Working Paper Series

Redistributive impacts of fiscal policies in Mexico: Corrections for top income measurement problems

Vladimir Hlasny

ECINEQ WP 2019 - 497

ECINEQ 2019 - 497
May 2019

www.ecineq.org

Redistributive impacts of fiscal policies in Mexico: Corrections for top income measurement problems*

Vladimir Hlasny
Ewha Womans University Seoul, Korea

Abstract

We assess the redistributive impacts of fiscal instruments in a 2014 Mexican household budget survey (ENIGH) correcting for potential top-income measurement problems. We use two correction methods – survey-sample reweighting for households' nonresponse probability, and replacing of top incomes using synthetic values from the Pareto distribution – to re-estimate the impacts of pensions, transfers, taxes, and subsidies. These corrections yield higher Ginis (0–9 pc.pt. increase) and top 1% and 10% income shares (0–5, and 1–5 pc.pt. increases), consistently between the reweighting and replacing methods, and consistently across all income concepts. Moving from pre-fiscal to post-fiscal income, corrections for nonresponse abate, while corrections for mismeasurement rise. Taxable income exhibits the highest inequality, further undergoing the highest upward correction for top income problems, potentially consistent with evidence of earnings misreporting among the rich. Nontaxable income has a strong equalizing impact of 3.3–4.5 points of the Gini further growing under the top-income corrections. The corrections confirm the inequality-neutral impact of contributory pensions, and equalizing impacts of transfers, taxes and subsidies. In-kind transfers, cash-like transfers and direct taxes have the strongest equalizing impacts: 4.7–5.7, 1.6–1.9, and 1.2–2.2 points of the Gini, respectively. Indirect taxes and subsidies are weakly equalizing, by 0.4-0.6 points. Finally, top-income measurement challenges retain their magnitude across the 2010, 2012 and 2014 ENIGH, but household nonresponse becomes more positively selected over time, causing serious biases.

Keywords: fiscal incidence, redistribution, inequality, top income measurement problems, Pareto distribution, Mexico, ENIGH.

JEL Classification: D31, D63, H22, I38, N36.

*I wish to thank Nora Lustig for motivating this work and offering technical feedback. Funding from the Commitment to Equity Institute is acknowledged. An earlier version of the paper appeared as the Commitment to Equity Institute working paper 84, March 2019.

I. Introduction

Mexico has a notoriously unequal society in various dimensions and according to a variety of indicators. At the same time, fiscal instruments of the Mexican state correct the inequality at most weakly, on account of the low redistributive effort (low shares of tax revenue and of social spending to GDP) and ineffective targeting of spending across income groups. The precise redistributive impact is best evaluated using survey data with household-level information on income components. However, these data are known to suffer from various measurement problems, particularly in surveys from developing countries like Mexico, and particularly among observations at the bottom and top of the income scale. Recent methodological literature has proposed advanced methods to correcting for some types of measurement problems, using information from within or outside of the surveys. These methods have led to nontrivial corrections to the distribution of incomes, and have typically increased the estimated inequality.

This study contributes by applying the correction methods to re-assess the redistributive impacts of fiscal policy in Mexico. We rely primarily on the 2014 round of the Mexican household budget survey (*Encuesta Nacional de Ingresos y Gastos de los Hogares*, ENIGH), but we cross-validate our results and comment on trends using also the 2010 and 2012 rounds. We implement two types of correction methods, namely reweighting of the survey sample to correct for households' estimated nonresponse probability, and replacing of potentially contaminated top incomes using synthetic estimates from smooth statistical distributions. Both of these methods rely on information within survey: survey response rates at the level of Mexican states, or the actual versus theoretically expected dispersion of top incomes, respectively.

The first approach (Mistiaen and Ravallion 2003; Korinek et al. 2006 and 2007) aims to inflate the weights of surveyed households to make them encompass even the mass of similar same-region households that have not responded to the survey, according to an estimated behavioral response-probability function. The second approach (Cowell and Victoria-Feser 1996a,b; Cowell and Flachaire 2007; Davidson and Flachaire 2007) mitigates the influence of individual income observations to inequality measurement by replacing them with synthetic values derived from estimated parametric distributions. This method aims to avoid the problems of data contamination by misreporting, and distortion of the income distribution by unit or item nonresponse. The method draws on a long-established literature confirming that the top tail of income distributions worldwide is well approximated by a general family of statistical distributions. In this study, replacing of top incomes is implemented either on all core income concepts, or on net market income alone, from which it is passed onto other income concepts through fiscal rules encompassed by the Commitment to Equity (CEQ) methodology.

After applying the corrections to eight core income concepts in Mexico – market, market+pensions, gross, taxable, net market, disposable, consumable and final incomes per capita – we estimate the corrected income distributions and inequality measures. With the corrected income distributions, we estimate the redistributive impacts of six different fiscal instruments on Mexican households – contributory pensions, cash-like transfers, direct taxes, indirect taxes and subsidies, and in-kind transfers. As a byproduct, we can comment on the relative redistributive role of taxable and nontaxable income in order to shed light on the redistributive effort of the Mexican state, and on the informality, tax evasion and access to fiscal resources in the Mexican economy. Finally, we cross-validate our results using the 2010 and 2012 rounds of the ENIGH, and comment on the persistence and trends in top-income measurement problems, and implications for the measurement of redistributive impacts.

The rest of our study is organized as follows. Section II reviews the existing evidence of income inequality in Mexico, inequality measurement under income reporting problems; and the redistributive impacts of instruments of the Mexican fiscal policy. Section III introduces two recently proposed techniques correcting for income mismeasurement, and Section IV briefly introduces the Mexican household budget survey data. Section V presents our main findings, and section VI discusses their policy implications.

II. Literature Review

Mexico, like Latin America at large, faces notoriously high economic inequality, and a growing body of literature has attempted to measure its degree, nature and sources. Distinct branches of this literature have tackled the sources of inequality in labor and non-labor market earnings, the redistributive impacts of various fiscal instruments, and the measurement challenges faced by tax authorities and survey administrators. This section briefly reviews these branches of literature, their broad findings, and our contributions to the current state of knowledge.

The degree and pattern of income inequality

Income inequality in Mexico rose sharply between the 1980s and the 1990s amid the austerity triggered by the 1980s debt crisis and structural reforms (Lustig 1998). Inequality then fell during the 2000s. During 2003–2008 the Mexican economy grew rapidly in a way that benefited the poor as well as the middle class. The returns to higher education as seen through labor earnings fell. During the 2008–2010 recession, demand for low-skill workers fell, and their purchasing power was held down by stagnating real minimum wage, less progressive public transfers and rising indirect taxes (Campos-Vázquez et al. 2018).

The Mexican public sector suffered from a dramatic fall and slow recovery of oil prices, and weak fiscal position including high public debt, declining tax revenues as share of GDP, and growing deficits (Martorano 2014). These factors limited the space for monetary and fiscal responses to the crisis, even with the support by the International Monetary Fund and the World Bank (Ros 2011). Mexican government implemented various programs, including public works and retraining programs, but these initiatives had insufficient coverage and funding, on account of restricted oil revenues (Freije et al. 2014). Recovery from the crisis was thus slow and uneven. Unemployment and poverty rate jumped even though overall inequality fell slightly between 2006 and 2012. Since then, inequality has remained stagnant or rising (Lopez-Calva and Lustig 2010; Esquivel 2011; Campos-Vázquez et al. 2012; Cornia 2014).

Despite the progress during the 2000s, inequality indicators including the Gini coefficient continue to make Mexico one of the more unequal countries. Thickness of the right tail of the income distribution, and income shares of the top 0.5–20 percent of households further attest to the extreme dispersion of incomes in Mexico (Alvaredo et al. 2017). They place Mexico at the high end globally, nearly on par with Brazil, Colombia, Guatemala, Peru, South Africa, Uruguay, and the US according to some measures (Hlasny 2016). The inverted Pareto coefficient among the richest percentile of households, indicating the degree of top-income dispersion, has been estimated at 2.6–2.7, and as high as 3.9 among the top 0.01 percent of households (Olascoaga 2015).

Our study confirms these findings and offers additional evidence of inequality and top income dispersion during the years 2010, 2012 and 2014. We estimate inequality for multiple core income concepts, and correct them for two specific top income measurement problems – unit nonresponse, and income misreporting among top-income households – using advanced but tractable methods.

Income measurement problems in tax records and budget surveys

One challenge with assessing fiscal redistribution, or measuring inequality per se in Mexico is that income data are widely suspected to suffer from statistical problems including misreporting. A number of studies have acknowledged this and some have attempted to address it in connection to measuring inequality. Among early investigations, Navarrete (1973) distinguished three strata in the household survey according to the relationship between households' incomes and expenditures, and national accounts data. She applied an adjustment to reconcile incomes and expenditures in the strata where incomes under-shot expenditures (ostensibly due to the omission of income in kind), and scaled up higher incomes proportionally to match national accounts data. Félix (1982) substituted consumption for income of the bottom 40 percent of households, and scaled the upper 60 percent of incomes up to account for the survey's aggregate shortfall vis-à-vis the System of National Accounts (SNA). Jimenez et al. (2010) proposed a proportional adjustment factor by which to scale up all incomes in household surveys in order to correct for the problem of tax evasion and to replicate the aggregate income figures in the SNA.

Olascoaga (2015) estimated top income shares in Mexico between 2009 and 2012 using the 2010 ENIGH as well as microdata with personal tax returns and employer-wage returns. He found that high incomes in Mexico are not as unequally distributed as in other Latin American countries, but the topmost incomes exhibit greater dispersion. The estimated inverted Pareto coefficients among Mexican topmost incomes, from 2.6 to as high as 3.9, are large by international standards.

Bustos (2015) used the 2012 ENIGH in combination with the Mexican SNA to correct for biases due to income under-reporting and under-representation of high-income households. To that aim, he fitted the generalized beta (type II), gamma, generalized gamma, and log-normal distributions. He concluded that the generalized beta distribution fit the data well, and outperformed the alternative parametric distributions, even though he cautioned that the statistical significance of these results is unclear. He found that the official poverty estimate was robust to the choice of the correction method, but the Gini coefficient was highly sensitive to the parametric choices, rising from 0.44 to 0.52 under the gamma distribution, or to as high as 0.80 under the GB2.

Bustos and Leyva (2017) adjusted 2012 ENIGH survey incomes to make them compatible with the SNA and the Mexican Tax Administration Service records. They fitted generalized gamma and generalized beta (type II) distributions, and found that these alternative specifications did not affect the measurement of inequality significantly. As a result of the correction, the Gini rose from 0.45 to 0.63, but the measure of poverty fell. Reyes et al. (2017) proposed a method for adjusting incomes to deal with income truncation in the top tail, and underreporting of various income components in the rest of the distribution. This adjustment increased the Mexican Gini from 0.52 to 0.74 or even as high as 0.97, making Mexico the most unequal country globally.

Campos-Vázquez et al. (2018) used the 2012 round of the household survey jointly with the national accounts to re-estimate inequality, accounting for the gap in aggregate income between the two sources. They estimated the residual income excluded from the household survey, and assigned a share of it to the top decile and the rest to the ninth decile. Using these new income shares of the two top deciles, they calculated the shape coefficient of the corresponding top-income Pareto distribution, and imputed the income shares of other fractiles of the top income distribution.

Alvaredo et al. (2017) corrected the ENIGH incomes for misreporting at the top by using a combination of taxpayer data, employer tax filings, national accounts statistics, and an employer survey (*Encuesta Nacional de Ocupación y Empleo*). They concluded that the top income shares in Mexico are among the highest in the world, with the top 10 percent of wage earners accounting for 45% of total wages, and top 10 percent of total-income earners accounting for a staggering 59–66 percent of aggregate income. Bourguignon (2018) applies several alternative adjustment methods to the entire income distribution to correct it for missing income or missing respondents.

