overall inequality, however the magnitudes are much smaller than in Bolivia. The average income in urban areas is a little bit less than twice as high as the average income in rural areas. Individuals with more than secondary education have an average income that is close to three times as high as the average income for individuals with less than primary education. Again, we don’t find any relevant income gradients for sex of the household head and household size.

We also find the gradients for education and rural/urban in Burkina Faso and Zambia. The magnitudes in both countries are somewhere between Bolivia and Indonesia.

4 Discussion

We made two important assumptions in order to simulate household income from an asset index and macroeconomic data: 1) National income distributions are log-normally distributed and 2) the individual ranks of households are the same in the income distribution as they are in the distribution of the asset index. Both assumptions need some discussion and justification.

We will start with the log-normality assumption. Of course, the log-normal distribution is only an approximation to the true form of national income distributions. With a large enough micro data set on income one could most likely reject the log-normal assumption – as one could reject any other simple parametric assumption. However, the fact that such data is not commonly available, in particular in the DHS, was the reason why we made the log-normal assumption in the first place. For the available data, the log-normal assumption turns out to be a quite good approximation for national income distributions. Lopez and Serven (2006) test the log-normal assumption systematically for a large number of countries and years (about 800 country-year observations), and they find that log-normality cannot be rejected for income data. One could also use other approaches to estimate national income distributions, for example the non-parametric approach of Sala-i-Martin (2006). However, the non-parametric approach would not significantly improve the fit to the data (this follows directly from Lopez and Serven's result) and it would complicate the method. We thus prefer the simple log-normal model.

We will turn to the assumption that the individual ranks of households are the same in the income distribution as they are in the distribution of the asset index. Several studies have compared the asset index with income or expenditure data. Filmer and Pritchett (2001) calculate the correlation between the asset index and expenditures. They use three household surveys from developing countries that have information on both assets and expenditures. They find correlations coefficient between 0.43 and 0.64. Filmer and Pritchett (2001) show that the asset index is internally coherent in the sense that it differs across poor, middle and rich households. Also Grimm et al. (2008, 2009) find a close relationship between the ranking of household income/expenditure and the asset index. Second, the asset index is robust to the assets included to derive the index. Third, the asset index performs well in comparing outcomes of poverty with outcomes based on income or expenditure.

However, the use of the asset index as a proxy for material welfare (i.e. income or expenditures) requires some careful discussion. First, the asset index might not correctly reveal differences between urban and rural areas. The asset index can be biased due to differences in prices and the supply of such assets as well as differences in preferences for assets between both areas. For example, urban households typically own more (and other) assets than rural households. Second, by definition, the asset index is a discrete function. This discreteness complicates the estimation of a (continuous) distribution. The extent of this problem depends on the number of assets incorporated in the asset index (Howe et al, 2009; McKenzie, 2005). In our case, similar to McKenzie (2005), we incorporate at least 25 assets into the index resulting in a more continuous character the distribution.

Stewart and Simelane (2005) validate the use of the asset index as a proxy for income to predict child mortality in South Africa. They find a high correlation between the asset index and income. In a recent paper, Filmer and Scott (2008) find that the gradient in education, health care, fertility and child mortality is very similar for the asset index and expenditures. They also find some differences in the ranking of households between the asset index and per capita expenditures, in particular for the lowest quintile. This is because asset indices are less suitable for capturing transitory effects. Filmer and Scott (2008) argue that targeting households on the basis of the asset index would only partly reach the same households as targeting based on income or expenditures – in particular in the lowest quintile. They found that the asset index is more likely to identify rural and smaller households as deprived than per capita expenditures. They conclude that welfare rankings based on an asset index do not lead to the exact same welfare rankings as per capita expenditure, but the gradient is quite similar for both measures.

