Top Incomes and the Measurement of Inequalityin Egypt

Vladimir Hlasny (Ewha Womans University, Seoul and UN-ESCWA)
Paolo Verme (World Bank)

Paper prepared for the IARIW-CAPMAS Special Conference “Experiences and Challenges in Measuring Income, Wealth, Poverty and Inequality in the Middle East and North Africa”

Cairo, Egypt
November 23-25, 2015

Session 5: Inequality I
Tuesday, November 24, 2015
08:30-10:00

Top Incomes and the Measurement of Inequality in Egypt

Vladimir Hlasny[1] and Paolo Verme[2]

Abstract

While the Egyptian revolution was partly motivated by claims of inequality, data show that income inequality was low and declining during the period that led to the 2011 revolution. This paper exploits unprecedented access to income data and a combination of newly developed statistical methods to evaluate income inequality in Egypt. Correcting for top incomes biases increases the Gini coefficient by just over one percentage point, a finding robust to several tests and methods. Also, Egyptian top incomes follow nicely the Pareto distribution and do not show any anomalies compared to surveys worldwide. These findings confirm that the pre-revolution income inequality in Egypt was low by regional and world standards and is not a good candidate to explain the Egyptian revolution.

JEL: D31, D63, N35.

Keywords: Top incomes, inequality measures, survey nonresponse, Pareto distribution, parametric estimation, Egypt.

[1] Ewha Womans University, Seoul

[2] The World Bank. The authors are grateful to Branko Milanovic and Francisco Ferreira for peer reviewing the work, Johan Mistiaen for useful advice, the CAPMAS of Egypt for providing access to the full 2008/2009 Household Income, Expenditure and Consumption Survey (HIECS) on site in Cairo, to Olivier Dupriez for estimating Pareto coefficients and Ginis using a World Bank harmonized data set of household income and expenditure data and to participants to a World Bank internal review meeting and a public seminar that took place in June and October 2013 respectively. Preliminary results were presented in Cairo during a workshop organized with the Egyptian Social Contract Center. The authors are grateful to the panel of Egyptian economists and statisticians that provided very rich comments during the workshop including Professors Heba El-Laithy, Sherine Al-Shawarby and Sahar El-Tawila. All remaining errors are ours.

1. Introduction

The question of income inequality has recently been at the forefront of the public discourse in the US and elsewhere. Joseph Stiglitz's recent book on inequality (2012) makes a convincing case that income inequality is a major impediment for future prosperity in the US. Paul Krugman recently argued that "(…) *if you take a longer perspective, rising inequality becomes by far the most important single factor behind lagging middle-class incomes*"[3], and President Obama stated that income inequality "*(…) is the defining challenge of our time.*"[4] This new stand on inequality finds its roots in concrete facts about inequality including three decades of increasing income inequality and a major financial crisis that had profound economic consequences for the middle-class in the US.

Beyond the US, research on top incomes in OECD countries based on tax records has shown that the last decades have seen an unprecedented rise in the shares of top incomes over total incomes (see Atkinson et al., 2011 for a review of this literature). Thomas Piketty (2014) built on these findings to propose a theory of capital that would explain historical trends in top incomes shares and the remarkable reception of his book has shown how this topic ranks high in people's minds globally. The evidence on top incomes and income inequality in the developing world is less rich but there is little doubt that these topics are important for people in this part of the world as well. One example is the Middle East and North-Africa (MENA) region where the issues of inequality and social justice became central themes of the 2011 revolutions and continue to remain widely debated in the region to this day.

In this global scenario, Egypt remains an anomaly if we distinguish factual and perceived inequality. While inequality and social justice have been two themes extensively debated since the start of the 2011 revolution, household data indicate that income inequality was neither high nor increasing during the years that led to the revolution. A recent study of inequality in Egypt (World Bank 2014) has confirmed that there is an important discrepancy between income inequality as measured by household surveys and the perception of income inequality as reported by people in values surveys. The study concluded that there are very good reasons to believe

[3] New York Times (December 15th, 2013). http://www.nytimes.com/2013/12/16/opinion/krugman-why-inequality-matters.html?ref=paulkrugman&_r=0
[4] http://www.usatoday.com/story/news/politics/2013/12/04/obama-income-inequality-speech-center-for-american-progress/3867747/

both household survey data and the people of Egypt and offered some leads that could explain the apparent mismatch including the absence of progress in terms of living standards during a period of sustained growth for the economy.

But mistrust for data continues to run high among observers. In a press article featured in the Al Shorouk newspaper in May 2014,[5] Egyptian scholars argued that believing that inequality in Egypt was low before the revolution amounted to supporting the former regime. A working paper by Alvaredo and Piketty (2014) argued that household data in Egypt should not be trusted because top incomes cannot be properly measured and proposed a parametric approach to re-estimate inequality. This yielded higher inequality estimates.[6]

Is income inequality poorly measured in Egypt because of the exclusion or biased inclusion of top incomes in household surveys? Measuring top incomes well is not easy and is a known challenge among economists and survey specialists. Household surveys are not particularly accurate at measuring top incomes because richer households are less likely to participate in household surveys or their incomes are mismeasured. It is therefore essential to understand how unit non-responses and nonrepresentative top incomes affect the measurement of income inequality in any particular survey.

This paper focuses on this question using two different approaches recently proposed to correct inequality measures for biases generated by top incomes. The joint use of the two approaches, the sensitivity analysis of their technical specifications and the joint performance of the two approaches are methodological innovations of this study. The first approach (Mistiaen and Ravallion, 2003; Korinek et al. 2006 and 2007) develops a probabilistic model to detect whether an income bias exists due to unit non-responses, and a re-weighting mechanism to correct for such a bias. Departing from previous studies, we will use information on unit non-responses at the PSU level rather than more aggregated levels. We will discuss implications of this choice and we will show that this level of aggregation is better suited to our data. The second approach (Cowell and Victoria-Feser, 1996, Cowell and Flachaire, 2007, and Davidson and Flachaire, 2007) tests how individual top income observations may affect measures of inequality and

[5] Al Shorouk (May 2014).
[6] In essence, the paper combines a Pareto distribution for top incomes with a lognormal distribution for the rest of incomes and creates an artificial distribution that is then used to measure inequality. The authors quote an earlier version of our paper and borrow our top incomes corrected Gini as one of the parameters used in the parametric estimation (see Appendix 3). This approach does not rely on any new data and outcomes rely on assumptions made about the parameters used.

combines a parametric estimation for the uppermost part of the income distribution and a non-parametric method for the rest of the distribution to correct for possible top incomes biases. As a robustness check, the weights derived from the first approach will also be applied to the distributions produced by the second approach. Sensitivity tests and comparisons with other national surveys are carried out for both approaches.

The paper finds a consistent and significant underestimation of the Gini inequality measure due to unit non-responses. The degree of underestimation is estimated at around 1.3 percentage points (confidence interval 1 to 2 percentage points), a finding robust to the different top incomes correction approaches proposed and to the use of income or expenditure measures. Applying this correction to the Egyptian Gini shows that the top incomes distribution of the Egyptian data follow rather closely the Pareto distribution and that inequality in Egypt remains low by regional and world standards.

These findings are important for Egypt today. If income inequality is effectively high and growing, the government of Egypt should be called to prioritize redistribution policies. But if inequality is low and declining, this may simply be a sign of widespread misery, low and stagnant income opportunities, low labor demand and ineffective markets. In this case, the focus of the government should be better placed on investments, inclusive growth measures, improving jobs and income opportunities, and better allocation of existing resources rather than a simple redistribution of government revenues. Our findings suggest that income inequality in Egypt is more likely to characterize this second scenario.

The paper is organized as follows. The next section introduces the inequality correction models proposed by the literature and our applications of the models. Section three describes the Egyptian survey data and compares them with surveys worldwide. Section four presents and discusses results. Section five summarizes main findings and discusses policy implications.

2. Models

The measurement of income inequality can be affected by a variety of issues including measurement errors, data input errors, item non-response, unit non-response, extreme observations, the shape of the top income distribution, post-survey top coding, post-survey reweighting and others (Atkinson and Micklewright, 1983; Juster and Kuester, 1991; Jolliffe, Datt and Sharma, 2004). The quality of the Egyptian household survey data has been assessed by the World Bank (2014), a study that reconstructed the welfare aggregates and carried out a number of quality tests on the data related to income and expenditure. The study did not find relevant issues related to measurement, data input and item non-response. The Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS) does not apply data modification methods such as top coding, imputation of values or trimming of sampling weights. However, unit non-response and extreme observations were found to be relevant issues, although no solutions were suggested or provided by the study to address these issues.

In what follows, we will use the same welfare aggregates constructed by the World Bank and focus on unit non-response, extreme values and the shape of the top income distribution, the three issues that have been recently studied by the literature described in the introduction. As for the previous literature, the Gini is our inequality measure of choice because of its popularity and because it is the most robust measure of inequality in the presence of extreme observations (Cowell and Victoria-Feser, 1996 and Cowell and Flachaire, 2007). We discuss the Gini correction for unit non-response and the correction for extreme values in this order.

Unit non-response

To test for the presence of a systematic non-response bias, we use a formal model to estimate the relationship between household income and the household’s probability of response. Unfortunately, unlike in the case of item non-response, we cannot simply infer households’ unreported income from their other reported characteristics, because we don’t observe any information for the non-responding households. Assigning the mean or median values to the missing items (even from within primary sampling units – PSUs) would be inappropriate, as the missing values may be systematically very different from the rest of the distribution.

Following Mistiaen and Ravallion (2003), and Korinek et al. (2006, 2007), we can use information about households’ response rates from across regions to infer the propensity of households with different characteristics, such as different incomes, to participate in the survey. This approach essentially takes advantage of the variation in household response rates and the variation in the distribution of observable variables (income or expenditure per capita) across geographical areas.

Assume that the survey probability of response of household *i*, P_i, is a logistic function of its arguments (Korinek et al. 2006, 2007):

$$P_i(x_i, \theta) = \frac{e^{g(x_i,\theta)}}{1 + e^{g(x_i,\theta)}}, \qquad (1)$$

where $g(x_i,\theta)$ is a stable function of x_i, the observable characteristics of responding households *i* that are used in estimations, and of θ, the corresponding vector of parameters from a compact parameter space. Variable-specific subscripts are omitted for conciseness. $g(x_i,\theta)$ is assumed to be twice continuously differentiable in θ.