These existing studies give rise to vastly different estimates of the degree of inequality in Mexico. Leyva-Parra (2004) sounded an early warning about the performance of alternative correction methods. He compared the existing approaches for realigning income distributions between household surveys and national accounts, and critically examined the assumptions behind the methods. He cautioned that these assumptions were likely violated, giving rise to biases of potentially large magnitudes.

Our study contributes to this growing literature by performing two alternative tractable corrections to the top tail of incomes in the Mexican household survey, using benchmarks from within the survey. Our aim is not to incorporate heterogeneous external information or produce inequality estimates correcting for all suspected biases. Instead, by using methods based on limited assumptions and a controlled information set, our aim is to produce estimates corrected for the specific top income biases, and linked clearly to specific modeling choices and comparison benchmarks. Alternative specifications and corrections performed on prior survey rounds are presented for reference. Finally, our study goes beyond correcting inequality indexes for various income concepts, and re-assesses the redistributive impacts of the core Mexican fiscal instruments.

Effectiveness of fiscal redistribution

A growing number of studies have evaluated the fiscal redistribution system in Mexico, and the prospect for using it to mitigate inequality. Existing fiscal policies have been found to correct inequality at most weakly, on account of the low redistributive effort (low shares of tax revenue and of social spending to GDP) and ineffective targeting across income groups. Lindert et al. (2006) review the universality and composition of transfers in Mexico, and report that 73% of Mexicans did not receive any social transfers, while some programs had benefits going all the way to the top income quintile.

Scott (2008, 2010, 2014) finds that a large share of transfers have limited redistributive effectiveness, including subsidies for social security pensions, exemptions in direct and indirect taxes, energy subsidies, and access to public higher education. Aranda and Scott (2016) report on the implications of public transfers for the poor and for ethnic minorities. Goñi et al. (2011) confirm that Mexico's fiscal system has a weak redistributive impact, both in terms of transfers and taxes. Mexico's fiscal revenues come largely from non-tax rents (e.g., from the state-owned oil company), and only limited revenues come from a narrow base of taxable units (e.g., tax evasion, and tax concessions and loopholes) (Goñi et al. 2011:12). Indeed, government revenues account for a mere fifth of the Mexican GDP. Indirect taxes in Mexico are near proportional according to the share of market income they take up (Goñi et al. 2011:16). Mexican public revenues are made up of nontax oil-sector revenues and direct taxes, and only a small share comes from indirect and other taxes, and social security contributions (Lustig 2018:481–482).

The weak redistributive impact of transfers means that market income and gross income (i.e., adding taxable transfers) exhibit small differences in inequality even by Latin American standards. Differences in inequality for gross income and disposable income (accounting for direct personal income taxes, and cash-like transfers) are larger in absolute value as well as by Latin American standards, but still much lower than is typical in western Europe (Goñi et al. 2011:7).

Fiscal redistribution does affect measures of inequality in Mexico, on account of high initial inequality in market incomes, but contributory social-security pensions are somewhat unequalizing (Lustig 2016). Disposable and post-fiscal incomes appear to be distributed as widely as market incomes, confirming the poor effectiveness of direct transfers and taxes, and the neutral effect of indirect taxes (Lustig et al. 2019). At the same time, final incomes exhibit lower inequality, due to the effect of in-kind education and health transfers.

If tax receipts and transfers were modestly increased, the distributional impact has been simulated to lower the Gini. However, were the tax structure to become more progressive or transfers more targeted, the Gini would fall only marginally, even compared to regional neighbors (Goñi et al. 2011). These results suggest that tax structure and transfer targeting have limited use as redistributive mechanisms, under the current regime of a low tax base, low income tax productivity, and a mixed package of fiscal transfers, some of which are unequalizing.

One limitation of the studies reviewed in this section is that they rely on survey data that suffer from misreporting problems, especially at the top, or on tax authority data that do not capture all households or income sources, especially at the bottom. Our study contributes by re-estimating the distributional impact of multiple core fiscal instruments using household survey data, while correcting for distinct known types of top income biases. The following section explains the correction methods, and their application to assessing the redistributive impact of fiscal instruments.

III. Methodology

Correcting inequality measures for household nonresponse

Household surveys are known to suffer from substantial and systematic nonresponse, which affects the observed distribution of household characteristics and outcomes. To correct the sampling weights and the income distribution in the Mexican survey for unit nonresponse, we apply a method proposed by Mistiaen and Ravallion (2003), and operationalized by Korinek et al. (2006, 2007) for the United States Current Population Survey, and validated by Hlasny (2016), and Hlasny and Verme (2018a,b,c) for other surveys worldwide. This method estimates households' response probabilities as a function of their observable characteristics, based on comparing the full distributions of those characteristics across regions with different mean response rates. The method then corrects households' sampling weights proportionally to their estimated inverse probability to respond – thus accounting for the density of nonresponding households expected to have similar characteristics as these responding households. With these augmented weights, we can obtain a corrected distribution of household characteristics or outcomes, such as incomes.

At the heart of the method is a behavioral equation linking households' observable characteristics x_i and their survey response probability P_i. Following existing studies we model this probabilistic relationship as a logistic function of household characteristics x_i multiplied by corresponding linear parameters θ:

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

where $g(x_i, \theta)$ is an arbitrary twice-continuously differentiable function of x_i, such as the commonly used $g = \theta_0 + \theta_1 \log(income)$. The best-fitting functional form will be selected from among various univariate and multivariate specifications.

Estimating parameters θ allows us to infer each household's probability to respond to the survey $\hat{P}_i$, and thus also the number of households from which survey respondents are drawn according to sampling design. The predicted number of households in a sampling frame for a region, $\hat{m}_j$, can be derived as the sum of densities – or the inverted response probabilities $\hat{P}_i$ multiplied by households' sampling weights – of all actually responding households in the region (written offhandedly as $i \in j$ below to save on notation).

$$\hat{m}_j = \sum_{i \in j} \hat{P}_{ij}^{-1} s_{ij}^{-1} \ . \quad (2)$$

Comparing these model-estimated populations $\hat{m}_j$ to known regional populations m_j according to sampling design, and weighting their deviations by a weight proportional to regional sum of sampling weights and inversely proportional to regional populations, w_j, provides a measure of model fit. The best-fitting coefficients $\hat{\theta}$ are those attaining the lowest weighted sum of squared population deviations:

$$\hat{\theta} = \arg\min_{\theta} \sum_j \left(\hat{m}_j - m_j\right)^2 / w_j \ . \quad (3)$$

Following Korinek et al.'s (2006, 2007) lead, we consider a number of specifications of the behavioral function of households' response. All models estimate households' survey-response probability as a nonlinear function of their characteristics, among the characteristics available in the household budget survey.

Correcting income distribution for potential misreporting among top incomes

Inequality measurement can be sensitive to the presence of even a few observations with misreported incomes (such as due to data-entry errors, underreporting by select individuals, or accounting rules that prescribe reporting of incomes when these are received rather than earned), or incomes distributed systematically differently than the true underlying values (say, due to top-coding, or omission of a nonmonetary income category). One method to deal with these suspected problems is to replace the observed incomes with values obtained under an expected counterfactual distribution, where the counterfactual distribution can be identified from within the survey itself (including Cowell and Flachaire 2007; Davidson and Flachaire 2007).[2]

Because income misreporting is thought to be a problem particularly among top incomes, we can refer to an established literature on parametric approximations of the tails of empirical distributions. After replacing top income observations with parametric estimates, we can compute a corrected measure of inequality among them, either using known parametric properties of the fitted distribution or quasi-nonparametrically, using income values drawn randomly and repeatedly from the fitted distribution. The inequality measure for these incomes can readily be combined with a non-parametric inequality measure for lower incomes, to arrive at an index of

[2] Another branch of literature has proposed deriving the counterfactual distribution from outside the survey. This branch used parametric smoothing to combine within-survey incomes (micro data or group data) with out-of-survey information such as top income shares or national accounts data (Lakner and Milanovic 2013; Piketty et al. 2016; Anand and Segal 2017; Novokmet et al. 2017). Our study uses only data from within the survey.

overall inequality corrected for possible top incomes biases. Its standard error can be computed by bootstrapping the estimation routine.

One candidate for the parametric form is the Pareto distribution. Over a century ago, Vilfredo Pareto (1896) noted that top incomes tend to be distributed subject to systematic polynomial decay adequately described by a few parameters. This empirical tendency has since been confirmed across many countries and years. The probability density function of the Pareto (type I) distribution is

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

The parameter α can be estimated by maximum likelihood as

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

where $X_{(j)}$ is the j_{th} order statistic in the sample of size n, and k is the count of observations classified as top incomes (Hill 1975). The estimation can be modified to allow for lower and upper truncation in situations where only some range of incomes are considered uncontaminated by measurement problems, and distributed Pareto-like. The estimation can also account for sampling weights.

An inverted equivalent of the Pareto parameter, often used as a measure of dispersion of top incomes, is $\beta = \alpha/(\alpha - 1)$. The values of the inverted Pareto coefficients can be compared to those found worldwide (Atkinson et al. 2011). The Gini coefficient under the estimated Pareto distribution for the k top income observations can be derived parametrically as $Gini_k = (2\alpha - 1)^{-1}$ and its standard error can be obtained by bootstrapping. This parametric Gini coefficient can then be combined with the non-parametric Gini for the *n-k* lower-income observations, $Gini_{n-k}$, using geometric properties of the Lorenz curves as

$$Gini = (1 + Gini_k)\frac{k}{n}s_k - (1 - Gini_{n-k})\left(1 - \frac{k}{n}\right)(1 - s_k) + \left(1 - \frac{2k}{n}\right) \tag{6}$$

where s_k refers to the estimated share of aggregate income represented by the parametrically-obtained k top incomes. s_k is estimated using the approximation of the inverted Pareto coefficient as the ratio of the mean top income $\bar{X}$ to the cutoff point for replacement L ($\hat{\beta} = \bar{X}/L$) (aka, van der Wijk's law; Atkinson et al. 2011). The mass of top incomes is thus estimated as $\hat{\beta} \times L \times k$.

The one-parameter Pareto (type I) functional form has been evaluated positively relative to more complex parametric choices, particularly at the topmost end of the income distribution.[3] An empirical question concerns the appropriate lower (and upper) cutoff points for the estimation of the Pareto distribution, and the cutoff points for the replacement of observed incomes with parametric estimates (Hlasny and Verme 2018a). Jenkins (2017), using data for the United Kingdom, concludes that the preferred lower cutoff point for the estimation of the generalized

[3] This includes the two-parameter generalized Pareto (type II) distribution and the four-parameter generalized beta distribution (Jenkins 2017; Hlasny and Verme 2018a,c). The empirical fit of other candidate distributions – including the lognormal, Fisk, Singh-Maddala and Dagum distributions (themselves limiting cases of the generalized beta – II) has been questioned (Butler and McDonald 1989; Brachmann et al. 1996; Jenkins 2007, 2009; Brzezinski 2013). On the other hand, Jenkins (2007, 2009) and Jenkins et al. (2011) make the case for using the generalized beta distribution (type II) as providing more flexibility and superior fit at the top end. In this study, we assess only the Pareto (type I) distribution, on account of its lower requirements on data size. The generalized Pareto (type II) distribution is compared against the Pareto (type I) in Appendix 1.

Pareto (type II) distribution is between the 95th and the 99th percentile of incomes. For the one-parameter Pareto (type I) distribution, the optimal cutoff may be at least as high or higher.