Therefore, the asset index should not be seen as an alternative measure of income or expenditure. But, in the absence of information on household income or expenditure, the ownership of assets can be used to approximate the household's living standard. In fact, one can even argue that the ownership of assets might be a better proxy for long-term income than current income, because it is less vulnerable to fluctuations over time than income or expenditure. In addition, assets are less vulnerable to measurement error based on underreporting, c.f. McKenzie (2005). We conclude that the household ranking for the asset index is not the same as the household ranking for income, but they are similar enough for our purpose.

5 Conclusion

In this paper, we propose a simple and intuitive approach to simulate household income based on a uni-dimensional asset index for household survey data, where no information on income or expenditure is available. This approach allows us to calculate any kind of poverty and inequality measures within and across countries, over time, and also by population subgroups and by socioeconomic characteristics. In a first best data world, there would be no use for our approach. It is born out of the unfortunate situation that the DHS data sets, which potentially could be a rich resource for analyses of inequality in sub-populations, do not contain information on household income or expenditure. We believe that the main advantages of our approach are its simplicity and transparency. It requires two main assumptions, and every researcher can decide for himself or herself whether or not these assumptions are appropriate for the research question at hand. The previous discussion has shown that our approach approximates household income quite well – at least under the given the data restrictions. We thus hope that this paper will help to improve our understanding of economic inequality in developing countries. However, we have to admit that we can only provide an approximation of household income, thus caution has to be applied if very precise estimates of income and inequality are needed.

References

Ainsworth, M. and D. Filmer (2006), Children's Schooling: AIDS, Orphanhood, Poverty, and Gender, World Development 34(6):1099-1128.

Atkinson, A.B. (1970), On the measurement of inequality, Journal of Economic Theory, 2 (3): 244-263.

Bicego, G., S. Rutstein, and K. Johnson (2003), Dimensions of the Emerging Orphan Crisis in Sub-Saharan Africa, Social Science and Medicine 56(6):1235-1247.

Bollen, K. A., J. L. Glanville, and G. Stecklov (2002), Economic Status Proxies in Studies of Fertility In Developing Countries: Does the Measure matter?, Population Studies 56(1):81-96.

Deaton, A. (1997), The Analysis of Household Surveys. Washington, DC: John Hopkins University Press.

Grimm, M., K. Harttgen, M. Misselhorn, and S. Klasen (2008). A Human Development Index by Income Groups. World Development, 36 (12): 2527-2546.

Grimm M., K. Harttgen, M. Misselhorn, T. Munzi, S. Klasen, and T. Smeeding (2009). Inequality in Human Development: An empirical assessment of thirty-two countries. Social Indicator Research DOI 10.1007/s11205-009-9497-7.

Filmer, D. and L.H. Pritchett (2001), Estimating Wealth Effects without Expenditure Data -or Tears: An Application to Educational Enrollments in States of India, Demography, 38 (1): 115-132.

Filmer, D. and K. Scott (2008), Assessing Asset Indices, World Bank Policy Research Working Paper No. 4605, World bank.

Harttgen, K. and M. Misselhorn (2006), A Multilevel Approach to Explain Child Mortality and

Undernutrition in South Asia and Sub-Saharan Africa, Ibero-America Institute Discussion Paper No. 152.

Houwelling T.A.J., A. E. Kunst, and J. P. Mackenbach (2003), Measuring Health Inequality among Children in Developing Countries: Does the Choice of the Indicator of Economic Status Matter?, International Journal for Equity in Health 2(8).

Howe, L.D., J.R. Hargreaves, S. Gabrysch, and S.R.A. Huttly (2009), Is the wealth index a proxy for consumption? A systemic review, J Epidemiol Community Health, 63: 871-880.

Lopez, H. and L. Serven (2006), A normal relationship? Poverty, inequality and growth. World Bank Policy Research Working Paper 3814.

McKenzie, D.J. (2005), Measuring inequality with asset indicators, Journal of Population Economics, 18: 229-260.

Montgomery, M.R., M. Gragnolati, K. Burke, and E. Paredes (1999), Measuring Living Standards with Proxy Variables, Policy Research Division Working Paper No. 129, New York University, New York.