Estimates of parameters θ affect each household's estimated probability of responding to the survey, and thus also the estimated number of households in sample design from which survey respondents are drawn. θ can be estimated by fitting the estimated and actual number of households in each region using the generalized method of moments (GMM) estimator

$$\hat{\theta} = \arg\min_{\theta} \sum_j \left[\left(\widehat{m}_j - m_j \right) w_j^{-1} \left(\widehat{m}_j - m_j \right) \right]. \qquad (2)$$

Here m_j is the reported number of households in region *j* according to sampling design, $\widehat{m}_j$ is its estimate conditional on the values of explanatory variables $x_i\ \forall i \in j$, and w_j is a region-specific analytic weight inversely proportional to the size of each region, m_j. The estimated number of households in sample design, $\widehat{m}_j$, can be imputed as the inverse of the estimated response probability of responding households in the region, $\hat{P}_{ij}$, summed over N_j responding households.[7]

$$\widehat{m}_j = s_j^{-1} \sum_{i=1}^{N_j} \hat{P}_{ij}^{-1}. \qquad (3)$$

[7] For instance, in the Egyptian HIECS, there are m_j=20 households in most PSUs *j* according to sampling design. Therefore, w_j=1/20. Across two PSUs with N_j=18 responding households each, we may estimate $\widehat{m}_1$=19.5 and $\widehat{m}_2$=20.5 because the latter PSU has higher values of $x_i\ \forall i \in j$ (such as higher incomes). Consistency of the estimator $\hat{\theta}$ for the true θ requires, among other things, that the incomes of the two nonresponding units in each region come from among the observed incomes of responding units in those regions.

Under the assumptions of random sampling within and across regions, inclusiveness of the observed values of explanatory variables x in each region with regard to all values of x of the nonresponding units, and stable functional form of $g(x_i,\theta)$ for all households across regions, the estimator $\hat{\theta}$ is consistent for the true θ. Estimated values of $\hat{\theta}$ that are significantly different from zero would indicate a systematic non-response bias as a function of variables x. In that case, we can use the imputed household response probabilities to correct for the bias. This entails imputing the true distribution of incomes by correcting the mass of each observation for its probability of being sampled. Inverses of the estimated response probabilities serve as the appropriate household weights. In the income distribution imputed in this way, the derived measures of inequality converge to their true values as long as the values of x of the nonresponding units are among the observed values of x for responding units in each region. This essentially requires a sufficient sample size in each region so that the unobserved incomes of the nonresponding units are among the values of incomes of responding units.

The model presented in equations 1-3 above uses *within-j* information as well as *between-j* information. It uses within-*j* information because the estimated number of households $\hat{m}_j$ is estimated within-*j* and it uses between-*j* information because the proportion of households observed within-*j* and the distribution of explanatory variables vary across j_s. Therefore, the choice of geographic disaggregation can affect the results, and involves a trade-off between the number of *j* data points, and the number and distribution of within-*j* observations vis-à-vis the characteristics of the nonresponding units. On the one hand, observations should be behaviorally similar to non-responding households within-*j*, and to households with similar values of explanatory variables in surrounding regions *j*, calling for smaller geographic areas. On the other hand, Equation 3 requires that the sample of respondents encompasses the entire range of values of explanatory variables of nonrespondents, potentially calling for larger geographic areas. In this paper we opt to use 2,526 PSUs as *j* regions with an average of 18.6 responding households per region, as compared to the 51 US states with an average of 1,649 households per state used by Korinek et al. (2006, 2007). While nonresponse rates are available consistently only at the state level in the US CPS, they are available even at the level of PSUs in the HIECS.

These are clearly two different approaches with different implications in at least three different respects: 1) In our case, the primary sampling units have relatively homogeneous households, with similar behavioral responses and presumably also similar survey-response probabilities. Because of the high response rate to the HIECS survey (96.3%), the observed range of incomes in each PSU is expected to contain the values of the few non-responding households. A higher level of geographic aggregation as in Korinek et al. (2006 and 2007) would make behavioral responses less likely to be stable within *j* areas, while offering little additional assurance that values of characteristics of responding households encompass values of non-responding units.

2) In our case, households' response probabilities are essentially inferred by comparing regions with similar, narrow ranges of explanatory variables. The response probability curve is constructed using 2,526 sets of probability estimates that are little overlapping on the curve. In Korinek et al.'s (2006, 2007) case, response probabilities are inferred by comparing fewer regions with greater heterogeneity of units within them. The response probability curve is constructed using 51 sets of probability estimates largely overlapping.

Put another way, because income inequality across large areas tends to dominate over inequality within regions, PSUs with similar ranges of incomes tend to be located near each other. As a result, two groups of households with similar ranges of incomes can be viewed as geographically related, be behaviorally similar and have a similar probability of survey response even if they reside in different PSUs. It is these groups of households whose frequencies of income observations are compared against each other to estimate response probability for their income level. This is appropriate because the two groups of households are behaviorally similar. But this would be difficult to argue in less regionally disaggregated samples, where two groups of households with similar income ranges may come from very different regions and different portions of their local income distributions and may be very different from each other behaviorally. Finally, because there are fewer regions *j* in less regionally disaggregated samples, fewer errors are estimated in Equation 2, and each cross-region comparison of frequencies of incomes and response rates is effectively more influential to the ultimate results.

3) In our case, the non-response bias correction is limited by the low observed non-response rate and by homogeneity of households in each PSU, which prevent the response probabilities to be estimated too low. In Korinek et al. (2006, 2007) case, response probabilities can be estimated very low for some households, because other households in the same region can be assigned very high probabilities in compensation.

This difference in methodologies is important because model errors are at the level of regions j. We think that our approach represents a more appropriate bias correction of the Gini coefficients in the HIECS data, that it is less likely to overshoot the correction, and that it is more consistent with the Pareto corrections illustrated in the next section. We will test these claims in the results section.

Extreme values

Measures of inequality can be influenced by the presence of even few observations with unusually high values that may not represent well the underlying population. In this section, we follow a procedure proposed by Cowell and Flachaire (2007) and Davidson and Flachaire (2007) to replace highest-income observations with values estimated under an expected distribution, and to combine the corresponding parametric inequality measure for these incomes with a non-parametric measure for lower incomes. Cowell and Flachaire (2007) consider a parametric Pareto distribution of the form

$$f(x) = \frac{\alpha}{x^{\alpha+1}}\ , 1 \leq x \leq \infty\ . \tag{4}$$

with the following formulation of α:

$$\alpha = \frac{1}{k^{-1}\sum_{i=0}^{k-1} \log X_{(n-i)} - \log X_{(n-k+1)}}\ , \tag{5}$$

where $X_{(j)}$ is the j_{th} order statistic in the sample of incomes n, and k is the delineation of top incomes such as the top 10% of observations. We can also estimate α using maximum-likelihood methods to obtain the estimate with its robust standard error. This method of parametric estimation of the income distribution still allows weighting of observations by their sampling probability. The Gini coefficient under the estimated Pareto distribution for the k top-income households can be derived from the expression of the corresponding Lorenz curve (expression inside the integral below) as

$$Gini = 1 - 2\int_0^1 1 - [1 - F(x)]^{1-1/\alpha}\, dF(x) = \frac{1}{2\alpha - 1}\ . \tag{6}$$

This parametric Gini coefficient can then be combined with the non-parametric Gini coefficient for the n-k lower-income observations using geometric properties of the Lorenz curves as

$$Gini = (1 + Gini_k)\frac{k}{n}s_k - (1 - Gini_{n-k})\left(1 - \frac{k}{n}\right)(1 - s_k) + \left(1 - \frac{2k}{n}\right). \qquad (7)$$

Here s_k refers to the share of aggregate income held by the richest k percent of households. As long as it was correct to assume that top incomes in the population are Pareto-distributed, this semi-parametric Gini coefficient can be compared to an uncorrected non-parametric estimate for the observed income distribution. A difference between the semi-parametric and non-parametric estimates would indicate that some observed high incomes may have been generated by a statistical process other than Pareto, and that our inequality measure is sensitive to this. A semi-parametric Gini that is lower than the non-parametric Gini can be interpreted as evidence that some top incomes in the sample are 'extreme' compared to those predicted under the Pareto distribution. A higher semi-parametric Gini would indicate that the observed top incomes are lower than predicted by the Pareto distribution, potentially implying under-representation of high-income units in the sample.

The unit non-response and extreme-observations conjectures thus yield opposite predictions about the influence of top incomes on inequality measures, to the extent that they may even cancel each other out. The former conjecture is that the observed top incomes are valid for the measurement of inequality, and should be even used to stand for unobserved incomes of non-responding households. The latter conjecture is that the observed top incomes may have been generated by processes different from those in the underlying population, by error or by different accounting practices, and should be replaced by values imputed from the data generating process in the population.

To comment on the validity of these opposite predictions and evaluate their relative significance, we can compare a set of four Gini coefficients: semi-parametric Gini accounting for the possibility of extreme observations but not for the non-response bias (i.e., Equation 7 where α and $Gini_k$ are derived from Equations 5 and 6 using an unweighted income distribution); semi-parametric Gini accounting for them both (with α and $Gini_k$ derived from Equations 5 and 6 in an income distribution weighted by households' response probability); non-parametric Gini correcting only for the non-response bias (Gini observed in a response-probability weighted income distribution); and the baseline uncorrected non-parametric Gini (observed in the unweighted income distribution). This comparison can inform us about the relative importance of extreme income observations versus non-response bias among high-income households and about their combined effect on the measurement of inequality in Egypt.

3. Data

This study relies on the Household Income, Expenditure and Consumption Survey (HIECS) administered by the Egyptian Central Agency for Public Mobilization and Statistics (CAPMAS). In this study, we use the 2008/2009 round of the HIECS. The sample comprises four quarterly independent subsamples of 12,000 households each that are nationally representative and stratified by governorate and urban and rural areas. The 48,000 households are selected via a multi-stage random process from a master sample constructed from the 2006 population census. The HIECS is one of the largest household surveys in the world and has been used extensively for the study of poverty and living standards in Egypt. For a full description of the data and for a discussion of comparability issues over time see World Bank (2014).

The CAPMAS has traditionally provided access to only 25% of the surveyed observations and it has recently posted on its website 50% of the sample. World Bank (2014) has shown that the analysis of such subsamples can potentially lead to a 2-3 percent overestimation or underestimation of the Gini coefficient.[8] However, for the benefit of this study, the CAPMAS has granted exceptional access to the 100% sample of the 2008/2009 data at the statistical agency's office in Cairo. Moreover, the 2008/2009 HIECS is also the only survey where household response rates were systematically collected for all PSUs, which is essential to implement some of the tests conducted in this paper. Improvements to the sampling and methodology made by the CAPMAS between 2000 and 2009, the fact that the 2008/2009 sample is the closest to the 2006 population census, and the fact that the following surveys occurred during or soon after the 2011 revolution and on smaller samples, also make this survey the most reliable household survey ever conducted in Egypt. As we will show, this survey compares favorably in terms of coverage and data quality even to surveys in industrialized countries.

[8] A Monte Carlo experiment on our data – extracting 25% or 50% of observations randomly from the 100% sample of the 2005 round of the HIECS – reveals a similar result (available on request). Gini coefficients estimated in individual extractions differ from the true Gini by up to 1.5 percentage points.