Estimating the redistributive impact of fiscal policy: top income measurement corrections

Households' income components from various sources and fiscal adjustments reported in the household budget survey must be combined in a consistent fashion to be comparable with income concepts used worldwide, and to be informative of the marginal redistributive impact of various fiscal instruments. Following the standardized methodology of the Commitment to Equity Institute studies (henceforth CEQ),[4] the following income concepts are adopted.

The broadest income concept representing households' *primary distribution* is *market income*, including all factor income sources, own production, imputed rent, and private transfers. Two alternative scenarios are distinguished regarding the character of contributory pensions: pensions are counted as pure transfers (excluded from *market income*), or pure deferred market income (as *market income+pensions*).

To this primary distribution, *secondary redistribution* is applied by the state. Adding direct cash and near-cash transfers (contributory pensions, conditional and unconditional cash transfers, school feeding programs, free food transfers, etc.), we obtain *gross income*. As an alternative measure of total income from tax authorities' perspective, we can disregard nontaxable earnings and report *taxable income*.

Subtracting direct personal income taxes from *gross income* we obtain *disposable income*. To see the impact of income taxation without the impact of public transfers, we can disregard cash transfers and report *net market income* after direct taxes. Using *disposable income*, subtracting indirect taxes (value added, excise, etc.) and adding indirect subsidies (energy, food, and other price subsidies) we get post-fiscal *consumable income*. Finally, adding monetized value of in-kind transfers at average government cost (in Mexico, notably, education and health), and subtracting co-payments and user fees, we obtain *final income* (Lustig 2018:17, 234).

Using this classification, we can identify the redistributive impact of various sets of fiscal interventions. Comparing the distribution of *market income* against *market income+pensions* shows the impact of contributory pensions. Comparing the distribution of *market income+pensions* against *gross income* shows the impact of cash-like transfers. Comparing the distribution of *gross income* against *disposable income* shows the impact of income taxes (including on taxable transfers). Comparing the distribution of *disposable income* against *consumable income* shows the impact of indirect taxes and subsidies. Finally, comparing the distribution of *consumable income* against *final income* shows the net impact of participation in in-kind programs.

Correcting the redistributive impacts for income measurement issues is done as follows. The reweighting approach is applied to each income concept of interest, to correct its distribution for unit nonresponse. In the behavioral equation of households' response probability, market or net market income appears to perform the best as an explanatory variable, since it is easily observable by households and can influence their survey compliance behavior (as observed empirically – refer to table 2). The correction weights estimated in the model are then applied to all income concepts

[4] Led by Nora Lustig since 2008, the Commitment to Equity (CEQ) project is an initiative of the Center for Inter-American Policy and Research (CIPR) and the Department of Economics, Tulane University, the Center for Global Development and the Inter-American Dialogue. The CEQ project is housed in the Commitment to Equity Institute at Tulane. The Bill & Melinda Gates Foundation has provided major funding for the preparation of the CEQ Handbook 2016 and updates. For more details see www.commitmentoequity.org, Lustig et al. (2014) and Lustig (2018).

to correct their respective distributions, and thus to observe the redistributive impacts of the fiscal adjustments linking them.

Under the replacing approach, two alternative modalities are applied. One, replacing is performed on each income concept in turn, and the respective corrected distributions are juxtaposed, to observe the redistributive impacts of the fiscal interventions linking them. This approach corresponds to an assumption that all income concepts decay approximately according to the power law, and that mismeasurements or contaminations can occur at any step in the transition from market income to final income. Indeed, prior literature has validated the Pareto approximation on various income concepts.

As an alternative, the replacing approach is applied to the distribution of net market incomes per capita, and fiscal adjustments are re-calculated from the corrected distribution using the fiscal rules encompassed in the CEQ methodology. One reason for relying on net market income is that it is the starting point for the construction of all CEQ Core Income Concepts in the Mexican ENIGH, because it is this income that is directly lifted from household questionnaires. Two, net market incomes per capita are thought to satisfy the power law intrinsic in the Pareto distributions because they are strongly driven by market forces compared to post-transfer and post-subsidy incomes. Three, mismeasurement problems are thought to come primarily from misreporting of factor incomes, rather than from neglected legal loopholes, errors in eligibility determination, or households' selective participation in fiscal programs. Correcting for mismeasurements or contaminations in net market incomes then allows us to accurately track the redistributive impact of the consecutive fiscal interventions.

To implement the correction, the top tail of the net market income distribution is replaced with random draws from the estimated Pareto distribution, and recombined with the bottom values of observed incomes. This pieced-together corrected distribution is then passed on to the CEQ algorithms to re-compute (in this order) taxable income, disposable income, consumable income, final income, gross income, and market income with/without pensions. The exercise is repeated a number of times to ensure that the results will not be contingent on a particular draw from the Pareto distribution.

Quantifying the redistributive impacts, and the effect of measurement corrections

To quantify the redistributive impact of fiscal policies, two alternative inequality indicators are reported consistently for all analyses: the aggregate-income shares of the top 1, 5 and 10 percent of households, sensitive to how heavy the topmost tail is, and the Gini index, more sensitive to the dispersion of incomes near the middle of income distributions than in their tails. Generalized entropy indexes (0,1,2), and standard deviation, skewness and kurtosis of incomes are reported on the margins but not discussed. Percentage point differences in the Ginis and in the top income shares between pairs of income concepts are used as the central measures of the redistributive impact of the fiscal instruments linking them. This corresponds with the practice in existing fiscal incidence studies. In fact, the difference between pre- and post-fiscal Ginis has the interpretation as the Reynolds-Smolensky index of vertical equity of tax and transfer systems (Reynolds and Smolensky 2013).

The rest of this study uses household income per capita as the welfare aggregate, in agreement with the practices in existing academic and policy literature (Deaton 1997:150). Households are weighted by their post-stratification sampling weights, accounting for their size.

IV. Data

This study relies on the 100-percent sample of the 2014 round of the Mexican National Household Income and Expenditure Survey (Encuesta Nacional de Ingresos y Gastos de los Hogares, ENIGH) administered by the National Institute of Statistics and Geography (Instituto Nacional de Estadística y Geografía, INEGI). Our version of survey microdata was obtained from the CEQ Institute, which standardized survey variables with other national surveys and with the CEQ methodology, and generated all core income concepts. This dataset is publicly available.

ENIGH is a high-quality nationally-representative survey of household wage and non-wage incomes, expenditures and consumption. Table 1 provides descriptive statistics of per capita incomes in the ENIGH sample. Their visual inspection does not reveal the presence of any clear outliers or misreported values. Both the pre-fiscal market income and the post-fiscal disposable income exhibit a heavy rightmost tail relative to a reference lognormal distribution (refer to figures A1-2 and A1-3), but this is not on account of a few outliers. Instead, the entire upper tail is dispersed widely.[5]

Unit nonresponse does appear to be a problem in the ENIGH, because 3.8 percent of households nationwide choose not to complete the survey (in addition to the other 6.0 percent of households listed falsely as residing in unoccupied dwellings, and 0.7 percent listed falsely as residing in what turn out to be invalid dwellings). The nonresponse rate varies by state, and in some states rises to 8–9 percent (Coahuila, Guerrero, Jalisco, and San Luis Potosi). Nonresponding households are likely to be systematically selected from among units with rare characteristics and rare income levels relative to the responding units (Mistiaen and Ravallion, 2003, show theoretically why), and their omission may affect our measurement of the income distribution, inequality and the redistributive impact of fiscal interventions. Figure A1-1 illustrates the positive association between the observed incomes and nonresponse rates across Mexican states. (Appendix 1 and table A1-1 show additional information on the ENIGH, including nonresponse rates.)

While ENIGH sampling weights provide some correction for unit nonresponse, INEGI does not publicize raw components of the weights, and so we cannot evaluate the extent of this correction. Moreover, the existing correction is likely to be inadequate, because the weights are at the level of PSUs, while each PSU contains heterogeneous households with vastly different income and demographic profiles, and different response probabilities. Applying the same weights to all households in a PSU means that each household's density is inflated slightly (by 3.8% on average across all PSUs) to incorporate the density of nonresponding units in the same PSU. However, since nonresponding units are thought to come from the extremes of the income distribution, inflating all households' density uniformly and by such a small margin will not correct the income distribution sufficiently for the systematic omission of the extreme-income households.

Additional survey rounds

Beside the 2014 dataset we have two additional rounds of the ENIGH at our disposal: ENIGH 2012 (descriptive statistics in tables A2-1 through A2-3), available also from the CEQ Institute, and ENIGH 2010 (descriptive statistics in tables A3-1 and A3-2), available from Luxembourg Income Study. These surveys can be used for cross-validation, as well as to comment on the

[5] Faulkner (2014) has recently provided a positive evaluation of survey incomes in the ENIGH, but found that incomes fell short of consumption estimates, and proposed that the income distribution's high skew called for the use of modified inequality indexes such as G2 (Gini×mean/median).

evolution of top income issues and measurement of fiscal redistributive impacts over time.[6] Across rounds of the ENIGH, we find that the household nonresponse problem did not clearly change in size over time. Total non-response rate was 10.1% in 2010, 11.8% in 2012, and 10.6% in 2014, while type A nonresponse was 2.9% in 2010, 2.5% in 2012, and 3.8% in 2014. On the other hand, nonresponse appears to have become more consistently associated with households' incomes across survey rounds (refer to figures A2-1 and A3-1). Correlation of statewide mean disposable income per capita and type A nonresponse rate was only 0.04 in 2010, while it was 0.24 in 2012 and 0.17 in 2014.

Comparing the dispersion of top disposable incomes across survey rounds, we find that 2012 incomes are distributed less smoothly subject to a significant discontinuity near MXN500,000 (or the 99.5th percentile) where the density suddenly drops off. The 2012 distribution covers an outlying individual whose per-capita income exceeds MXN3 million and is more than twice the following value (figures A2-2 and A2-3). On the contrary, top disposable incomes in the 2010 sample have a spike in density above the MXN500,000 level.

The differences in nonresponse rates and their association with regional incomes suggest that the reweighting approach may produce more significant correction weights in recent surveys than in 2010. However, when applied to non-smooth distributions of incomes with outliers, it is an empirical question how the corrections of inequality and of fiscal impacts will compare across surveys. The replacing approach may be sensitive to the bunching of top incomes for the estimation of parameters as well as correction of inequality measures.

V. Results

The observed degree of inequality and redistributive impacts of fiscal programs

According to the 2014 ENIGH, Mexican incomes per capita are distributed widely subject to a high right skew, particularly among the richest few households. Refer to table 1 and figure A1-2 (for ENIGH 2010–2012, see the corresponding tables and figures in Appendices 2 and 3). In the income distribution weighted using post-stratification weights and household size, the top decile of Mexican households account for 42.1 percent of aggregate market income or 36.5 percent of aggregate final income. The top ventile accounts for 30.3 percent of market income or 25.7 percent of final income, and the top percentile accounts for 13.6 percent of market income or 12.1 percent of final income. The Gini coefficient is 52.8 for market income, and falls to 49.4 and to 44.2 as one moves from market income to disposable income and to final income, respectively.[7]

[6] Results for these surveys are reported on the margins and in Appendices 2 and 3. The difference in data sources does present some compatibility challenges. In the 2010 survey, the only income concepts available are disposable household income and disposable household income net of after-tax private and social security transfers (in LIS nomenclature, which differs from CEQ), because all incomes are surveyed net of taxes, social contributions and other deduction such as union fees. Source: LIS Data Center, Original survey information – Mexico 2010, www.lisdatacenter.org/wp-content/uploads/our-lis-documentation-by-mx10-survey.pdf (accessed 24 January 2019).

[7] Beside the benchmark Gini on the fully weighted income distribution, we have assessed the Ginis on an income distribution unweighted using post-stratification weights, and on an entirely unweighted (=equal household weights) distribution. These are available on request.