Sahn, D.E. and D.C. Stifel (2003), Exploring Alternative Measures of Welfare in the Absence of Expenditure Data, Review of Income and Wealth, 49 (4): 463-489.

Sahn, D.E. and D. C. Stifel (2001), Poverty Comparisons Over Time and Across Countries in Africa, World Development, 28 (12): 2123-2155.

Sala-i-Martin, X. (2006), The World Distribution of Income: Falling Poverty and … Convergence, Period. Quarterly Journal of Economics 121(2): 351-397.

Sastry, N. (2004), Trends in Socioeconomic Inequalities in Mortality in Developing Countries: The case of Child Survival in Sao Paulo, Brazil, Demography 41(3): 443-464.

Schellenberg, J. A., C. G. Victora, A. Mushi, D. de Savigny, D.Schellenberg, H. Mshinda, and J. Bryce (2003), Inequities among the Very Poor: Health Care for Children in Rural Southern Tanzania, The Lancet 361(9357): 561-566.

Stewart, T. and S. Simelane (2005), Are Assets a Valid Proxy for Income? An Analysis of Socioeconomic Status and Child Mortality in South Africa, Indiana University, mimeo.

Stifel, D. and L. Christiaensen (2007), Tracking Poverty Over Time in the Absence of Comparable Consumption Data, The World Bank Economic Review 21(2):317-341.

Tarozzi, A. and A. Mahajan (2005), Child Nutrition in India in the Nineties: A Story of Increased Gender Inequality?, Unpublished manuscript. Duke University and Stanford University.

Vollmer, S., H. Holzmann and J. Weisbrod (2010), A Global Growth Incidence Analysis. Working paper. Available at: http://works.bepress.com/sebastianvollmer/

World Bank (2006), World Development Report 2007: Development and the Next Generation, The World Bank and Oxford University Press. Washington, DC.

Table 1: Data Sources

Country	Year	Type of Survey
Bolivia	2003	Demographic and Health Survey (DHS)
Burkina Faso	2003	Demographic and Health Survey (DHS)
Indonesia	2003	Demographic and Health Survey (DHS)
Zambia	2002	Demographic and Health Survey (DHS)
Data on mean income and Gini		PovCalNet
Data on Purchasing Power parities (PPP)		International Comparison Program (ICP)