The main welfare measures used in this paper are income and expenditure per capita. It is common practice in developing countries to use expenditure or consumption as a proxy of income rather than income itself given that income tends to be underreported and given that consumption is smoother than income, especially in rural areas. The World Bank (2014) report on income inequality in Egypt and our own work found the income variable in the HIECS to be rather good. The distribution of income is very similar in shape to that of consumption while the central moments of the distribution of income are higher than those of expenditure. The difference between income and consumption (savings) is also an increasing linear function of income as one should expect. As a validation of our approach, we will use both income and expenditure throughout the paper.

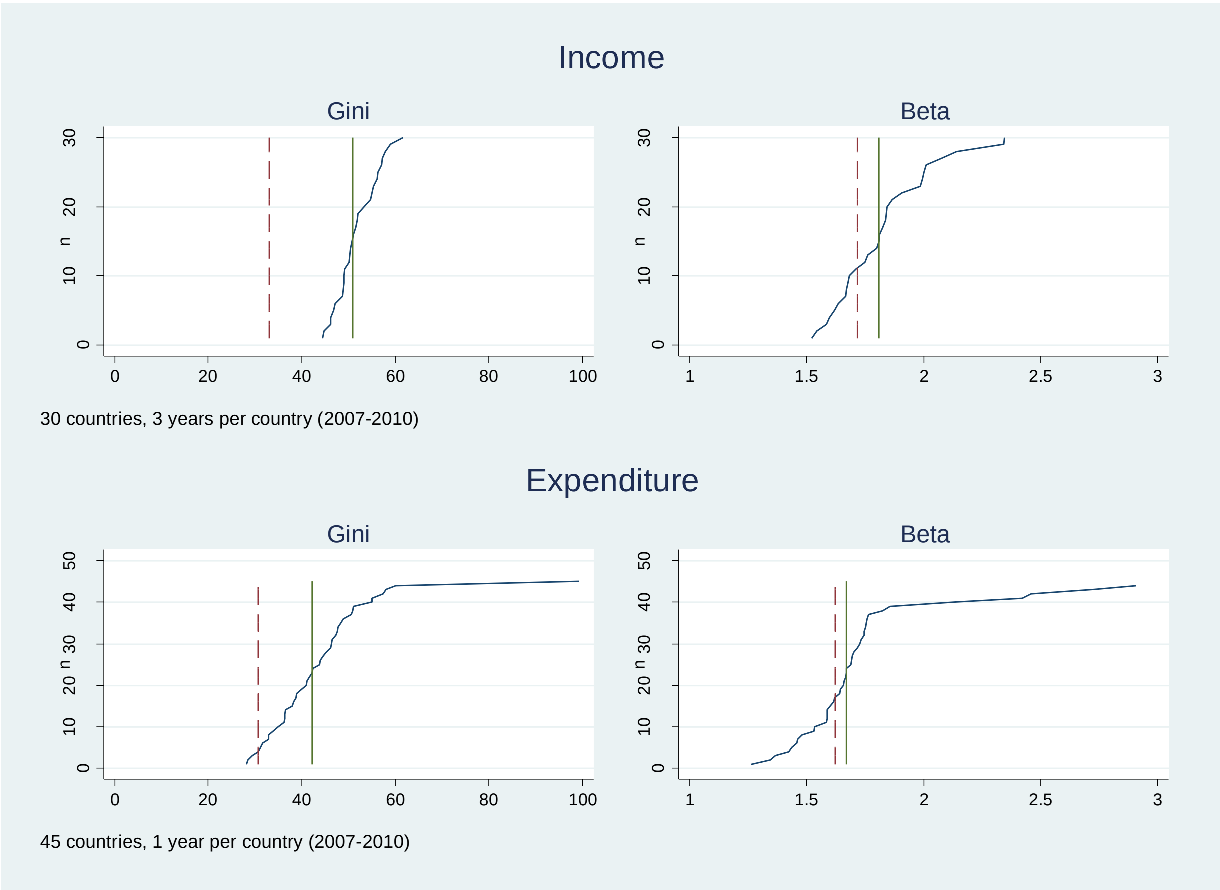
Our measure of income includes six main groups of items: wages and salaries, income from non-agricultural activities, cash transfers, income from agricultural activities, income from non-financial assets and income from financial assets. Our measure of expenditure includes food and beverages, alcoholic drinks and smokes, clothes, textiles and shoes, residence and its accessories, furniture, durables, health care and services, transportation, telecommunications, culture and entertainment, education, restaurants and hotels, various services and commodities. Both income and expenditure are used in per capita terms to avoid normative decisions about adult equivalence scales that could influence the measurement of inequality.[9]

If we compare the 2008/2009 Egyptian distributions of income and expenditure with those of other emerging economies, we find that the Gini coefficient for Egypt is very low while the inverted Pareto coefficient (a measure of thickness of the right-hand tail of the distribution) is very close to the median values of other countries. This is shown in Figure 1 where we compare the Egyptian Gini and Beta Pareto coefficient (dashed lines) with the respective median values (solid lines). This is done using a panel of Latin American countries for income and a panel of world emerging economies for expenditure.[10] For both income and expenditure, Egypt exhibits a Gini coefficient much lower than in other countries while the inverted Pareto coefficient is close to the median value of other countries, especially for expenditure. Results are robust to different

[9] For a more accurate description of the welfare aggregates and an analysis of their distributions see World Bank (2014).
[10] The data are extracted from a harmonized data set prepared by the World Bank research department and available to World Bank staff. The list of countries is provided in the annex, Table A1. Harmonized income data are only available for Latin American countries since most emerging economies tend to use consumption or expenditure rather than income as welfare measure. A larger set of harmonized data is available for expenditure.

choices of countries or years within the available data set. (Refer to Figure A1 in the annex for an expanded set of countries and years – "All" in the figure, and for comparisons with countries at similar levels of GDP per capita – "Selected" in the figure.) In essence and based on this initial data exploration, while the Egyptian Gini is low by world standards, the shape of the right-hand tail of the distribution is not atypical. Comparing the 2008/2009 round of the HIECS to the 1999/2000, 2004/2005 and 2010/2011 rounds also suggests that the right-hand tail in 2009 is typical of the surrounding years.

Figure 1. Gini and inverted Pareto coefficients for Egypt and the rest of the world



Source: HIECS 2008/2009. Dashed line=Egypt; Solid line=Median for the rest of the world.

4. Results

Unit non-response

Unit non-response is a problem in the HIECS data, particularly in some regions. Across governorates, the survey non-response rate in 2009 ranged from 0.0% to 10.5% with a mean of 3.7%. While the nationwide average non-response rate in the HIECS data is lower than in household surveys in other countries (for instance, refer to the literature surveyed in Korinek et al. 2006), it still leads to biases in statistics based on the observed sample. Out of 48,635 households contacted for the 2008/2009 survey, only 46,857 responded to the survey, while 1,778 reportedly did not respond. Secondly, the problem may be more serious in some governorates than in others, and so interregional demographic comparisons based on the sample may be flawed. Table 1 illustrates the differences in regional non-response rates and mean incomes of reporting households.

Table 1. Non-response rates and mean incomes and expenditures by governorate

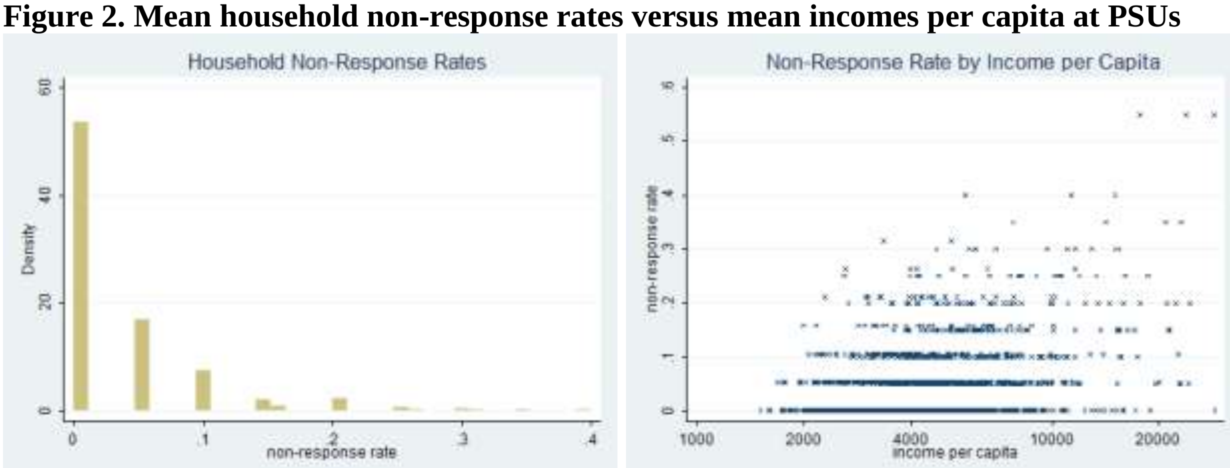
Governorate	PSUs	Households	Non-Response Rate (%)	Mean Income per Capita	Mean Expenditure per Capita
Alexandria	149	2,801	6.0	5,393.10	5,082.83
Assiut	101	1,872	2.4	2,665.06	2,216.75
Aswan	52	978	1.0	3,635.79	2,713.95
Behera	152	2,871	0.6	3,680.44	3,035.94
Beni Suef	69	1,294	1.3	2,887.36	2,557.90
Cairo	285	5,194	8.9	6,499.94	5,794.74
Dakahlia	176	3,289	1.6	4,467.94	3,768.32
Damietta	52	959	2.9	5,460.37	4,654.69
Fayoum	78	1,466	1.1	3,071.68	2,784.29
Gharbia	139	2,584	2.2	4,606.58	4,025.31
Giza	215	3,939	6.5	4,347.80	3,821.73
Ismailia	52	967	2.1	5,401.84	3,810.80
Kafr ElSheikh	85	1,547	4.2	4,279.37	3,497.10
Kalyoubia	145	2,668	3.2	4,137.20	3,642.90
Luxor	14	263	1.1	4,704.10	3,591.63
Matrouh	11	209	0.0	5,861.38	4,525.81
Menia	128	2,371	2.5	3,451.37	2,876.04
Menoufia	107	1,977	2.8	4,147.15	3,324.27
New Valley	8	146	3.9	5,322.18	4,458.13
North Sinai	14	243	10.5	3,768.41	2,829.52
Port Said	50	925	7.4	6,501.37	5,844.91
Qena	88	1,628	2.6	3,302.03	2,637.08

Red Sea	13	239	3.2	7,050.69	5,151.85
Shrkia	175	3,262	1.9	3,662.45	3,093.52
South Sinai	4	69	9.2	10,969.95	6,357.09
Suez	50	951	4.9	7,269.37	6,370.75
Suhag	114	2,145	1.0	2,809.37	2,391.82
Mean	94	1,735	3.7	4,653.03	3,974.44

Note: Nonresponse rate, reported in the survey at the PSU level, is weighted by the number of responding households in each PSU. Per-capita income and expenditure are further weighted by household size. These mean incomes and expenditures may not be representative of those for the entire governorates, as they omit non-responding households.