The results in table 1 can be contrasted with those using prior surveys. In 2012, the top decile of households accounted for 40.9% of market income and 35.8% of final income, the top ventile accounted for 28.4% and 24.3%, and the top percentile accounted for 11.3% and 9.1%, respectively. The Gini was 52.1 for market income, 44.0 for final income. In 2010, for disposable income net of private and social security transfers, the top decile accounted for 38.8%, the top ventile – for 26.7%, and the top percentile – for 10.8%. The Gini was 49.7. For disposable income inclusive of transfers, the top decile accounted for 36.5%, the top ventile – for 24.9%, the top percentile – for 9.8%, Comparing columns in table 1 confirms that contributory pensions have a neutral or slightly unequalizing impact on general inequality measured by the Gini (columns 1 & 2), while cash-like transfers (col. 2 & 3), direct taxes (col. 3 & 6), indirect taxes and subsidies (col. 6 & 7) and in-kind programs (col. 7 & 8) have equalizing impacts. The Gini falls by 1.8 percentage points on account of cash-like transfers, by another 1.6 percentage points on account of progressive income taxation, and by a significant 4.8 percentage points on account of in-kind programs (refer to table 7 for summary). The impacts on top income shares are analogous, and large in magnitude. Cash-like transfers decrease the top 10% income share by 0.9 percentage points (0.3pc.pt. for top 1% income share), direct taxes decrease the top 10% income share by 1.3 percentage points (0.7pc.pt. for top 1% income share), and in-kind transfers decrease the top 10% income share by 3.1 percentage points (1.2pc.pt. for top 1% income share).

To evaluate how sensitive these findings are to top income measurement issues, the following sections correct the income distributions for two distinct types of expected measurement problems – unit nonresponse, and income misreporting – and re-estimate the redistributive impacts of fiscal interventions.

Correcting for unit nonresponse by reweighting

First we attempt to correct the income distribution for the tendency of rich households not to complete surveys. In this analysis we disregard unoccupied or invalid dwellings, and restrict our attention to households that were successfully contacted, because only their probability of survey response is amenable to behavioral analysis. Even among these households, we ignore instances when interview was impossible due to climatological, political, social or security – that is, non-behavioral – problems (13 out of 838 nonrespondents nationwide).

Table 2 reports the results of univariate model specifications using as an explanatory variable households' market, gross or disposable income, either at the household level or per capita. The estimated models are enumerated in the first column. The table shows the estimated values of the model intercept $\widehat{\theta_0}$, slope coefficient on the income measure $\widehat{\theta_1}$, and selected measures of model fit: sum of weighted squared deviations of regional populations, factor of proportionality σ^2 related to the typical squared deviation between predicted and actual regional populations, and the Akaike and Schwarz (Bayesian) information criteria. The Akaike information criterion is used to guide model selection because of its good consistency properties.

To illustrate the implications of each model for inequality, the last three columns in table 2 show the Gini indexes for market income per capita on the survey sample reweighted using the nonresponse-correction weights $\hat{P}_{ij}^{-1}$ (refer to equation 1). These Ginis can be compared to those in table 1, repeated for convenience in the first row in table 2, with the 'weighted data' Gini viewed as the benchmark.[8]

Individual rows in table 4 show alternative specifications of $g(x_i, \theta)$: logarithmic, linear, polynomial or square root. Across all models, we find that income has a consistent significant negative effect on the probability of response ($\widehat{\theta_1} < 0$). The corrected Ginis are always higher than the uncorrected ones (by 0.9–8.2 percentage points, and 4.4 on average in the ENIGH-weighted

and the Gini was 47.1. The degree of dispersion among top disposable incomes, as well as general income inequality, thus appears to have risen in Mexico during both periods 2010–2012 and 2012–2014.

[8] Because survey weights in ENIGH already provide limited correction for nonresponse, it is a standard practice to report Ginis derived under composite weights ($\hat{P}_{ij}^{-1} \times s_j^{-1} \times HHsize$, 'weighted data'); Ginis using only the nonresponse-correction weights ($\hat{P}_{ij}^{-1} \times HHsize$); or Ginis on an unweighted sample where each household is given equal weight and only nonresponse correction is applied ($\hat{P}_{ij}^{-1}$). These alternative Ginis are available on request.

data). The difference is statistically insignificant at the 5 percent level, but borderline significant at the 10 percent level in a number of models. Logarithmic functional form appears to outperform linear, square root, or polynomial forms in terms of various model statistics, and market income outperforms gross and disposable incomes, as well as incomes per capita.

The logarithmic and quadratic logarithmic models of market income, models 1 and 7 (underlined for ease of reference), provide the best fit among the considered models.[9] They show response probability gradually declining with households' income level, slightly more dramatically so in model 7. The richest households in the sample have a predicted response probability as low as 0.182 in model 1, and 0.065 in model 7 (refer to figure A1-5).

In what follows, model 1 is used as the model of choice, for its simplicity and conformance with prior studies. Table 3 shows the central implications of this model for the estimated redistributive impact of fiscal policies. Compared to the results in table 1, the Gini rises by 3.5–4.0 percentage points for all income concepts. The largest correction occurs with taxable income. Whether this large correction could be attributed to the clustering of top taxable incomes or the absence of nontaxable incomes in the top of the gross income distribution is unclear, and should be explored further. Across all income concepts, the upward correction in the Gini tends to slightly increase as one moves from market income to final income, lowering the equalizing impact of cash-like transfers, direct taxes and in-kind transfers found in table 1 (refer to the summary in table 7).

In regard to top income shares, the aggregate-income share of the top 0.1 percent of households rises by 1.9–2.4 percentage points, representing a 56–76 percent gain on the uncorrected share (3.3 pc.pt. for taxable income representing an 84% gain). Analogously, the share of the top percentile of households rises by 3.9–5.3 percentage points (33–37% gain), the share of the top ventile rises by 4.3–5.5 percentage points (15–17% gain), and the share of the top decile rises by 4.0–4.9 percentage points (10–11% gain). These increases are similar in magnitude across all income concepts.[10] Combined with our findings for the Gini, we conclude that the redistributive impact of fiscal instruments is robust to the correction for unit nonresponse but becomes slightly less equalizing.

In regard to the impact of individual fiscal instruments, our conclusions from table 1 remain valid with three notable differences. One, after correcting for rich-households' nonresponse, contributory pensions are found to have a weak equalizing impact as gauged by the Gini. Two, we now find that households' nontaxable income – the difference between taxable and gross income – has an even more equalizing impact than in table 1. This is the case with all top income shares and the Gini, and particularly with the income share of the top 0.1 and 1 percent of households.

[9] Table A1-2 extends this analysis by considering multivariate functions of income and additional characteristics of household heads: age, age squared, gender, household size, and binary indicators of urban/rural residence, current attendance of post-secondary schools, and formal employment status. Comparing tables 4 and 5 suggests that simple univariate functions of income in table 4 have better efficiency properties with respect to various model statistics. They yield greater significance of coefficients, lower value of the minimization objective function, lower values of the information criteria, and lower standard errors on the Ginis.

These findings are supported in the ENIGH 2012 data (tables A2-4, A2-5). On the other hand, in the ENIGH 2010, income is consistently estimated to have a positive effect on response probability ($\widehat{\theta_1} < 0$), and the corrected Ginis are typically lower than the uncorrected ones (tables A3-3, A3-4) – confirming what we observed in figure A3-1.

[10] Figure A1-4 illustrates the change in the Lorenz curves for market income and for disposable income – the corrected curves are shown to be strictly dominated and show more inequality than the uncorrected curves across different parts of the income distribution. In ENIGH 2012, we find much lower differences in Lorenz curves (figure A2-5). In ENIGH 2010, interestingly, the nonresponse-corrected Lorenz curve dominates the uncorrected curve in the top third of the income distribution, implying that the correction lowers the estimated inequality at the top (figure A3-1 column 2).

Three, we also find that direct taxes – the difference between gross and disposable income – have an unequalizing impact as gauged by a higher top 0.1% share (and GE(2) index), even though they preserve their equalizing impact on other inequality indexes from table 1.[11]

Figure 2 panels *i-ii* illustrate the cumulative redistributive impacts of fiscal instruments between pre-fiscal (market) income and post-fiscal (disposable) income, both before and after the correction for rich-households' nonresponse. Disposable income Lorenz curves dominate those for market income, confirming the equalizing impact of fiscal policies taken together. Comparing the nonresponse-corrected and uncorrected Lorenz curves shows that the estimated equalizing impact increases – the difference between market and disposable Lorenz curves increases – among the top 25 percent of households. (Figures A2-8 and A3-6 show the analogous results for ENIGH 2012 and 2010, respectively.)

Correcting top incomes through replacing with Pareto (type I) estimates

In this section we embark on correcting the distributions of various income concepts for impurities in their top tails due to misreporting or mis-recording of some income components. The correction is done by the means of replacing a suspected range of top incomes with smooth estimates from a known statistical distribution, here the Pareto distribution of type I.

Table 4 reports the summary results of this procedure (table A1-3 and figure A1-6 provide additional results). The first two blocks of rows in table 4 show the observed distribution of bottom and top incomes in the 2014 ENIGH sample – the Gini coefficients among incomes classified as 'bottom' and 'top,' under alternative cutoff points. 'Top' incomes are those used to fit the Pareto distribution and to be replaced with synthetic values from that distribution.

The third block of rows thus reports the estimated Pareto (type I) coefficients α for individual income concepts and various lower cutoff points. Coefficients α estimated with a lower cutoff point at the 75th or even the 90th percentile (1.59–1.85 across all but one income concept) are substantially lower than the α estimated with a cutoff at the 95th or the 99th percentile (1.84–2.09 across all income concepts). Among the richest 0.1 percent of households, the α is estimated high, at 2.73–3.61. This indicates that the dispersion of incomes is wide among incomes in the 75th to 95th percentile, narrower among 95–99.9th percentile incomes, and narrower still in the uppermost tail, relative to what would be predicted under a single underlying Pareto distribution.

Correspondingly, inequality measures estimated in the rest of table 4 show low degrees of inequality among the topmost 0.1 percent of households, fair inequality among the top 1–5 percent, and higher inequality among the top 10 or especially all top 25 percent. The inverted Pareto coefficient β is estimated at 1.92–2.19 among the top 1–5 percent of households, but rises to 2.18–2.70 among the top 10–25 percent (with one exception), and falls to 1.38–1.58 among the top 0.1

[11] The same analysis performed on the 2012 round of the ENIGH (tables A2-1 and A2-6) reveals essentially the same qualitative patterns, but of much smaller magnitudes on account of a weaker link between incomes and response probability. Across all considered models, the Ginis increase by 0.3-6.7 pc.pt. (see table A2-4). Using the preferred model (same as in 2014), the Ginis increase by 0.4 pc.pt. across all income concepts, and the top income shares increase by 0.2–0.3 pc.pt. for the top 1% share, and by 0.4–0.5 for the top 10% share (see table A2-6). Because of the weak corrections for nonresponse, pensions are found to have an unequalizing impact as measured by the top 5% and 10% income share as well as the Gini, both before and after the correction.

In the 2010 round, the only two available income concepts allow us to comment only on the redistributive impact of cash-like after-tax transfers (see tables A3-1 and A3-3). The correction for nonresponse leads to non-positive changes in the Gini and the top income shares in most estimated models. The benchmark estimates in table A3-1 show that the Gini increases by 0.03–0.15 pc.pt., the top percentile share increases by 0.4pc.pt., and the top decile share increases by 0.6–0.7pc.pt., all very small effects. The redistributive impact of cash-like after-tax transfers is equalizing, at 2.6 points of the Gini, which increases to 2.7 points after the correction for unit nonresponse.

percent (refer to table A1-3). These differences are statistically highly significant, as likelihood ratio tests confirm. The top income shares, the parametric Gini, and half the coefficient of variation squared (when available; refer to table A1-3) show the analogous patterns of a low degree of income dispersion among the top 1–5 percent of households, but high inequality among the top 10–25 percent or among the topmost 0.1 percent.