Table 2: Scoring factors of asset index

	Bolivia			Burkina Faso			Indonesia			Zambia		
Asset	**Scoring factor**	**Mean**	**SD**	**Scoring factor**	**Mean**	**SD**	**Scoring factor**	**Mean**	**SD**	**Scoring factor**	**Mean**	**SD**
Own radio	0.254	0.835	0.371	0.293	0.688	0.463	0.293	0.688	0.463	0.454	0.444	0.497
Own television	0.747	0.617	0.486	0.791	0.126	0.332	0.791	0.126	0.332	0.788	0.187	0.39
Own bicycle	0.161	0.512	0.500	-0.092	0.850	0.357	-0.092	0.850	0.357	-0.011	0.367	0.482
Own car	0.419	0.131	0.337	0.556	0.034	0.181	0.556	0.034	0.181	0.384	0.028	0.166
Own motocylce	0.162	0.057	0.231	0.539	0.303	0.460	0.539	0.303	0.460	0.159	0.006	0.074
Own refrigerator	0.574	0.172	0.377	0.700	0.048	0.213	0.654	0.036	0.186	0.543	0.035	0.185
Own phoine	0.678	0.301	0.458	0.654	0.036	0.186	0.700	0.048	0.213	0.778	0.095	0.293
Main source of lightening is electric	0.728	0.712	0.453	0.810	0.114	0.318	0.810	0.114	0.318	0.867	0.158	0.365
Drinking water from piped source	0.661	0.716	0.451	0.679	0.064	0.245	0.679	0.064	0.245	0.812	0.141	0.348
Drinking water from open source	-0.356	0.116	0.321	0.093	0.040	0.195	0.093	0.040	0.195	-0.201	0.469	0.499
Drinking water from other source	-0.433	0.111	0.314	-0.440	0.527	0.499	-0.440	0.527	0.499	-0.388	0.39	0.488
Flush toilet	0.601	0.294	0.456	0.302	0.014	0.119	0.302	0.014	0.119	0.823	0.136	0.343
Pit latrine	-0.103	0.293	0.455	0.099	0.086	0.281	0.099	0.086	0.281	-0.267	0.57	0.495
No toilet facility	-0.579	0.338	0.473	-0.676	0.693	0.461	-0.676	0.693	0.461	-0.361	0.28	0.449
Dwelling has low quality of floor material	-0.789	0.368	0.482	-0.662	0.587	0.492	-0.662	0.587	0.492	-0.787	0.665	0.472
Dwelling has high quality of floor material	0.789	0.632	0.482	0.662	0.413	0.492	0.662	0.413	0.492	0.787	0.335	0.472
Dwelling has low quality of wall material	-0.690	0.646	0.478	na	na	na	na	na	na	na	na	na
Dwelling has high quality of wall material	0.690	0.354	0.478	na	na	na	na	na	na	na	na	na
Dwelling has low quality of roof material	-0.365	0.641	0.480	na	na	na	na	na	na	na	na	na
Dwelling has high quality of roof material	0.251	0.034	0.181	na	na	na	na	na	na	na	na	na
Percentage of the covariance explained by the first principal component	0.301			0.312			0.28			0.351		
Eigenvlaue of first principal component	6.02			4.992			5.592			5.62		

Source: Demographic and Health Surveys (DHS); calculations by the authors.

Table 3: Summary of Macroeconomic Data

PovCalNet Input Data				
Country	**Survey year**	**Year**	**Means $**	**Gini(%)**
Bolivia	2002, 2003	2002	185.99	60.24
Burkina Faso	2002, 2003	2003	46.85	39.6
Indonesia	2002, 2003	2002	61.83	30.39
Zambia	2002	2002	41.07	42.08

Source: PovCalNet and International Comparison Program (ICP).

Table 4: Inequality Decomposition – Bolivia

	Population Share	Income per capita (PPP)	Gini	Atkinson e0.5	Atkinson e1	Atkinson e2
Full Sample	100.0%	188.41	0.610	0.307	0.519	0.758
Rural	40.2%	55.41	0.483	0.191	0.340	0.552
Urban	59.8%	277.74	0.533	0.232	0.400	0.622
Between				0.103	0.208	0.373
Within				0.227	0.393	0.613
Female Head	16.5%	194.58	0.575	0.271	0.475	0.728
Male Head	83.5%	187.185	0.616	0.314	0.527	0.762
Between				0.000	0.001	0.004
Within				0.307	0.518	0.757
Household Size 0-4	36.0%	191.273	0.579	0.277	0.483	0.739
Household Size 5-8	53.0%	195.05	0.628	0.326	0.542	0.771
Household Size 8+	11.0%	146.84	0.606	0.298	0.508	0.736
Between				0.002	0.004	0.006
Within				0.306	0.517	0.756
No Education	8.7%	70.63	0.559	0.261	0.441	0.664
Primary	55.2%	119.67	0.568	0.267	0.455	0.686
Secondary	24.4%	235.96	0.517	0.219	0.390	0.630
Higher	11.7%	508.27	0.506	0.210	0.382	0.634
Between				0.094	0.182	0.303
Within				0.234	0.412	0.652