Unit non-response is associated positively with income of responding households across regions. Figure 2a reports that, at the level of individual PSUs, survey non-response rate ranges from 0.0% to 55% with a heavy right tail. Figure 2b shows the systematic relationship between household non-response rate and mean per-capita income of responding households at PSUs. Non-response rates greater than 33% occur only among the richest 25% of PSUs in terms of income per capita, and only among the richest 15% of PSUs in terms of expenditure per capita. Because of these findings, it is likely that mean incomes and expenditures are even higher in the underlying populations of regions with high non-response rates, and that the associations are even stronger with the incomes and expenditures of the underlying populations.

Figure 2. Mean household non-response rates versus mean incomes per capita at PSUs



(a) Histogram of mean non-response rates (b) Mean non-response rate by income per capita

Note: The unit of observation is a PSU. Average household non-response rate and average income per capita in a PSU are shown.

Table 2 shows the results of estimation of households' survey response as a function of household income or expenditure. Following Korinek et al.'s (2006, 2007) lead, all models estimate survey-response probability as a nonlinear function of income or expenditure. Models 1 and 2 make *g(x)* in Equation 1 a function of household income or expenditure. Models 3-10 use imputed income or expenditure per capita as explanatory variables, by dividing household-level variables by household size. Model specifications in Table 2 were selected in concurrence with Korinek et al.'s models, and with the aim to evaluate a variety of functional forms, from linear to highly non-linear.

The main finding is that households' survey response is related negatively to income and expenditures. The coefficients on income and expenditures are consistently negative and significant. The simplest univariate logarithmic functions are thought to be more robust, and exhibit better fit than more complex or polynomial functions. They yield greater significance of all coefficients, lower value of the minimization objective function, and lower values of the Akaike and Schwarz Information Criteria, implying more efficient overall model fit.

Household expenditure appears to have a better explanatory power than household income, yielding lower values of the Akaike and Schwarz Criteria. Income and expenditure per capita provide better fit than household-level income and expenditure, implying that per-capita variables are more predictive of householders' decision to respond than the household-level equivalents, without introducing additional noise into the model.

The negative relationship between income and response probability is particularly strong at high incomes. The same is true for expenditures. The estimated relationship is highly nonlinear, with the response rate dropping rapidly in the highest range of expenditures. Models using linear, quadratic or polynomial functions (such as square-root or cubic-root of expenditures) rather than logarithmic functions achieve inferior measures of fit. Linear, quadratic and square-root models (Models 7-9) exhibit the poorest fit.

The various models correcting for non-response bias yield similar estimates for the measure of income inequality. The last two columns in Table 2 report the estimated Gini coefficients for income and expenditure per capita across models. They range from 0.329 to 0.351 for income, and from 0.305 to 0.320 for expenditure. Considering the differences in specifications used and fit achieved, these ranges are quite narrow, particularly for expenditure. Across models, 95% confidence intervals of the income Gini coefficients have lower bounds of 0.324-0.336 and upper bounds of 0.333-0.365. Expenditure Gini coefficients have lower bounds of 0.302-0.313 and upper bounds of 0.309-0.327. With the exception of the Gini coefficients from the poorly performing Models 7-9, all Ginis fit within the 95% confidence intervals of each other. This provides some evidence of consistency of the estimates.

Table 2. Estimation results for various logistic models of response probability

Specification of *g(x)*	$E(\theta_1)$ / s.e.	$E(\theta_2)$ / s.e.	Objective Value: Sum of Squared Weighted Errors	Factor of Proportio -nality (σ^2)	Akaike Informat. Criterion	Schwarz Informat. Criterion	Per-Capita Income Gini / s.e.	Per-Capita Expendit. Gini / s.e.
Household level								
1: $\theta_1+\theta_2$log(income)	14.9909	-1.1853	85,079.65	.0776	8,887.82	8,885.20	.3506	.3151
	(.0169)	(.0016)					(.0072)	(.0024)
2: $\theta_1+\theta_2$log(expenditure)	17.2057	-1.4232	81,219.50	.0753	8,770.53	8,767.92	.3426	.3200
	(.0184)	(.0017)					(.0035)	(.0033)
Per capita								
3: $\theta_1+\theta_2$log(income)	11.6554	-.9939	83,400.47	.0757	8,837.46	8,834.85	.3488	.3151
	(.0122)	(.0013)					(.0062)	(.0023)
4: $\theta_1+\theta_2$log(expenditure)	13.0790	-1.1742	80,554.84	.0737	8,749.77	8,747.16	.3423	.3181
	(.0142)	(.0015)					(.0035)	(.0025)
5: $\theta_1+\theta_2$log(exp.)^2	7.4535	-.0603	81,623.97	.0744	8,783.08	8,780.46	.3421	.3176
	(.0066)	(.0001)					(.0039)	(.0026)
6:	1.5485	-.1391	83,644.60	.0757	8,844.85	8,842.23	.3418	.3168
$\theta_1$log(exp.)+$\theta_2$log(exp.)^2	(.0013)	(.0001)					(.0045)	(.0028)
7: $\theta_1+\theta_2 10^{-3}$expenditure	3.3528	-.0254	95,919.03	.0845	9,190.73	9,188.11	.3338	.3084
	(.0019)	(.0000)					(.0044)	(.0023)
8: $\theta_1+\theta_2 10^{-9}$expenditure^2	3.2832	-.0026	99,480.83	.0873	9,282.83	9,280.21	.3289	.3054
	(.0020)	(.0189)					(.0023)	(.0017)
9: $\theta_1+\theta_2$expenditure^{½}	4.0854	-.0137	88,808.82	.0792	8,996.18	8,993.56	.3388	.3130
	(.0023)	(.0000)					(.0052)	(.0029)
10: $\theta_1+\theta_2$expenditure$^{1/3}$	5.1798	-.1224	85,366.91	.0768	8,896.33	8,893.72	.3408	.3153
	(.0035)	(.0001)					(.0049)	(.0029)

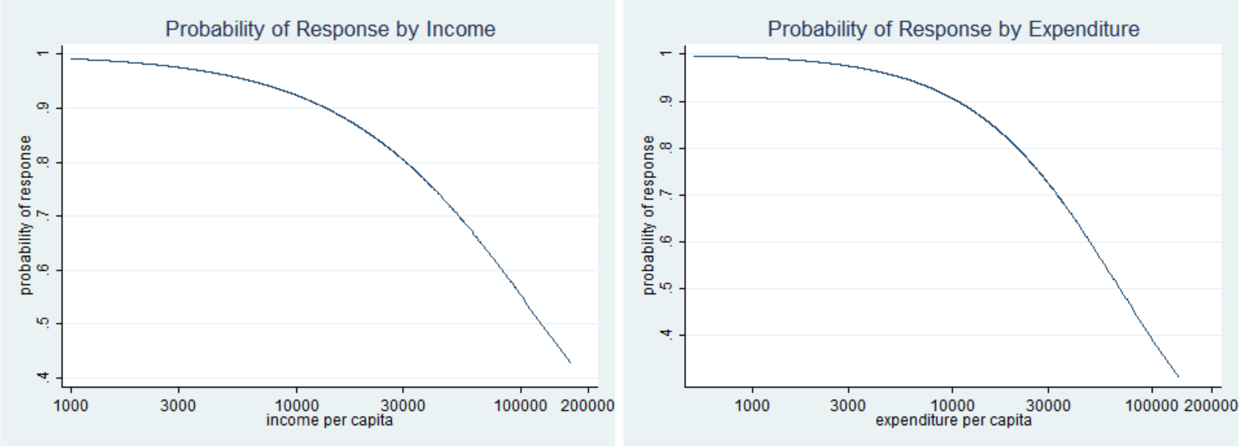
Note: Sample size is 2,526 PSUs, containing 46,857 household observations. PSU populations are fitted using response probabilities estimated for all households. Standard errors on Gini coefficients are bootstrapped estimates.

Beside the ten models in Table 2, we have considered other polynomial specifications as well as a model controlling for the four quarterly rounds in the 2008/2009 HIECS. While some coefficients in these models were significant, the models' overall fit was worse than in Models 1-4, and the corresponding Gini coefficients did not depart significantly from those in Table 2. The imputed household response probabilities and Gini coefficients are thus not too sensitive to the addition of more variables into *g(x)*.

In the rest of the analysis, we will use Model 4 as a benchmark specification due to its superior fit and similarity to the model used by Korinek et al. (2006, 2007). The following figures provide additional results for this model, as well as other comparison models. Figure 3 shows households' probability of survey response by income or expenditure per capita estimated in

Models 3 and 4. In agreement with negative estimates of θ_2 in the logarithmic specifications, the estimated response-probability falls with income, most rapidly in the highest range of incomes (expenditures). Figure 3 thus confirms the central premise that richer households are systematically less likely to participate in surveys, and that this issue is particularly grave for top-income households. The response probabilities shown here will be used as the appropriate household weights for the imputation of income distribution and measures of inequality.

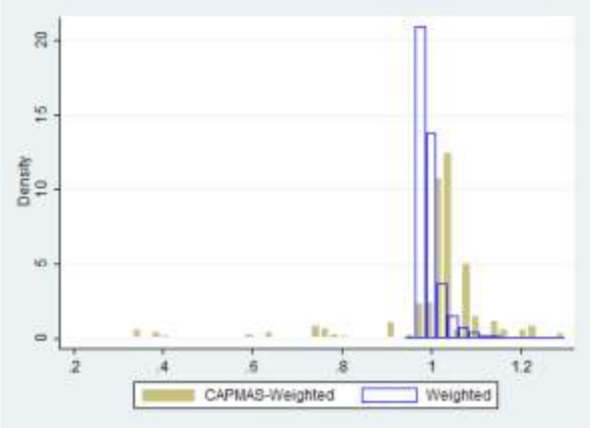
Figure 3. Household response probability by income or expenditure per capita (Models 3,4)



(a) Model 3 (b) Model 4

The corrected weights differ significantly from the CAPMAS-provided sampling weights. The CAPMAS provides sampling weights that correct for unit non-response by simply expanding the weight for the non-response rate at PSUs. CAPMAS-provided sampling weights are normalized to 1, have standard deviation of 0.173, and are identical for all households within a PSU. Weights from Model 4, obtained as the inverse response probabilities estimated in that model, have a mean of 1.041, standard deviation of 0.057, and vary across all households even within PSUs. Figure 4 reports the distribution of households' sampling weights provided by the CAPMAS and those derived from Model 4.