Using the above parametric estimates among the top tail of incomes, we re-estimate inequality under the entire income distribution. The bottom rows of table 4 report the results. Compared to the original uncorrected Ginis in table 1, the Ginis corrected for suspected top income mismeasurements are systematically higher, by 0.2–4.5 percentage points across all income concepts and all choices regarding cutoff points (except when only the top 0.1% of observations are replaced). These are systematic and sizeable corrections.[12,13] (The original and corrected Lorenz curves for market income and disposable income are illustrated in figure A1-6.)

The upward corrections to the Gini are highest for taxable income, just as we saw with the correction for unit nonresponse. In tandem, these findings suggest that the distribution of taxable income may be less smooth, and is affected when the nonresponse weights or the parametric replacing are applied to its top values. In contrast to the corrections for unit nonresponse, the corrections in table 4 appear to slightly decline as we move from market income to final income. The impact of fiscal policies is thus estimated to be more equalizing. Cash-like transfers are estimated to reduce the Gini by 1.8–1.9 points, and direct taxes by 1.5–2.2 (see table 7, and figure 2 panel *iii*). In-kind transfers reduce the Gini by 4.9–5.7 points, mean 5.6 across the different cutoff points. Contributory pensions are found to be unequalizing as gauged by the rising Gini as well as the top 10 and top 25-percent income share, but equalizing at the complete top, as gauged by the falling top 5, 1 or 0.1-percent shares.

---

[12] In the 2012 round of the ENIGH, a different pattern emerges. Among the richest 0.1 percent of households, the coefficient α becomes low again (1.52–1.76), in part due to the outlier in the sample (see table A2-7; figure A2-2). The inverted Pareto coefficient β is estimated at 1.57–1.90 among the top 1–5% of households, but rises to 2.12–2.92 among the top 10–25% or among the top 0.1%. The corrected semiparametric Ginis for the entire income distribution are systematically higher than the uncorrected Ginis in table A2-1 (with a single exception), by 0.0–5.1 pc.pt.

In the 2010 round, the estimated Pareto coefficient rises gradually as the lower cutoff is increased, and the corresponding measures of inequality fall (see table A3-5). The corrected Gini for disposable income net of transfers is 4.0 pc.pt. higher than the uncorrected Gini when all of the top 25% of incomes are replaced, but by as little as 0.2 pc.pt. when only the top 1% are replaced (2.9 and 0.1 pc.pt., respectively, for disposable income). In light of the sensitivity of the α parameter to the cutoff points, across all survey rounds and various income concepts, the Pareto (I) approximation may not be adequate when we include lower incomes such as those in the 75–90th percentile.

[13] As a robustness test, we replace top incomes with estimates from the generalized Pareto (type II) distribution. Its cumulative density function is $F(x) = 1 - [1 + x/\sigma]^{-\xi}$ where ξ is a shape parameter and σ is a scale parameter (e.g., Arnold 2008). The Gini under the parametric distribution can be computed as:
$Gini = 1 - [L + 2\sigma\xi^{-2}G(2\xi^{-1} - 1)G(2)G(2\xi^{-1} + 1)]/[L + \sigma\xi^{-2}G(\xi^{-1} - 1)G(2)G(\xi^{-1} + 1)]$ where L is the lower cutoff income and G is the gamma function. Refer to table A1-4 and figure A1-8 in the appendix (for ENIGH 2010 and 2012, refer to tables A2-10 and A3-6, and figures A2-7 and A3-6). Here, the shape coefficient does not change much across different income concepts and different delineations of top incomes, except for the extreme cutoff at the 99.9th percentile. Using the 99.9th percentile cutoff, the shape parameter is estimated much lower for market income+pensions and gross income, suggesting that a simple exponential distribution may describe the upper-most dispersion for these income concepts adequately. For the rest of income concepts, the shape parameter among the top 0.1% of incomes is estimated higher.

The estimated generalized Pareto (type II) distribution yields higher shape coefficients, lower inverted Pareto coefficients, and lower Ginis than the Pareto I specification (refer to figure A1-8; or figures A2-7 and A3-6 for prior surveys). Because the Pareto I specification is viewed as less sensitive to individual income values, we use Pareto I as the main model.

As in tables 1 and 3, nontaxable income has an equalizing impact in terms of most inequality indicators. However, here the impact becomes neutral or even slightly unequalizing at the top tail gauged by the income share of the top 0.1 percent of households, and the equalizing impact becomes stronger lower down in the income distribution – for a 4 percentage point decrease in the Gini (compared to 3.3 in tables 1 and 3).

Correcting top net-market incomes through Pareto (type I) replacing, and using CEQ Method

One modality of the Pareto replacing method is to replace the top tail for the income concept thought to be most susceptible to mismeasurement, and then re-calculate the rest of the derived income concepts using the CEQ Method. Net market income per capita, being taken directly from survey questionnaires and being the starting concept in the CEQ methodology from which other income concepts are imputed, is the natural choice for this exercise.

The model used here is the same as in the previous section for net market income per capita (table 4 column 5). For the main model specification, we choose the lower cutoff at the 90th percentile. Hence, reported incomes above MXN80,362 are viewed as potentially misreported and are replaced with random draws from the estimated Pareto distribution. This model specification is expected to produce intermediate estimates of the Pareto coefficient and intermediate corrections of inequality compared to the ranges presented in table 4. (Tables 6 and 7 report on confidence intervals using alternative cutoff points for replacing).

Table 5 reports the corrected distributions of all income concepts. By design, the Gini coefficient and the top 10 percent income share obtained for net market income per capita is nearly identical to that in table 4 (52.68 vs. 52.60 for the Gini, 42.49 vs. 42.40 for the top 10% share), the minute differences being due to random drawing from the Pareto distribution. Across income concepts and the intervening fiscal instruments, our findings are similar to those in table 4. Pensions have a neutral impact, slightly equalizing among the topmost one percent of incomes, and slightly unequalizing among lower incomes, leading to no change in the Gini. Adding of nontaxable incomes has an equalizing effect of a similar magnitude as in table 4, of 3.8 points of the Gini. The equalizing impact of direct taxes is estimated to be around 1.4 points of the Gini, and that of in-kind programs around 4.9 points, slightly lower than in tables 3 and 4. (Figure 2 panel *iv* illustrates the cumulative redistributive impact of fiscal instruments, using the entire Lorenz curve. Figure A1-7 illustrates the impact on the distribution of market incomes and disposable incomes, similarly as in figures A1-4 and A1-6.)[14]

The main results for the redistributive impact of fiscal instruments

Tables 6 and 7 summarize the main results of the corrections implemented in this study, including the ranges of inequality estimates and of the estimated redistributive impacts of fiscal programs. Table 6 shows the full ranges (mean, median and the extrema) of the Ginis for each income concept estimated under various behavioral specifications of the reweighting model in table 2, and various cutoffs for the Pareto replacing in table 4. For convenience, the table reports the estimated changes on the uncorrected Ginis (in bold). The reweighting method is shown to correct the Gini upward by 4.3–5.2 percentage points on average (median 4.1–4.8), and by as much as 10.1. The replacing method corrects the Ginis by 0.7–1.4 percentage points on average (median 0.5–0.9), and by as much as 4.5. Reassuringly, the two modalities of the replacing method lead to similar magnitudes of corrections for all income concepts.

[14] Table A2-9 and figures A2-7 and A2-8 show the analogous results for the 2012 round of the ENIGH.

Table 7 offers a slightly different perspective on the estimated redistributive impact of fiscal policies. Instead of showing the estimates of inequality, percentage point differences across income concepts are shown. Moreover, instead of showing the full range of estimates including outliers, three parametric forms of the correction methods are chosen as representing the low, central and high among reasonable specifications. Under the reweighting approach, the per-capita model 4, the household market income model 1, and the quadratic model 7 – exhibiting good theoretical justification, consistency with one another, and empirical fit – are used as the low, central and high specifications. Of course, the redistributive impacts estimated under these low-to-high specifications may not be related monotonically to one another and may not be ranked from low to high. Under the replacing approach, lower cutoffs at the 75th, 90th, and 99th percentiles – again, models showcasing some theoretical justification and adequate empirical fit – are used as the low, central and high specifications.

For all three specifications under each correction approach, table 7 reports the percentage point changes in the Ginis and in the top income shares attributable to specific fiscal instruments – that is, the differences in inequality indices between the pairs of adjacent income concepts. Table 7 confirms that pensions have a negligible impact on inequality, and in fact the impact varies across quantiles of the income distribution (as seen by the changes in the Gini vs. the top income shares). Cash-like transfers have a strong equalizing impact of 1.6–1.9 percentage points of the Gini. The impact of nontaxable income is stronger still, at 3.3–4.5 percentage points, where the corrected estimates are universally larger than or as large as the uncorrected figure (3.3pc.pt.). Progressive direct taxes account for another 1.2–2.2 percentage point drop in the Gini. Indirect taxes and subsidies have a weak equalizing impact of 0.4-0.6 points of the Gini, but again this is universally larger than the uncorrected figure (0.4pc.pt.). In-kind transfers have a strong equalizing impact of 4.7–5.7 points of the Gini, again typically larger than the uncorrected impact (4.8pc.pt.).[15]

VI. Summary and Discussion

This study has evaluated the redistributive impacts of various instruments of fiscal policy in Mexico, using the 2010–2014 ENIGH surveys, and applying two specific corrections for potential top-income measurement problems. We have first reweighted the survey sample to correct the income distribution for selective nonresponse by rich households, and then we have replaced potentially mismeasured top incomes with synthetic values from a smooth parametric distribution. By comparing the uncorrected measures of inequality and degrees of redistributive fiscal impacts with the two alternative sets of corrected figures, we have evaluated the robustness of the uncorrected figures and provided improved estimates.

The key result of the study is that pensions in Mexico are confirmed to be inequality-neutral, while in-kind transfers, cash-like transfers and direct taxes have strong equalizing impacts, of 4.7–5.7, 1.6–1.9, and 1.2–2.2 points of the Gini, respectively. Indirect taxes and subsidies are equalizing only weakly, by 0.4-0.6 points of the Gini.

The new estimates should not be considered to be accurate or unbiased, as they correct for a single source of imprecision at a time. However, because the corrected estimates were obtained

[15] Tables A2-10 and A2-11 show the analogous results for the 2012 round of the ENIGH, while table A3-6 shows the results for 2010. Qualitative results for 2012 are very similar to those in tables 6 and 7, but the magnitudes are smaller. These results suggest that the top-income biases are at least as worrying in 2014 as in 2012. In 2010, after-tax cash-like transfers have an equalizing impact of 3pc.pt. of the Gini, and 2.3pc.pt. of the top-10% income share, but these figures are not comparable to the impacts of before-tax transfers assessed in rounds 2012–2014.

using established and transparent methods, and using rather conservative modeling specifications, they can be viewed as improved baseline estimates which can be evaluated for other biases. Both the uncorrected and corrected estimates have large standard errors, suggesting that sampling error tends to dominate estimation error, but the differences in estimates are quite consistent, and significant in a number of cases.

Tables 6 and 7 summarize the main results of this paper for the estimates of inequality and fiscal redistributive impacts. Across the board, corrections to the Gini coefficients and top income shares are positive, suggesting that the uncorrected statistics suffer from a downward bias, and the corrected estimates of the redistributive impacts are qualitatively similar to the uncorrected impacts, which helps to validate our methods. The corrected estimates of the redistributive impacts do differ quantitatively from the uncorrected ones in a number of cases, and the differences are systematic.