Table 5: Inequality Decomposition – Burkina Faso

	Population Share	Income per capita (PPP)	Gini	Atkinson e0.5	Atkinson e1	Atkinson e2
Full Sample	100.0%	41.58	0.410	0.135	0.249	0.423
Rural	78.6%	28.67	0.305	0.075	0.148	0.286
Urban	21.4%	88.93	0.291	0.069	0.135	0.265
Between				0.068	0.125	0.202
Within				0.072	0.142	0.277
Female Head	6.0%	57.84	0.419	0.141	0.273	0.483
Male Head	94.0%	40.55	0.406	0.133	0.245	0.415
Between				0.002	0.003	0.003
Within				0.134	0.247	0.421
Household Size 0-4	14.7%	40.97	0.421	0.142	0.261	0.440
Household Size 5-8	36.5%	40.18	0.432	0.150	0.272	0.447
Household Size 8+	48.8%	42.81	0.389	0.122	0.227	0.395
Between				0.000	0.002	0.005
Within				0.135	0.248	0.420
No Education	84.2%	34.49	0.364	0.107	0.202	0.360
Primary	9.7%	56.64	0.356	0.102	0.201	0.383
Secondary	5.2%	102.37	0.285	0.067	0.136	0.280
Higher	1.0%	175.48	0.255	0.051	0.099	0.189
Between				0.040	0.073	0.117
Within				0.099	0.189	0.346

Table 6: Inequality Decomposition – Indonesia

	Sample Size	Income per capita (PPP)	Gini	Atkinson e0.5	Atkinson e1	Atkinson e2
Full Sample	100.0%	63.28	0.307	0.074	0.144	0.267
Rural	55.7%	47.14	0.262	0.054	0.107	0.204
Urban	44.3%	83.54	0.267	0.056	0.110	0.211
Between				0.020	0.039	0.074
Within				0.055	0.109	0.208
Female Head	8.3%	62.20	0.307	0.074	0.144	0.268
Male Head	91.7%	63.38	0.307	0.074	0.144	0.267
Between				0.000	0.000	0.000
Within				0.074	0.144	0.267
Household Size 0-4	38.5%	59.90	0.294	0.069	0.133	0.249
Household Size 5-8	52.7%	65.12	0.312	0.077	0.149	0.276
Household Size 8+	8.9%	67.00	0.317	0.079	0.153	0.284
Between				0.000	0.001	0.001
Within				0.074	0.143	0.267
No Education	8.8%	44.70	0.261	0.054	0.107	0.205
Primary	49.2%	51.47	0.265	0.056	0.110	0.212
Secondary	34.6%	74.29	0.283	0.064	0.125	0.241
Higher	7.4%	112.62	0.241	0.047	0.095	0.193
Between				0.018	0.034	0.060
Within				0.058	0.114	0.221

Table 7: Inequality Decomposition – Zambia

	Sample Size	Income per capita (PPP)	Gini	Atkinson e0.5	Atkinson e1	Atkinson e2
Full Sample	100.0%	43.87	0.416	0.138	0.260	0.453
Rural	71.5%	29.50	0.333	0.091	0.177	0.334
Urban	28.5%	79.92	0.307	0.079	0.158	0.325
Between				0.059	0.111	0.184
Within				0.085	0.167	0.329
Female Head	19.5%	36.42	0.403	0.129	0.240	0.409
Male Head	80.5%	45.67	0.415	0.138	0.261	0.459
Between				0.002	0.003	0.003
Within				0.137	0.258	0.451
Household Size 0-4	22.7%	36.18	0.403	0.130	0.241	0.414
Household Size 5-8	53.0%	43.04	0.420	0.141	0.263	0.454
Household Size 8+	24.3%	52.88	0.394	0.125	0.244	0.451
Between				0.005	0.009	0.013
Within				0.134	0.253	0.446
No Education	15.0%	25.71	0.337	0.094	0.176	0.315
Primary	51.1%	32.58	0.347	0.098	0.190	0.357
Secondary	26.4%	58.37	0.350	0.101	0.205	0.412
Higher	7.5%	105.41	0.308	0.077	0.156	0.336
Between				0.048	0.088	0.132
Within				0.095	0.188	0.369