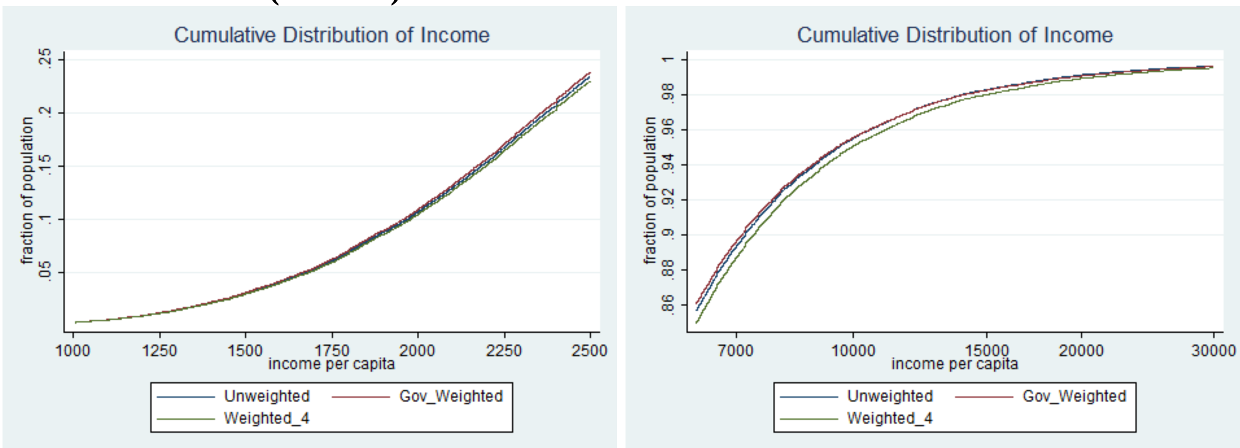
Figure 4. Distribution of CAPMAS-provided sampling weights and weights correcting for non-response bias from Model 4



Use of the corrected weights affects the imputed income distribution. Figures 5-6 show the implications of our estimation for the imputed distribution of per-capita incomes and the corresponding Lorenz curves, for the entire population as well as for the poorest and richest households. These figures show that our correction of the survey-nonresponse bias increases our measurement of income-inequality. The Lorenz curve imputed using our weights first-order dominates both the uncorrected Lorenz curve as well as the CAPMAS sampling-weights corrected Lorenz curve on the entire domain. The uncorrected and CAPMAS-corrected Lorenz curves do not exhibit clear dominance over one another.

Under our corrected income distribution, the estimated fraction of households in the highest income range increases, and the fraction of households in all lower income ranges – including the lowest-income range (less than LE2,500 in Figure 5 panel a) – falls. The share of national income held by top earners is also corrected upward from the unweighted distribution, from 7.18% to 8.15% for the top one percent of earners, and from 27.83% to 29.30% for the top ten percent of earners. (These shares are 7.40% and 28.26%, respectively, in the distribution corrected using CAPMAS sampling weights.)

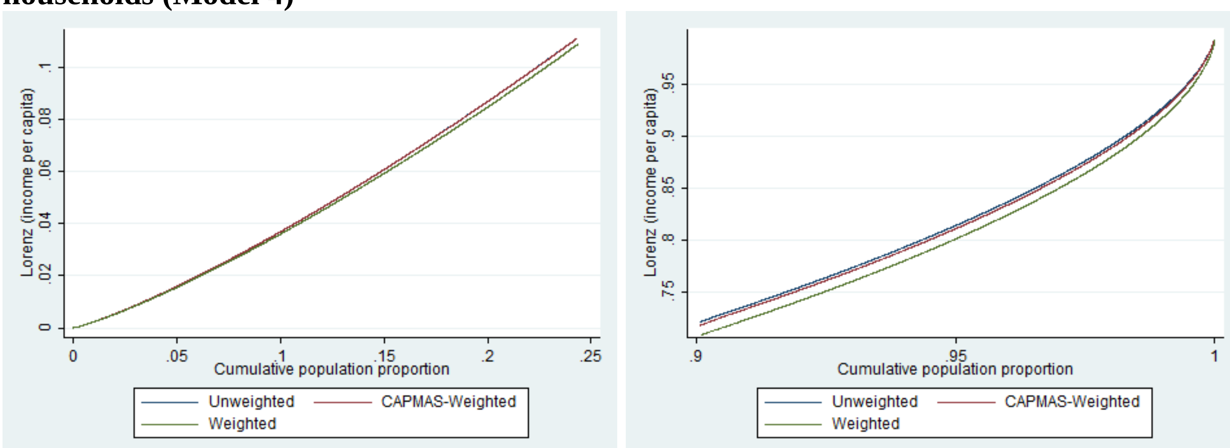
Figure 5. Cumulative distribution of per-capita income among the poorest 25% and richest 10% of households (Model 4)



(a) Poorest 25% per-capita incomes (Model 4)

(b) Richest 10% per-capita incomes (Model 4)

Figure 6. Lorenz curves in the population, and for the poorest 25% and richest 10% of households (Model 4)



(a) Lorenz curve for the poorest 25% (Model 4)

(b) Lorenz curve for the richest 10% (Model 4)

Correspondingly, use of the corrected weights affects the imputed Gini index of inequality positively. By reweighting income distribution to account for households' endogenous survey response, we obtain significantly higher measures of income inequality. The Gini coefficient for per-capita incomes using simple household-size weights is 0.3289 (s.e. 0.0023). The Gini coefficient using the CAPMAS-provided sampling weights is 0.3305 (s.e. 0.0024). The Gini coefficient using response-probability weights estimated in our Model 4 is 0.3423 (s.e. 0.0035). This corrected Gini coefficient is statistically higher than both of the uncorrected ones at the 1% level of significance (p-values of 0.002).

For per-capita expenditure, the Gini coefficient for the unweighted distribution is 0.3054 (s.e. 0.0017), while that using the CAPMAS-provided sampling weights is 0.3070 (s.e. 0.0019). The Gini coefficient using response-probability weights estimated in Model 4 is 0.3181 (s.e. 0.0025). Again, this corrected Gini coefficient is statistically higher than either of the uncorrected ones at the 1% level of significance (p-values of 0.001).

Use of the corrected weights also significantly affects the estimated distribution of top incomes. The Pareto coefficient for unweighted per-capita incomes is 2.428, and the inverted Pareto coefficient is 1.700. For incomes weighted by the CAPMAS-provided weights, these coefficients are 2.392 and 1.718, respectively. For incomes weighted by the response-probability weights estimated in Model 4, these coefficients are 2.250 and 1.800. For per-capita expenditure, the Pareto and inverted Pareto coefficients are 2.685 and 1.593 in the unweighted income distribution, 2.606 and 1.623 in the income distribution weighted using the CAPMAS weights, and 2.478 and 1.677 in the income distribution weighted as per Model 4.

The corrected weights estimated across the alternative models in Table 2 give rise to very different estimates of top-income distribution. This variation can be explained by the differential treatment of top-income households across models. By estimating households' survey-response probability as a function of their log-expenditure (or log-income), versus regular or squared expenditure, we assign very different weights to the highest-income households, while keeping weights of lower-income households similar.

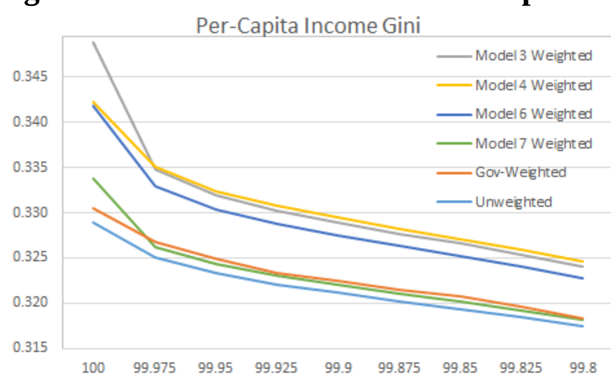
Extreme observations

In this section we test sensitivity of the Gini coefficients to extreme observations on the right-hand side of the distribution, in the raw data as well as in the income distribution corrected for unit non-response. If top incomes turn out to be influential, we then correct for their presence using an estimated Pareto distribution as discussed in the methodological part.

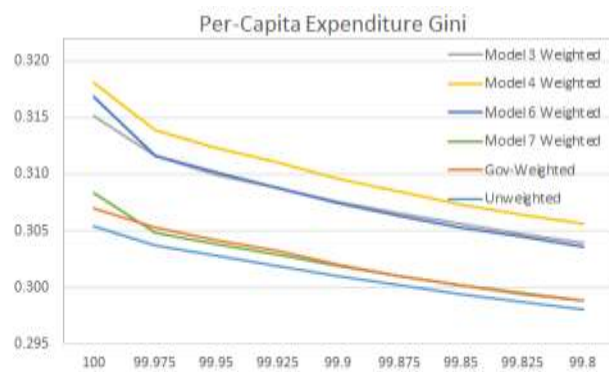
Extreme values in the Egyptian data cannot be ignored. On the one hand, removing some of the top observations may contribute to underestimation of inequality if these observations are accurate and representative of the underlying population. On the other hand, by keeping top observations that arise from data errors or those that do not represent the underlying population well may lead to overestimation of inequality. In either case our inequality estimates may be biased.

Distribution of income and expenditure in the HIECS is indeed very sensitive to top observations, and this sensitivity persists when we correct for unit non-response. A sensitivity analysis reported in Figure 7 shows that inequality measures are very sensitive to the top 0.025% of observations. In this analysis, we recalculate the Gini and Pareto coefficients after removing 0.025%-0.2% households with the highest incomes (12-96 households in the 100% sample of the 2008/2009 HIECS). A significant portion of the difference in Gini coefficients across models disappears as we remove the highest-earning 0.025%-0.05% of households (12-24 households). Exclusion of additional high-income households does not yield significant changes. The difference in statistics across models appears to converge to a particular level, which decreases at a much slower rate with exclusion of additional households. (Figure A2 in the annex reports the same patterns for the Pareto and inverted Pareto coefficients.)

Figure 7. Gini index for income and expenditure in trimmed distributions (Models 3-7)



(a) Gini coefficient for income per capita



(b) Gini coefficient for expenditure per capita

Note: '100' indicates full, untrimmed distribution. '99.975' is distribution with 0.025% of top-value households trimmed (12 households in the 100% 2008/2009 HIECS), …, '99.8' is distribution with 0.2% of top-value households trimmed (96 households).

The discussion above suggests that observations with the highest incomes affect the measurement of inequality. Excluding these observations yields lower and more homogeneous estimates of inequality across models. A question arises whether exclusion is appropriate theoretically, given that it reduces sample size and may result in the censoring of meaningful observations. Here we address these questions by comparing actual observations of top incomes with values imputed under their expected Pareto distribution, and estimating the effects on Gini coefficients. This provides an alternative way to evaluate robustness of our Gini coefficients to the presence of extreme incomes in the sample. In view of our results about survey non-response, this also allows us to comment on the relative significance of the two statistical issues with top-income households – unit non-response and extreme income observations.