As we move from pre-fiscal toward post-fiscal incomes, the corrections to inequality estimates under the reweighting method somewhat increase, while the corrections under replacing tend to fall or stagnate. This suggests that measurement problems differ under the different income concepts, or that households' pre-fiscal and post-fiscal incomes are associated in non-obvious ways. Correcting for unit nonresponse through monotonic reweighting of top observations reduces the equalizing redistributive impact of fiscal programs. This may be due to limited progressivity of taxes and transfers, or fiscal loopholes among households with top taxable incomes, whose weight is corrected the most.

Income measurement issues as evidenced by comparisons to smooth Pareto distributions appear to affect most seriously the distribution of pre-fiscal incomes, for which the estimated biases are larger. Taxable income is the income concept most heavily affected by both unit nonresponse and mismeasurement. Whether the large estimated biases are on account of unreported taxable earnings or some clustering of top taxable incomes, is unclear and should be explored further. Nontaxable income is shown to be even more equalizing after correcting gross income than in the uncorrected distribution, suggesting that nontaxable incomes are not too prevalent in the top tail of the gross income distribution where the bulk of the upward corrections – by reweighting or replacing – takes place.

Finally worth noting, the corrections for possible misreporting of top incomes, by Pareto replacing, can be compared to the corrections for unit nonresponse by reweighting, to judge the relative gravity of these two distinct problems (as shown in tables 6 and 7). Interestingly, the mean, median and maximum corrections to the Gini are substantially higher under the reweighting method.

Unit nonresponse leads to substantial underestimation of mean incomes and measures of inequality. The Gini coefficient of market income per capita is found to be biased downward by up to 8.2 percentage points, and typically by 4.3 points across all estimations. The Gini for final income is biased downward by up to 8.8 points, and typically by 4.4 points. By contrast, the suspected tainting of the distribution of top incomes by income mismeasurement biases the Gini of market income per capita by up to an estimated 3.1 points, and typically by only 0.7 points, and these biases fall to 2.3 and 0.5 points for final income per capita.

The study confirms that unit nonresponse is a systematic and non-negligible problem in the Mexican ENIGH survey. Along with other top-income measurement challenges, unit nonresponse retains its magnitude across the 2010, 2012 and 2014 rounds of the ENIGH. Moreover, household nonresponse becomes more positively selected over time, causing more serious measurement biases. Analysts and policymakers relying on ENIGH would be wise to take note.

References

Alvaredo, Facundo, Santiago Garriga, and Florencia Pinto (2017). Household Surveys, Administrative Records, and National Accounts in Mexico 2009–2014. Is a Reconciliation Possible? Paris: Paris School of Economics unpublished paper.

Anand, Sudhir, and Paul Segal (2017) Who Are the Global Top 1%? *World Development* 95:111–126.

Aranda, Rodrigo, and John Scott (2015) CEQ Master Workbook for Mexico. Mimeo, Commitment to Equity Institute.

Aranda, Rodrigo, and John Scott (2016) Fiscal Incidence by Race and Ethnicity: Master workbook for Mexico, prepared by the Commitment to Equity Project for the Inter-American Development Bank, Programa para Mejorar Las estadísticas de Raza y Etnicidad para el análisis y formulación de Políticas.

Arnold, Barry C. (2008) Pareto and generalized Pareto distributions. In: Duangkamon Chotikapanich (ed.) Modeling Income Distributions and Lorenz Curves, pp. 119–145. New York: Springer.

Atkinson, Anthony, Thomas Piketty, and Emmanuel Saez (2011) Top incomes in the long run of history, *Journal of Economic Literature* 49(1):3–71.

Bourguignon, Francois (2018) Simple adjustments of observed distributions for missing income and missing people, *Journal of Economic Inequality* 16(2):171–188.

Brachmann, Klaus, Andreas Stich, and Mark Trede (1996) Evaluating parametric income distribution models. *Allgemeines Statistisches Archiv* 80:285–98.

Brzezinski, Michał (2013) Parametric modelling of income distribution in Central and Eastern Europe. *Central European Journal of Economic Modelling and Econometrics* 3:207–230.

Bustos, Alfredo (2015) Estimation of the distribution of income from survey data, adjusting for compatibility with other sources, *Statistical Journal of the IAOS* 31(4):565–577.

Bustos, Alfredo, and Gerardo Leyva-Parra (2017) Towards a more realistic estimate of the income distribution in Mexico, *Latin American Policy* 8(1):114–126.

Butler, Richard J., and James B. McDonald (1989) Using incomplete moments to measure inequality, *Journal of Econometrics* 42:109–119.

Campos-Vázquez, Raymundo M., Gerardo Esquivel, and Nora Lustig (2012) The Rise and Fall of Income Inequality in Mexico, 1989-2010. UNU-WIDER Technical Report 010.

Campos-Vázquez, Raymundo M., Nora Lustig, John Scott (2018) Inequality in Mexico: Labour markets and fiscal redistribution 1989–2014, WIDER Working Paper 2018/188.

Castelletti, Barbara (2013) How Redistributive Is Fiscal Policy in Latin America? The Case of Chile and Mexico, OECD Development Centre Working Paper 318.

Cornia, Giovanni A., ed. (2014) Falling Inequality in Latin America: Policy Changes and Lessons. Oxford: Oxford University Press.

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

Davidson, Russell, and Emmanuel Flachaire (2007) Asymptotic and bootstrap inference for inequality and poverty measures, *Journal of Econometrics* 141(1):141–166.

Deaton, A. (1997), The analysis of household surveys: a microeconometric approach to development policy, Washington: World Bank.

Esquivel, Gerardo (2011) The Dynamics of Income Inequality in Mexico since NAFTA, Economia: Journal of the Latin American and Caribbean Economic Association 12(1):155–188.

Faulkner, Christina M. (2014) Using G2 to measure income inequality in two Latin American upper middle income countries, *Statistical Journal of the IAOS* 30:321–329.

Félix, David (1982) Income Distribution Trends in Mexico and the Kuznets Curves, in Brazil and Mexico: Patterns in Late Development, Sylia Ann Hewlett and Richard S. Weinert (eds.), Philadelphia: Institute for the Study of Human Issues.

Freije, S., G. López-Acevedo, G., and E. Rodríguez-Oreggia (2014). Effects of the 2008–09 Economic Crisis on Labor Markets in Mexico, in Banerji, Arup, David Newhouse, Pierella Paci, and David Robalino, (eds.) Working through the Crisis: Jobs and Policies in Developing Countries during the Great Recession. Directions in Development. Washington: World Bank.

Goñi, Edwin, J. Humberto López, and Luis Servén (2011) Fiscal redistribution and income inequality in Latin America, Fiscal Redistribution and Income Inequality in Latin America, *World Development* 39(9):1558–1569.

Hill, Bruce M. (1975) A simple general approach to inference about the tail of a distribution. Annals of Statistics 3:1163–1174.

Hlasny, Vladimir (2016) Unit nonresponse bias to inequality measurement: worldwide analysis using Luxembourg Income Study database. Luxembourg Income Study technical WP 8.

Hlasny, Vladimir, and Paolo Verme (2018a) Top incomes and inequality measurement: a comparative analysis of correction methods using the EU SILC data, *Econometrics* 6(2):30, June.

Hlasny, Vladimir, and Paolo Verme (2018b) Top Incomes and the Measurement of Inequality in Egypt, *World Bank Economic Review* 32(2):428–455.

Hlasny, Vladimir, and Paolo Verme (2018c) The impact of top incomes biases on the measurement of inequality in the United States. ECINEQ WP 452.

Jenkins, Stephen P. (2007) Inequality and the GB2 income distribution. IZA Discussion Papers 2831, Institute for the Study of Labor (IZA), Bonn.

Jenkins, Stephen P. (2009) Distributionally-sensitive inequality indices and the GB2 income distribution, *Review of Income and Wealth* 55: 392–98.

Jenkins, Stephen P., Richard V. Burkhauser, Shuaizhang Feng, and Jeff Larrimore (2011) Measuring inequality using censored data: a multiple-imputation approach to estimation and inference, *Journal of the Royal Statistical Society* 174(1):63–81.

Jenkins, Stephen P. (2017) Pareto distributions, top incomes, and recent trends in UK income inequality, *Economica* 84(334):261–289.

Korinek, Anton, Johan A. Mistiaen, and Martin Ravallion (2006) Survey nonresponse and the distribution of income, *Journal of Economic Inequality* 4:33–55.

Korinek, Anton, Johan A. Mistiaen, and Martin Ravallion (2007) An econometric method of correcting for unit nonresponse bias in surveys, *Journal of Econometrics* 136:213–235.

Lakner, Christoph, and Branko Milanovic (2013) Global income distribution from the fall of the Berlin Wall to the great recession, World Bank Policy Research working paper series #6719.

Leyva-Parra, Gerardo (2004) El ajuste del ingreso de la ENIGH con la Contabilidad Nacional y la medición de la pobreza en México, Documentos de Investigación 19, SEDESOL, México.

Lindert, Kathy, Emmanuel Skoufias, and Joseph Shapiro (2006) Redistributing Income to the Poor and Rich: Public Transfers in Latin America and the Caribbean. Social Protection Discussion Paper 0605. Washington: World Bank.

Lopez-Calva, Luis Felipe, Nora Lustig, John Scott, and Andres Castañeda (2018) Gasto social, redistribución del ingreso y reducción de la pobreza en México: evolución y comparación con Argentina, Brasil y Uruguay. In Rodolfo de la Torre, Eduardo Rodríguez-Oreggia e Isidro Soloaga (Eds.), *Política social y bienestar: México desde el año 2000.* México: CIDE-FCE.

Lustig, Nora (1998) *Mexico: The Remaking of an Economy*, Washington: Brookings Institution, 2nd Edition.

Lustig, Nora (2016) Inequality and Fiscal Redistribution in Middle Income Countries: Brazil, Chile, Colombia, Indonesia, Mexico, Peru and South Africa, *Journal of Globalization and Development* 7(1):17–60.

Lustig, Nora, Ed. (2018) Commitment to Equity Handbook: Estimating the Impact of Fiscal Policy on Inequality and Poverty, New Orleans: CEQ Institute at Tulane University, and Washington: Brookings Institution Press.

Lustig, Nora, Judith Morrison, and Adam Ratzlaff (2019) Splitting the Bill: Taxing and Spending to Close Ethnic and Racial Gaps in Latin America, Inter-American Development Bank Gender and Diversity Division (SCL/GDI) monograph.

Lustig, Nora, Carola Pessino, and John Scott (2014) The Impact of Taxes and Social Spending on Inequality and Poverty in Argentina, Bolivia, Brazil, Mexico, Peru, and Uruguay, Public Finance Review 42(3):287–303.

Martorano, Bruno (2014), Pre-crisis Conditions and Government Policy Responses: Chile and Mexico during the Great Recession, UNICEF Office of Research Working Paper 2014-15.

Mistiaen, Johan A., and Martin Ravallion (2003) Survey compliance and the distribution of income, Policy Research Working Paper #2956, Washington: World Bank.

Navarrete, Ifigenia Martinez de (1973) La Distribución del Ingreso en México: Tendencias y Proyeccion a 1980, SOLIS, L. La economía mexicana. Vol. 1, México, p.278–291.

Novokmet, Filip, Thomas Piketty, and Gabriel Zucman (2017) From Soviets to oligarchs: inequality and property in Russia 1905-2016. Journal of Economic Inequality 16(2):189–223.