Table 3 presents semi-parametric estimates of Gini coefficients, obtained by replacing the highest top ten percent of income observations (alternatively, 5% or 20%) with values imputed from the corresponding Pareto distribution as per Cowell and Flachaire (2007), and Davidson and Flachaire (2007). The first three rows show the benchmark non-parametric estimates – unweighted; corrected for sampling probability using CAPMAS weights; and corrected for non-response bias as per Model 4. These three rows again illustrate the importance of correcting for survey non-response. These rows serve as a benchmark to which the following semi-parametric estimates will be compared. The next three rows present the main results – semi-parametric estimates with the top ten percent of incomes imputed from corresponding Pareto distributions. The three rows differ in the definition of the top ten percent of incomes and in the estimated α, as they assign different weights to each income observation (i.e., unity, CAPMAS weights, and non-response correcting weights). The last six rows report on a robustness check, where such imputation is performed on top five or twenty percent of incomes.

The main finding is that the CAPMAS data do not appear to suffer from extreme income observations relative to what would be predicted if our top-income data followed the Pareto distribution exactly. The corrected Gini coefficients are essentially unchanged, falling or rising by a very small amount. This suggests that the exclusion of top incomes in the previous section is not warranted on the grounds that they are outliers, but simply as a robustness test of the Gini estimates to individual income observations. The size of the correction for extreme observations is trivial compared to the correction for unit non-response. The results for expenditure, available on request, are analogous.

In the income distribution uncorrected for non-response bias, the semi-parametric Gini coefficient – corrected for the possible presence of extreme observations among the top 10% of incomes – is 0.3278 compared to the non-parametric value of 0.3289. When we increase the range of top incomes to be imputed, from 10% to 20% of households, the semi-parametric Gini falls to 0.3273. In the income distribution sampling-corrected using CAPMAS weights, the semi-parametric Gini coefficient is the same as the non-parametric estimate, 0.3305. Finally, in the income distribution corrected for non-response bias using weights from Model 4, the corrected Gini is again the same as the uncorrected value, 0.3423. When we increase the range of top incomes to be imputed, from 10% to 20% of households, the semi-parametric Gini rises slightly to 0.3425.

Table 3. Non-parametric and semi-parametric estimates of Gini coefficients

Modeling of top incomes	Correction for extreme observations	k	Sampling correction	Pareto coefficient *a* (s.e.)	$Gini_{n-k}$, (s.e.)	$Gini_k$, (s.e.)	Gini (s.e.)
Non-parametric	No	k=10%	No	2.4279 (.0309)	.2191 (.0007)	.2584 (.0069)	.3289 (.0023)
	No	k=10%	Yes, CAPMAS	2.3919 (.0326)	.2175 (.0007)	.2654 (.0070)	.3305 (.0024)
	No	k=10%	Yes, Model 4	2.2501 (.0329)	.2214 (.0007)	.2844 (.0112)	.3423 (.0035)
Semi-parametric	Yes	k=10%	No	2.4279 (.0309)	.2191 (.0007)	.2594 (.0043)	.3278 (.0005)
	Yes	k=10%	Yes, CAPMAS	2.3919 (.0326)	.2175 (.0007)	.2643 (.0047)	.3305 (.0005)
	Yes	k=10%	Yes, Model 4	2.2501 (.0329)	.2214 (.0007)	.2857 (.0056)	.3423 (.0005)
Semi-parametric	Yes	k=5%	No	2.4638 (.0937)	.2463 (.0008)	.2546 (.0128)	.3288 (.0006)
	Yes	k=5%	Yes, CAPMAS	2.4378 (.0969)	.2452 (.0008)	.2580 (.0137)	.3305 (.0006)
	Yes	k=5%	Yes, Model 4	2.2507 (.0961)	.2503 (.0008)	.2856 (.0167)	.3422 (.0006)
Semi-parametric	Yes	k=20%	No	2.4190 (.0223)	.1864 (.0007)	.2606 (.0031)	.3273 (.0004)
	Yes	k=20%	Yes, CAPMAS	2.3811 (.0234)	.1849 (.0007)	.2658 (.0034)	.3306 (.0004)
	Yes	k=20%	Yes, Model 4	2.2603 (.0232)	.1876 (.0007)	.2840 (.0039)	.3425 (.0005)

We can now come back to the question of within-*j*/between-*j* trade-off discussed in the methodological section. We argued that using a highly aggregated *j* would be likely to overshoot the Gini correction and would lead to results that are less consistent with the Pareto corrections proposed. Indeed, our non-response correction – of 1-2 percentage points – is smaller than that reported by Korinek et al. (2006, 2007) for the US Current Population Survey (CPS) – of 4-5 percentage points.

To test the claims regarding appropriate geographic aggregation, we have re-estimated the models in Table 2 using governorates by urban and rural substrata (50 areas) rather than PSUs. Refer to Table A2 in the annex. If we compare the models with the best fit (model 4 in the two tables) we find that using governorates by urban and rural areas raises the corrected Gini for income from 0.3423 (s.e. 0.0035) to 0.3714 (0.0129) and the corrected Gini for expenditure from 0.3181 (0.0025) to 0.3419 (0.0075). Across most models, the estimated Ginis rise by 3-5 percentage points for income, and by 1-4 percentage points for expenditure.

In our view, Table 2 provides more accurate estimates for the HIECS data than Table A2 in the annex. First, Ginis estimated at the governorate by urban/rural areas are consistently higher than the semi-parametric Ginis estimated using the alternative Cowell and Flachaire (2007) and Davidson and Flachaire (2007) methodology proposed while the Ginis estimated with PSUs are very much in line with those estimates. Second, in Table A2, all Ginis show significantly higher standard errors. Third, the Egyptian data exhibits a different distribution of incomes and demographics than the US CPS, with significant heterogeneity across governorates, and relative homogeneity within them. For the imputation of response probabilities, it is more meaningful to compare the frequencies of incomes of households with their counterparts in neighboring PSUs within a governorate than with households from across governorates. Fourth, the HIECS data have a much higher household response rate (96.3%) than the US CPS (91.7%), implying that incomes of responding households are likely to encompass incomes of nonresponding households even within smaller regions, and that there is less bias to correct. And fifth, inequality is much lower in the HIECS data, suggesting that the percentage-point correction in the Gini may be lower. The optimal tradeoff of the within-*j*/between-*j* number of bins depends on the nature of the model and on the nature of the data at hand. This paper has proposed a different approach and applied this approach to a different data set as compared to Korinek et al. (2006 and 2007). Clearly, the question of optimal within-*j*/between-*j* trade-off will require testing in a separate paper to be fully resolved but this paper shows that an alternative path is possible and appears preferable in the case of the HIECS data.

5. Discussion

This paper has evaluated income inequality and the distribution of top incomes in Egypt in the presence of a variety of potential statistical issues, including unit nonresponse and representativeness of top income observations. The joint use of two distinct statistical methods for correcting top incomes biases, and sensitivity analysis of their technical specifications and joint performance were methodological innovations of this study. We first tested and corrected for the problem of unit non-response by top income households. Correction for unit non-response increased the estimate of inequality by 1.3 percentage points (confidence interval 1 to 2 percentage points). The estimated Gini coefficient for income per capita rose from 0.329 to 0.342 in the main model, while the Gini for expenditure per capita rose from 0.305 to 0.318.

Given the importance of representation of top incomes in the sample, we next evaluated how influential are individual income observations at the upper tail of the Egyptian distribution, and whether they present a measurement issue. We found that the Egyptian distribution of top incomes follows rather closely the Pareto distribution. The observed top incomes appear to be representative of the underlying population and need to be considered when measuring inequality. This analysis reinforces the case for assigning a greater weight to the observed top incomes to correct for the systematic non-response of top income households in the population.

There are several policy implications of these results that are relevant for Egypt today. First, the paper has validated the quality of the Egyptian HIECS with respect to top observations, the income and expenditure aggregates and the measurement of income inequality. Also, in the world of household surveys, the Egyptian data stand out as particularly good data. Second, these findings motivate the search for factors that could explain popular perceptions about income inequality in Egypt and elsewhere. As the World Bank (2014) report has shown, there are many factors that could explain perceptions of inequality that are little related with the measurement of inequality itself and that are little researched, including the role of expectations about the future, changes in the reference groups, the expansion and penetration of the social media or the lack of GDP trickle-down effects. The priority for Egypt today may not be the reduction of income inequality but the expansion of the growth base, providing more opportunities to economically marginalized groups such as the youth and women, providing more voice to the media-excluded groups such as the poor and rural residents and others. Inequality of opportunities, inequality of rights, inequality of aspirations and inequality of values are some of the inequality dimensions that are easily confounded with income inequality but that should be carefully distinguished by the policy maker.

References

Alvaredo, F. and Piketty, T. (2014) Measuring Top Incomes and Inequality in the Middle East: Data Limitations and Illustration with the Case of Egypt. Economic Research Forum, Working Paper 832, May 2014.

Atkinson, A.B. and Micklewright, J. (1983) On the reliability of income data in the family expenditure survey 1970–1977, Journal of the Royal Statistical Society Series A 146(1), 33-61.

Atkinson, A, Piketty, T. and Saez, E. (2011) Top incomes in the long run of history, Journal of Economic Literature, 49, 3-71.

Cowell, F.A. and Victoria-Feser, M.-P. (1996) Poverty measurement with contaminated data: A robust approach, European Economic Review, 40, 1761-1771.

Cowell, F.A. and Victoria-Feser, M.-P. (1996) Robustness properties of inequality measures, Econometrica, 64, 77- 101.

Cowell, F.A. and Victoria-Feser, M.-P. (2007) Robust Lorenz curves: a semiparametric approach, Journal of Economic Inequality, 5, 21–35.

Cowell, F.A. and Flachaire, E. (2007) Income distribution and inequality measurement: The problem of extreme values, Journal of Econometrics, 141(2), 1044–1072.

Davidson, R. and Flachaire, E. (2007) Asymptotic and bootstrap inference for inequality and poverty measures, Journal of Econometrics, 141(1), 141-166.

Deaton, A. (2005) Measuring Poverty in a growing world (or measuring growth in a poor world), The Review of Economics and Statistics, 87(1), 1-19.

Jolliffe, D., Datt, G., and Sharma, M. (2004) Robust poverty and inequality measurement in Egypt: Correcting for spatial-price variation and sample design effects, Review of Development Economics, 8(4), 557-572.

Juster, F.T., and Kuester, K.A. (1991) Differences in the measurement of wealth, wealth inequality and wealth composition obtained from alternative US wealth surveys, Review of Income and Wealth, 37(1), 33-62.

Korinek, A., Mistiaen, J.A. and Ravallion, M. (2006) Survey nonresponse and the distribution of income, Journal of Economic Inequality, 4, 33-55.

Korinek, A., Mistiaen, J.A. and Ravallion, M. (2007) An econometric method of correcting for unit nonresponse bias in surveys, Journal of Econometrics, 136, 213-235.

Mistiaen, J.A., and Ravallion, M. (2003) Survey compliance and the distribution of income, Policy Research Working Paper #2956, The World Bank.