Olascoaga, Sebastián Sandoval (2015) The Distribution of Top Incomes in Mexico: How rich are the richest? Paris School of Economics, Public Policy and Development Master's Thesis, http://piketty.pse.ens.fr/files/Sandoval2015.pdf (accessed 4 September 2018).

Pareto, Vilfredo (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, Thomas, Li Yang, and Gabriel Zucman (2016) Capital Accumulation, Private Property and Rising Inequality in China 1978-2015, WID.world Working Paper.

Reyes, Miguel, Graciela Teruel, and Miguel López (2017) Measuring True Income Inequality in Mexico. Latin American Policy 8(1):127–148.

Reynolds, Morgan, and Eugene Smolensky (2013) Public Expenditures, Taxes, and the Distribution of Income: The United States, 1950, 1961, 1970, London: Academic Press.

Ros, J. (2011). The Economic Crisis of 2008–09 and Development Strategy: the Mexican Case, in Haddad, M. and B. Shepherd (Eds.), Managing Openness: Trade and outward-oriented growth after the crisis, Washington: World Bank.

Scott, John (2008) Redistributive Constraints under High Inequality: The Case of Mexico, CIDE WP 441.

Scott, John (2010) Commitment to Equity: An Assessment of Fiscal Policies in Mexico. CEQ Assessment Report, Washington DC: Inter-American Dialogue; New Orleans, LA: Center for Inter-American Policy & Research (CIPR), Tulane University.
Scott, John (2013) CEQ Master Workbook: Mexico. Version: September 2, 2013, CEQ Data Center on Fiscal Redistribution (CEQ Institute, Tulane University).
Scott, John (2014) Redistributive Impact and Efficiency of Mexico's Fiscal System. In: Lustig, Nora, Carola Pessino and John Scott, eds., The Redistributive Impact of Taxes and Social Spending in Latin America. Special Issue. Public Finance Review 42(3):368–390.

Table 1. Income summary statistics, various income concepts: uncorrected data

	(1) Market income per cap.	(2) Market income + pensions per cap.	(3) Gross income per cap.	(4) Taxable income per cap.	(5) Net market income per cap.	(6) Disposable income per cap.	(7) Consumable income per cap.	(8) Final income per cap.
99.9th %ile	1,052,184	1,051,861	1,051,861	776,087	918,303	918,303	885,516	887,620
99th %ile	304,486	311,864	311,864	221,505	273,162	273,162	262,714	265,543
95th %ile	125,755	133,943	135,000	90,492	120,148	120,867	115,837	120,053
90th %ile	83,714	88,416	89,443	60,492	80,362	81,033	78,427	83,243
75th %ile	46,477	48,042	48,594	32,983	44,590	45,144	43,769	48,405
Mean	43,178	44,912	45,985	30,242	41,045	42,117	40,767	45,374
Median	25,983	26,941	27,747	18,010	25,443	26,286	25,659	30,340
25th %ile	14,685	15,298	16,642	9,192	14,754	16,044	15,782	20,377
10th %ile	8,115	8,454	10,353	2,779	8,354	10,226	10,112	14,357
5th %ile	4,822	5,006	7,323	492	4,950	7,255	7,207	11,480
1st %ile	1,751	1,796	3,944	0	1,796	3,903	3,917	7,098
Std. dev.	81,471	82,794	82,602	64,453	73,396	73,222	69,851	69,697
Skewness	18.29	17.53	17.63	27.29	20.75	20.88	20.57	20.63
Kurtosis	685.48	641.69	646.90	1,456.07	925.56	933.40	909.34	914.54
Sample	73,508	73,508	73,508	73,508	73,508	73,508	73,508	73,508
Top 0.1% inc. share	3.65	3.54	3.46	3.93	3.26	3.18	3.14	2.82
0.1-1% inc. share	9.94	9.67	9.45	10.41	9.27	9.04	8.93	8.08
1-5% inc. share	16.72	16.98	16.64	17.16	16.55	16.20	16.05	14.75
5-10% inc. share	11.79	11.95	11.74	12.07	11.80	11.58	11.53	10.87
Gini (HH-size & sampling weighted data)	52.75 (0.82)	52.79 (0.79)	50.99 (0.80)	54.31 (0.86)	51.33 (0.76)	49.43 (0.77)	49.00 (0.77)	44.17 (0.76)
Mean log dev. (GE0)	0.515 (0.009)	0.516 (0.009)	0.456 (0.008)	0.594 (0.010)	0.484 (0.008)	0.426 (0.008)	0.417 (0.008)	0.328 (0.007)
Theil index (GE1)	0.598 (0.020)	0.593 (0.019)	0.559 (0.019)	0.637 (0.024)	0.557 (0.018)	0.522 (0.018)	0.513 (0.018)	0.424 (0.016)
Half coef. of var. squared (GE2)	1.780 (0.198)	1.699 (0.184)	1.613 (0.176)	2.271 (0.365)	1.599 (0.207)	1.511 (0.197)	1.468 (0.190)	1.180 (0.154)

Notes: Sampling-weighted sample used. 0MXN incomes (19 household observations for market income, 7 for market+pensions & net market income, 2 for gross & disposable income, and 1,465 for taxable income) are omitted in computations of the Gini. Gini standard errors are jack-knife estimates on household-level data (recognizing that household-member incomes are copies of one another), accounting for household size except in last row. Ginis and standard errors are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2014, CEQ database.

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

Specification of $g(x_i,\theta)$	$\widehat{\theta_0}$ (s.e.)	$\widehat{\theta_1}$ (s.e.)	Sum of squared wted errors	Factor of proportionality (σ^2)	AIC SIC	Gini (s.e.): Market income per capita , weighted data

Uncorrected						52.75 (0.82)
1: $\theta_0+\theta_1$log(market inc.)	14.464	-0.960	35,242	0.448	228.14	56.41 (1.92)
	(0.078)	(0.006)			225.52	
2: $\theta_0+\theta_1$log(gross inc.)	14.685	-0.972	36,116	0.460	228.92	56.30 (1.90)
	(0.082)	(0.007)			226.31	
3: $\theta_0+\theta_1$log(dispos. inc.)	16.087	-1.091	36,080	0.454	228.89	57.64 (2.71)
	(0.074)	(0.006)			226.27	
4: $\theta_0+\theta_1$log(mkt. inc.pc)	8.667	-0.537	38,450	0.486	230.92	53.78 (0.98)
	(0.062)	(0.006)			228.31	
5: $\theta_0+\theta_1$log(gross inc.pc)	8.594	-0.524	39,263	0.496	231.59	53.68 (0.96)
	(0.065)	(0.006)			228.98	
6: $\theta_0+\theta_1$log(disp. inc.pc)	8.645	-0.532	39,412	0.499	231.72	53.65 (0.96)
	(0.067)	(0.006)			229.10	
7: $\theta_0+\theta_1$log(market inc.)^2	9.261	-0.043	34,425	0.420	227.39	60.17 (4.12)
	(0.024)	(0.000)			224.77	
8: $\theta_0+\theta_1 10^{-6}$ market inc.	3.087	-0.373	40,078	0.476	232.25	59.76 (7.08)
	(0.002)	(0.001)			229.64	
9: $\theta_0+\theta_1 10^{-15}$market inc.^{2}	3.029	-22.056	41,092	0.483	233.05	59.36 (6.97)
	(0.002)	(0.056)			230.44	
10:$\theta_0+\theta_1(10^{-3}$ mkt. inc.$)^{½}$	3.675	-0.053	36,301	0.451	229.08	60.94 (7.20)
	(0.003)	(0.000)			226.47	

Note: Standard errors on Gini coefficients are jackknife estimates.
Source: Own analysis of ENIGH 2014, CEQ database.

Table 3. Income summary statistics for various income concepts, nonresponse corrected weights

	(1) Market income per cap.	(2) Market inc + pensions per cap.	(3) Gross income per cap.	(4) Taxable income per cap.	(5) Net market income per cap.	(6) Disposable income per cap.	(7) Consumable income per cap.	(8) Final income per cap.
99.9th %ile	1,265,377	1,310,623	1,310,623	802,174	1,028,962	1,028,962	977,821	979,925
99th %ile	361,257	370,574	370,574	269,022	324,344	324,344	312,031	315,450
95th %ile	138,716	148,594	148,885	98,607	132,807	133,243	128,139	131,867
90th %ile	90,263	94,975	95,498	64,206	85,414	86,008	82,928	88,227
75th %ile	48,361	49,977	50,557	34,843	46,127	46,623	45,272	49,935
Mean	48,630	50,329	51,365	34,568	45,822	46,858	45,287	49,888
Median	26,802	27,653	28,562	18,689	26,023	26,879	26,209	31,017
25th %ile	15,090	15,739	17,025	9,443	15,164	16,424	16,110	20,676
10th %ile	8,364	8,698	10,590	3,005	8,626	10,412	10,330	14,603
5th %ile	5,015	5,179	7,467	600	5,157	7,420	7,398	11,686
1st %ile	1,823	1,866	4,033	0	1,866	3,960	4,026	7,162
Std. dev.	123,595	124,475	124,313	106,891	114,799	114,655	109,068	108,893
Skewness	20.01	19.59	19.65	26.58	22.84	22.91	22.70	22.75
Kurtosis	588.16	569.91	572.26	946.92	748.75	751.75	740.70	743.40
Sample	73,467	73,467	73,467	73,467	73,467	73,467	73,467	73,467
Top 0.1% inc. share	5.70	5.51	5.40	7.24	5.68	5.56	5.45	4.95
0.1-1% inc. share	12.31	12.01	11.77	12.43	11.18	10.94	10.83	9.88
1-5% inc. share	17.09	17.22	16.91	17.33	16.78	16.46	16.32	15.11
5-10% inc. share	11.38	11.57	11.39	11.47	11.41	11.22	11.18	10.60
Gini (HH-size & sampling weighted data)	56.41	56.29	54.62	58.31	54.86	53.10	52.65	47.89
	(1.92)	(1.86)	(1.90)	(2.39)	(2.01)	(2.04)	(2.03)	(2.02)
Mean log dev.(GE0)	0.591	0.587	0.528	0.680	0.554	0.495	0.486	0.390
	(0.021)	(0.021)	(0.020)	(0.027)	(0.021)	(0.021)	(0.020)	(0.018)
Theil index (GE1)	0.761	0.748	0.713	0.840	0.715	0.679	0.667	0.566
	(0.055)	(0.053)	(0.052)	(0.078)	(0.058)	(0.057)	(0.056)	(0.051)
Coef. of var. (GE2)	3.230	3.058	2.929	4.781	3.138	2.993	2.900	2.382
	(0.667)	(0.624)	(0.602)	(1.262)	(0.755)	(0.724)	(0.698)	(0.583)

Notes: Statistics are based on non-response correction weights estimated in the logarithmic model of market income (model 1). These statistics exclude 19 household (41 individual) observations with market income of 0. The statistics are still comparable to those in table 1, which are extremely robust to this exclusion (changing by 0.01 at most.) Another 1,446 household (2,889 individual) 0-income observations are omitted in computations of the Gini for taxable income. Gini standard errors are jack-knife estimates on household-level data (recognizing that household-member incomes are copies of one another), accounting for household size except in last row. Ginis and standard errors are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2014, CEQ database.