Neri, L., Gagliardi, F., Ciampalini, G., Verma, V. and Betti, G. (2009) Outliers at upper end of income distribution (EU-SILC 2007), DMQ Working Paper n. 86, November 2009.

Pareto, V. (1896) La courbe de la repartition de la richesse, Ecrits sur la courbe de la repartition de la richesse, (writings by Pareto collected by G. Busino, Librairie Droz, 1965), 1-15.

Piketty, T. (2014) Capital in the Twenty-First Century, Cambridge, MA: Belknap Press.

Stiglitz, J. (2012) The Price of Inequality: How Today's Divided Society Endangers Our Future, New York: W. W. Norton.

World Bank (2014) Inside Inequality in the Arab Republic of Egypt: Facts and Perceptions across People, Time and Space, World Bank Studies, Washington DC.

Annex

Table A1. Countries and years used to generate Figure 1

Income (3 obs./country, 2007-2009)	Expenditure (1 obs./country)		
arg	afg (2007)	idn (2008)	per (2007)
bra	alb (2008)	ind (2009)	per (2008)
col	arm (2009)	jam (2007)	per (2009)
cri	aze (2008)	khm (2008)	per (2010)
dom	bfa (2009)	lao (2007)	png (2009)
ecu	bgd (2010	lbr (2007)	rus (2008)
hnd	bol (2007)	mdv (2009)	slv (2010)
par	bra (2008)	mex (2008)	srb (2007)
per	btn (2007)	mng (2007)	swz (2010
ury	civ (2008)	moz (2008)	tha (2009)
	cmr (2007)	mwi (2010)	tjk (2009)
	col (2008)	ner (2007)	tza (2007)
	cpv (2007)	nga (2009)	uga (2009)
	egy (2009)	npl (2010)	vnm (2008)
	gin (2007)	pak (2007)	zwe (2007)

Table A2. Estimation Results for Various Logistic Models of Response Probability (by governorate and urban/rural areas)

Specification of *g(X)*	$E(\theta_1)$ / s.e.	$E(\theta_2)$ / s.e.	Objective Value: Sum of Squared Weighted Errors	Factor of Proportio -nality (σ^2)	Akaike Informat. Criterion	Schwarz Informat. Criterion	Per-Capita Income Gini / s.e.	Per-Capita Expendit. Gini / s.e.
Household level								
1: $\theta_1+\theta_2$log(income)	20.8870	-1.7686	780,896	.8543	486.81	484.19	.4411	.3398
	(.0088)	(.0008)					(.0389)	(.0070)
2: $\theta_1+\theta_2$log(expenditure)	25.5496	-2.2284	299,122	.3321	438.83	436.21	.3798	.3625
	(.0073)	(.0007)					(.0151)	(.0181)
Per capita								
3: $\theta_1+\theta_2$log(income)	15.8384	-1.4714	577,654	.6505	471.74	469.12	.4210	.3375
	(.0063)	(.0007)					(.0301)	(.0086)
4: $\theta_1+\theta_2$log(expenditure)	18.6483	-1.7947	299,994	.3321	438.97	436.36	.3714	.3419
	(.0062)	(.0006)					(.0129)	(.0075)
5: $\theta_1+\theta_2\log(\text{exp.})^2$	9.9506	-.0916	344,805	.3828	445.94	443.32	.3784	.3452
	(.0028)	(.0000)					(.0188)	(.0101)
6: $\theta_1\log(\text{exp})+\theta_2\log(\text{exp})^2$	2.0269	-.1934	450,540	.5036	459.31	456.70	.3862	.3481
	(.0005)	(.0001)					(.0266)	(.0134)
7: $\theta_1+\theta_2 10^{-3}$expenditure	3.1297	-.0344	2,189,226	2.2715	538.35	535.74	.3594	.3202
	(.0007)	(.0000)					(.0256)	(.0104)
8: $\theta_1+\theta_2 10^{-9}\text{expenditure}^2$	2.9787	-.1329	2,599,937	2.6735	546.95	544.34	.3375	.3089
	(.0008)	(.0005)					(.0089)	(.0037)
9: $\theta_1+\theta_2\text{expenditure}^{½}$	4.3705	-.0195	1,107,645	1.2019	504.29	501.67	.3859	.3399
	(.0009)	(.0000)					(.0373)	(.0165)
10: $\theta_1+\theta_2\text{expenditure}^{1/3}$	6.1436	-.1785	667,983	.7437	479.00	476.39	.3889	.3459
	(.0014)	(.0001)					(.0335)	(.0156)

Note: Sample size is 50 governorate-urban/rural strata containing 46,857 household observations. Standard errors on Gini coefficients are bootstrapped estimates.

Figure A1. Gini and inverted Pareto coefficients for Egypt and the rest of the world

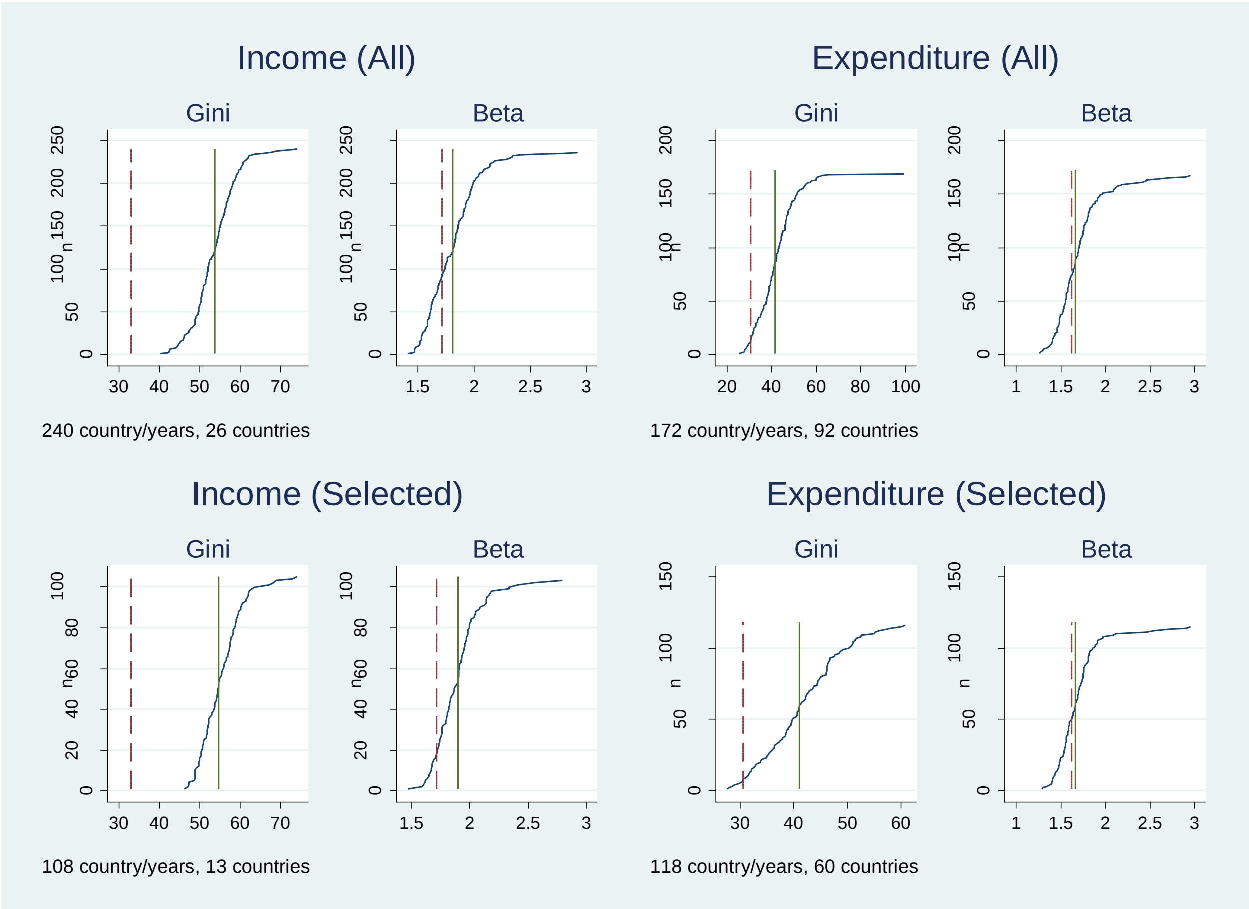
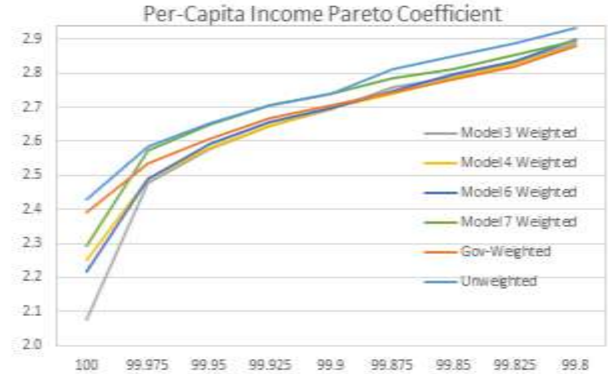


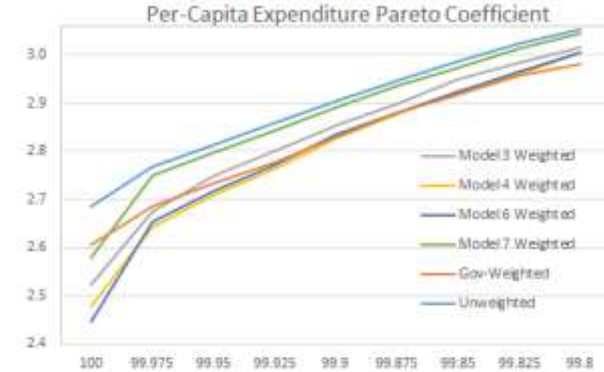
Table A3. Countries and years used to generate Figure A1

Income (All, 240 obs.)					Income (Selected, 105 obs)			Expenditure (All, 172 obs.)				Expenditure (Selected, 118 obs.)		
arg_1992	chl_1987	dom_2006	mex_1992	slv_1996	blz_1993	guy_1993	slv_2004	afg_2005	gnb_1993	moz_2008	tha_2009	alb_2007	jor_2002	tcd_2003
arg_1993	chl_1990	dom_2007	mex_1994	slv_1998	blz_1994	hnd_1991	slv_2005	afg_2007	gtm_2000	mrt_2000	tjk_2003	alb_2008	ken_2005	tgo_2006
arg_1994	chl_1992	dom_2008	mex_1996	slv_1999	blz_1997	hnd_1992	slv_2006	alb_2007	hnd_2004	mrt_2004	tjk_2007	arm_2004	kgz_2003	tha_2002
arg_1995	chl_1994	dom_2009	mex_1998	slv_2000	blz_1998	hnd_1993	slv_2007	alb_2008	hrv_2004	mus_2006	tjk_2009	arm_2008	khm_2003	tha_2006
arg_1996	chl_1996	ecu_1994	mex_2000	slv_2001	blz_1999	hnd_1994	slv_2008	arg_1996	idn_2002	mwi_2004	tls_2001	arm_2009	khm_2006	tha_2009
arg_1997	chl_1998	ecu_1995	mex_2002	slv_2002	bol_1992	hnd_1995		arm_2004	idn_2008	mwi_2010	tls_2006	aze_2001	khm_2007	tjk_2003
arg_1998	chl_2000	ecu_1998	mex_2004	slv_2003	bol_1997	hnd_1996		arm_2008	idn_2009	mys_2004	tur_2002	aze_2008	khm_2008	tjk_2007
arg_1999	chl_2003	ecu_1999	mex_2005	slv_2004	bol_1999	hnd_1997		arm_2009	idn_2010	ndj_i199	tza_2000	ben_2003	lao_2002	tjk_2009
arg_2000	chl_2006	ecu_2000	mex_2006	slv_2005	bol_2000	hnd_1998		aze_2001	ind_1993	ner_2005	tza_2007	bfa_2003	lao_2007	tls_2001
arg_2001	chl_2009	ecu_2003	mex_2008	slv_2006	bol_2001	hnd_1999		aze_2008	ind_2000	ner_2007	uga_1999	bfa_2009	lka_2002	tls_2006
arg_2002	col_1992	ecu_2004	nic_1993	slv_2007	bol_2002	hnd_2001		bdi_1998	ind_2004	nga_2003	uga_2002	bgd_2000	lka_2006	tza_2000
arg_2004	col_1996	ecu_2005	nic_1998	slv_2008	bol_2003	hnd_2003		bdi_2006	ind_2009	nga_2009	uga_2005	bgd_2005	lso_2002	tza_2007
arg_2005	col_1999	ecu_2006	nic_2001	sur_1999	bol_2005	hnd_2004		ben_2003	irq_2006	nic_1998	uga_2009	bgd_2010	mar_1998	ukr_2003
arg_2006	col_2000	ecu_2007	nic_2005	ury_1989	bol_2007	hnd_2005		bfa_2003	jam_2002	nic_2001	ukr_2003	bih_2004	mar_2000	vnm_2004
arg_2007	col_2001	ecu_2008	pan_1991	ury_1992	col_1992	hnd_2006		bfa_2009	jam_2007	nic_2005	vnm_2004	bol_2002	mkd_2002	vnm_2006
arg_2008	col_2003	ecu_2009	pan_1995	ury_1995	col_1996	hnd_2007		bgd_2000	jor_2002	npl_2003	vnm_2006	bol_2007	mkd_2003	vnm_2008
arg_2009	col_2004	gba_1980	pan_1997	ury_1996	col_1999	hnd_2008		bgd_2005	kaz_2003	npl_2010	vnm_2008	btn_2003	mli_2006	yem_2005
arg_2010	col_2006	gba_1986	pan_1998	ury_1997	col_2000	hnd_2009		bgd_2010	ken_2005	pak_2001	yem_2005	btn_2007	mng_2002	zmb_2006
bhs_01_S	col_2007	gba_1988	pan_2000	ury_1998	col_2001	hti_2001		bgr_2003	kgz_2003	pak_2004	zaf_2000	civ_2002	mng_2005	
bhs_2001	col_2008	gba_1991	pan_2001	ury_2000	col_2003	jam_1990		bih_2004	khm_2003	pak_2007	zaf_2005	civ_2008	mng_2007	
blz_1993	col_2009	gtm_2006	pan_2002	ury_2001	col_2004	jam_1996		bol_2002	khm_2006	pak_2010	zmb_2006	cmr_1996	mrt_2000	
blz_1994	cri_1989	gua_2000	pan_2003	ury_2002	col_2006	jam_1999		bol_2007	khm_2007	per_2003	zwe_2007	cmr_2001	mrt_2004	
blz_1997	cri_1990	gua_2002	pan_2004	ury_2003	col_2007	jam_2001		bra_2002	khm_2008	per_2004		cmr_2007	nga_2003	
blz_1998	cri_1991	gua_2003	pan_2005	ury_2004	col_2008	jam_2002		bra_2008	lao_2002	per_2005		cog_2005	nga_2009	
blz_1999	cri_1992	gua_2004	pan_2006	ury_2005	col_2009	nic_1993		btn_2003	lao_2007	per_2006		col_2003	nic_1998	
bol_1992	cri_1993	guy_1993	pan_2009	ury_2006	dom_1996	nic_1998		btn_2007	lbr_2007	per_2007		col_2008	nic_2001	
bol_1997	cri_1994	hnd_1991	par_1999	ury_2007	dom_1997	nic_2001		civ_2002	lka_2002	per_2008		col_2010	nic_2005	
bol_1999	cri_1995	hnd_1992	par_2002	ury_2008	dom_2000	nic_2005		civ_2008	lka_2006	per_2009		com_2004	pak_2001	
bol_2000	cri_1996	hnd_1993	par_2003	ury_2009	dom_2001	per_1997		cmr_1996	lso_2002	per_2010		cpv_2001	pak_2004	
bol_2001	cri_1997	hnd_1994	par_2004	ven_1989	dom_2002	per_1998		cmr_2001	ltu_2003	phl_2003		cpv_2007	pak_2007	
bol_2002	cri_1998	hnd_1995	par_2005	ven_1992	dom_2003	per_1999		cmr_2007	mar_1998	phl_2006		dji_1996	pak_2010	
bol_2003	cri_1999	hnd_1996	par_2006	ven_1995	dom_2004	per_2000		cod_2004	mar_2000	png_2009		egy_2009	per_2003	
bol_2005	cri_2000	hnd_1997	par_2007	ven_1998	dom_2005	per_2001		cog_2005	mdg_2001	pry_2000		fji_2002	per_2004	
bol_2007	cri_2001	hnd_1998	par_2008	ven_2000	dom_2006	per_2002		col_2003	mdg_2005	rus_2003		gha_1998	per_2005	
bra_1992	cri_2002	hnd_1999	par_2009	ven_2001	dom_2007	per_2003		col_2008	mdv_2003	rus_2008		gha_2006	per_2006	
bra_1993	cri_2003	hnd_2001	per_1997	ven_2002	dom_2008	per_2004		col_2010	mdv_2004	rwa_2000		gmb_2003	per_2007	
bra_1995	cri_2004	hnd_2003	per_1998	ven_2003	dom_2009	per_2005		com_2004	mdv_2009	rwa_2005		gnb_1993	per_2008	
bra_1996	cri_2005	hnd_2004	per_1999	ven_2004	ecu_1994	per_2006		cpv_2001	mex_2000	sen_2001		gtm_2000	per_2009	
bra_1997	cri_2006	hnd_2005	per_2000	ven_2005	ecu_1995	per_2007		cpv_2007	mex_2004	sen_2005		hnd_2004	per_2010	
bra_1998	cri_2007	hnd_2006	per_2001	ven_2006	ecu_1998	per_2008		dji_1996	mex_2005	sle_2003		idn_2002	phl_2003	
bra_1999	cri_2008	hnd_2007	per_2002		ecu_1999	per_2009		egy_2009	mex_2008	slv_2004		idn_2008	phl_2006	
bra_2001	cri_2009	hnd_2008	per_2003		ecu_2000	slv_1991		eth_2000	mex_2010	slv_2010		idn_2009	png_2009	
bra_2002	dom_1996	hnd_2009	per_2004		ecu_2003	slv_1995		eth_2004	mkd_2002	srb_2007		idn_2010	pry_2000	
bra_2003	dom_1997	hti_2001	per_2005		ecu_2004	slv_1996		fji_2002	mkd_2003	stp_2000		ind_1993	sen_2001	
bra_2004	dom_2000	jam_1990	per_2006		ecu_2005	slv_1998		gab_2005	mli_2006	swz_2000		ind_2000	sen_2005	
bra_2005	dom_2001	jam_1996	per_2007		ecu_2006	slv_1999		gha_1998	mne_2004	swz_2010		ind_2004	slv_2004	
bra_2006	dom_2002	jam_1999	per_2008		ecu_2007	slv_2000		gha_2006	mng_2002	tcd_2003		ind_2009	slv_2010	
bra_2007	dom_2003	jam_2001	per_2009		ecu_2008	slv_2001		gin_2002	mng_2005	tgo_2006		irq_2006	stp_2000	
bra_2008	dom_2004	jam_2002	slv_1991		ecu_2009	slv_2002		gin_2007	mng_2007	tha_2002		jam_2002	swz_2000	
bra_2009	dom_2005	mex_1989	slv_1995		gtm_2006	slv_2003		gmb_2003	moz_2002	tha_2006		jam_2007	swz_2010	

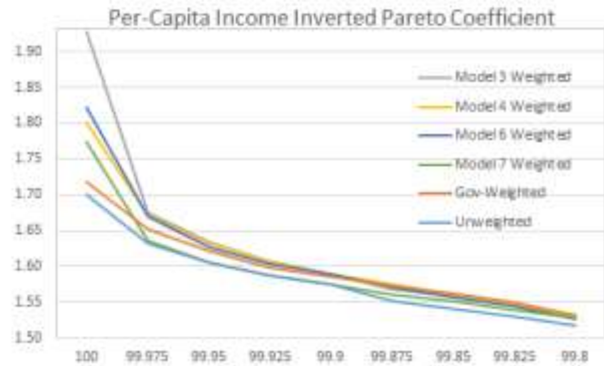
Figure A2. Pareto and inverted Pareto coefficients for income and expenditure per capita in trimmed distributions (Models 3-7)



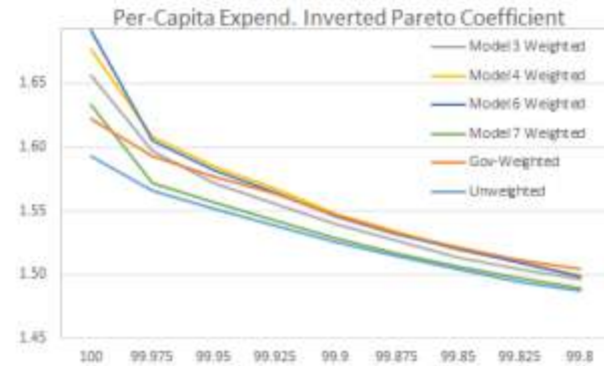
(a) Pareto coefficient for income per capita



(b) Pareto coefficient for expend. per capita



(c) Inverted Pareto coefficient for income per capita



(d) Inverted Pareto coef. for expend. per capita

Note: ‘100’ indicates full, untrimmed income distribution. ’99.975’ indicates income distribution with the 0.025% households with the highest incomes trimmed (12 households in the 100% sample of the 2008/2009 HIECS). Similarly, ’99.8’ indicates the trimming of 0.2% of highest-earning households (96).