Table 4. Replacing top incomes of each income concept with Pareto I distribution: corrected inequality indexes

Cutoff percentile	(1) Market income per cap.	(2) Market inc + pensions per cap.	(3) Gross income per cap.	(4) Taxable income per cap.	(5) Net market income per cap.	(6) Disposable income per cap.	(7) Consumable income per cap.	(8) Final income per cap.
Nonparametric Gini among incomes classified as 'bottom'								
Top 25%	30.97 (.24)	30.86 (.25)	28.10 (.22)	33.31 (.28)	29.92 (.24)	27.13 (.21)	26.88 (.21)	22.68 (.17)
Top 10%	36.80 (.23)	36.89 (.23)	34.61 (.21)	38.63 (.25)	35.86 (.23)	33.52 (.21)	33.22 (.21)	28.84 (.18)
Top 5%	40.67 (.24)	40.89 (.24)	38.79 (.23)	42.27 (.26)	39.76 (.24)	37.59 (.22)	37.25 (.22)	32.73 (.20)
Top 1%	46.90 (.31)	47.24 (.32)	45.29 (.31)	48.33 (.32)	45.93 (.31)	43.94 (.30)	43.55 (.30)	38.81 (.29)
Top 0.1%	51.22 (.60)	51.33 (.57)	49.50 (.58)	52.74 (.61)	50.01 (.55)	48.10 (.56)	47.68 (.56)	42.88 (.55)
Nonparametric Gini among incomes classified as 'top'								
Top 25%	37.89 (1.48)	37.67 (1.43)	37.47 (1.42)	38.22 (1.59)	36.59 (1.38)	36.39 (1.37)	36.11 (1.37)	34.24 (1.35)
Top 10%	35.10 (2.03)	34.23 (1.98)	34.09 (1.98)	35.75 (2.21)	33.37 (1.94)	33.23 (1.94)	33.04 (1.94)	31.98 (1.92)
Top 5%	33.70 (2.52)	32.43 (2.49)	32.34 (2.49)	34.54 (2.81)	31.64 (2.49)	31.52 (2.48)	31.40 (2.48)	30.82 (2.47)
Top 1%	30.55 (3.70)	30.14 (3.73)	30.15 (3.72)	31.16 (5.02)	29.10 (4.33)	29.09 (4.32)	29.08 (4.28)	28.93 (4.26)
Top 0.1%	21.66 (11.82)	20.93 (11.91)	20.93 (11.89)	30.74 (16.65)	26.85 (15.20)	26.86 (15.17)	26.46 (15.26)	26.43 (15.25)
Pareto (type I) coefficient among 'top' incomes								
Top 25%	1.62 (.03)	1.59 (.03)	1.60 (.03)	1.59 (.03)	1.65 (.03)	1.66 (.03)	1.67 (.03)	1.78 (.03)
Top 10%	1.76 (.05)	1.77 (.05)	1.79 (.06)	1.77 (.06)	1.82 (.05)	1.83 (.06)	1.85 (.06)	1.94 (.06)
Top 5%	1.84 (.09)	1.89 (.09)	1.91 (.09)	1.84 (.09)	1.93 (.09)	1.94 (.09)	1.94 (.09)	2.00 (.09)
Top 1%	1.99 (.22)	2.03 (.22)	2.06 (.22)	1.97 (.20)	2.04 (.20)	2.04 (.20)	2.06 (.21)	2.09 (.21)
Top 0.1%	3.61 (1.22)	2.73 (1.05)	2.73 (1.05)	3.42 (2.23)	3.33 (2.04)	3.32 (2.04)	3.37 (2.02)	3.38 (2.07)
Parametric Gini among 'top' incomes								
Top 25%	44.48	45.78	45.37	45.92	43.53	43.14	42.70	39.17
Top 10%	39.76	39.37	38.70	39.28	37.94	37.65	37.02	34.73
Top 5%	37.22	35.85	35.35	37.40	34.95	34.74	34.86	33.44
Top 1%	33.65	33.44	32.95	33.97	32.72	32.80	32.20	31.56
Top 0.1%	19.21	20.80	20.84	16.53	16.60	16.63	16.73	16.47
Income share (%) among incomes classified as 'top'								
Top 25%	65.57	66.24	65.03	67.96	64.59	63.31	62.90	58.64
Top 10%	43.72	43.90	42.85	44.84	42.40	41.45	40.93	37.36
Top 5%	31.37	31.13	30.40	32.40	30.00	29.27	29.03	26.29
Top 1%	14.09	13.60	13.32	14.77	12.97	12.67	12.46	11.22
Top 0.1%	3.33	3.25	3.17	3.16	2.64	2.57	2.52	2.26
Semiparametric Gini (combining nonparametric 'bottom' & parametric 'top' Gini)								
Top 25%	55.86 (1.37)	56.64 (1.46)	54.77 (1.51)	58.77 (2.18)	54.56 (1.96)	52.62 (2.06)	52.10 (2.10)	46.42 (1.37)
Top 10%	54.10 (1.22)	54.25 (1.22)	52.32 (1.14)	55.78 (1.30)	52.60 (1.27)	50.67 (1.33)	50.11 (1.31)	44.92 (1.30)
Top 5%	53.47 (1.03)	53.45 (1.00)	51.58 (1.00)	55.15 (1.28)	51.95 (1.17)	50.04 (1.08)	49.65 (1.17)	44.65 (1.17)
Top 1%	53.03 (1.00)	53.08 (0.90)	51.24 (0.94)	54.60 (1.27)	51.60 (1.02)	49.71 (1.14)	49.24 (1.00)	44.37 (1.01)
Top 0.1%	52.71 (0.76)	52.78 (0.72)	50.97 (0.74)	54.11 (0.77)	51.21 (0.71)	49.31 (0.71)	48.88 (0.70)	44.05 (0.69)
Uncorrected	52.75 (0.82)	52.79 (0.79)	50.99 (0.80)	54.31 (0.86)	51.33 (0.76)	49.43 (0.77)	49.00 (0.77)	44.17 (0.76)
Gini correction								
Top 25%	+3.11	+3.85	+3.78	+4.46	+3.23	+3.19	+3.10	+2.25
Top 10%	+1.35	+1.46	+1.33	+1.47	+1.27	+1.24	+1.11	+0.75
Top 5%	+0.72	+0.66	+0.59	+0.84	+0.62	+0.61	+0.65	+0.48
Top 1%	+0.28	+0.29	+0.25	+0.29	+0.27	+0.28	+0.24	+0.20
Top 0.1%	–0.04	–0.01	–0.02	–0.20	–0.12	–0.12	–0.12	–0.12

Notes: Nonparametric (semiparametric) Gini standard errors are jack-knife (bootstrap) estimates. Ginis and standard errors are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2014, CEQ database.

Table 5. Income summary statistics: Replacing top net market incomes with Pareto I estimates, and imputing other income concepts by CEQ Method

	(1) Market income per cap.	(2) Market inc + pensions per cap.	(3) Gross income per cap.	(4) Taxable income per cap.	(5) Net market income per cap.	(6) Disposable income per cap.	(7) Consumable income per cap.	(8) Final income per cap.
99.9th %ile	1,586,003	1,585,680	1,585,680	1,261,436	1,377,565	1,377,565	1,320,081	1,322,184
99th %ile	317,481	322,046	322,046	232,907	286,984	286,984	277,402	281,135
95th %ile	123,626	130,824	131,988	87,958	117,226	118,115	113,030	117,428
90th %ile	83,668	87,845	88,601	60,407	80,362	80,680	78,196	82,721
75th %ile	46,473	48,042	48,594	32,975	44,590	45,144	43,769	48,405
Mean	44,321	46,055	47,127	31,397	42,187	43,260	41,848	46,454
Median	25,983	26,941	27,747	18,031	25,443	26,286	25,659	30,340
25th %ile	14,675	15,298	16,642	9,237	14,754	16,045	15,782	20,377
10th %ile	8,113	8,454	10,353	2,833	8,354	10,226	10,112	14,357
5th %ile	4,820	5,006	7,323	525	4,950	7,255	7,207	11,480
1st %ile	1,751	1,796	3,944	0	1,796	3,903	3,917	7,098
Std. dev.	98,330	99,735	99,571	80,198	89,632	89,485	85,228	85,071
Skewness	17.76	17.27	17.34	22.34	18.72	18.80	18.60	18.63
Kurtosis	505.22	482.50	485.23	817.68	591.49	594.94	582.52	584.64
Sample	73,508	73,508	73,508	73,508	73,508	73,508	73,508	73,508
Top 0.1% inc. share	5.00	4.85	4.74	5.85	4.72	4.61	4.53	4.09
0.1-1% inc. share	11.10	10.86	10.62	11.87	10.58	10.33	10.19	9.23
1-5% inc. share	16.19	16.45	16.12	16.36	15.96	15.63	15.50	14.29
5-10% inc. share	11.34	11.42	11.24	11.50	11.23	11.02	10.99	10.40
Gini (HH-size & sampling weighted data)	54.00	54.00	52.21	55.94	52.68	50.80	50.35	45.50
	(1.04)	(1.00)	(1.02)	(1.13)	(0.99)	(1.01)	(1.00)	(0.99)
Mean log dev.(GE0)	0.541	0.541	0.481	0.630	0.511	0.452	0.443	0.351
	(0.012)	(0.012)	(0.011)	(0.014)	(0.011)	(0.011)	(0.011)	(0.009)
Theil index (GE1)	0.672	0.664	0.629	0.737	0.634	0.598	0.586	0.490
	(0.028)	(0.027)	(0.027)	(0.033)	(0.026)	(0.026)	(0.025)	(0.023)
Coef. of var. (GE2)	2.461	2.345	2.232	3.262	2.257	2.139	2.074	1.677
	(0.253)	(0.240)	(0.230)	(0.396)	(0.247)	(0.236)	(0.227)	(0.187)

Notes: Statistics are based on non-response correction weights estimated in the logarithmic model of market income (model 1). These statistics exclude 19 household (41 individual) observations with market income of 0. The statistics are still comparable to those in table 1, which are extremely robust to this exclusion (changing by 0.01 at most.) Another 1,446 household (2,889 individual) 0-income observations are omitted in computations of the Gini for taxable income. Gini standard errors are jack-knife estimates on household-level data (recognizing that household-member incomes are copies of one another), accounting for household size except in last row. Ginis and standard errors are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2014, CEQ database.

Table 6. Summary results of correction methods: corrected Ginis for all income concepts

	(1) Market income per cap.	(2) Market inc + pensions per cap.	(3) Gross income per cap.	(4) Taxable income per cap.	(5) Net market income per cap.	(6) Disposable income per cap.	(7) Consumable income per cap.	(8) Final income per cap.
Uncorrected	52.8	52.8	51.0	54.3	51.3	49.4	49.0	44.2
Correction by reweighting (models in table 2)								
Minimum	53.7 **+0.9**	53.7 **+0.9**	51.9 **+0.9**	55.2 **+0.9**	52.2 **+0.9**	50.4 **+1.0**	49.9 **+0.9**	45.2 **+1.0**
Mean	57.2 **+4.4**	57.1 **+4.3**	55.4 **+4.4**	59.6 **+5.2**	55.9 **+4.5**	54.1 **+4.7**	53.6 **+4.6**	48.9 **+4.7**
Median	57.0 **+4.3**	56.9 **+4.1**	55.3 **+4.3**	59.1 **+4.8**	55.6 **+4.2**	53.8 **+4.4**	53.4 **+4.4**	48.6 **+4.4**
Maximum	60.9 **+8.2**	60.7 **+7.9**	59.1 **+8.1**	64.4 **+10.1**	60.0 **+8.6**	58.2 **+8.8**	57.7 **+8.7**	53.0 **+8.8**
Correction by Pareto (type I) replacing of own income concept								
Minimum	52.7 **+0.0**	52.8 **+0.0**	51.0 **+0.0**	54.1 **–0.2**	51.2 **–0.1**	49.3 **–0.1**	48.9 **–0.1**	44.1 **–0.1**
Mean	53.8 **+1.1**	54.0 **+1.3**	52.2 **+1.2**	55.7 **+1.4**	52.4 **+1.1**	50.5 **+1.0**	50.0 **+1.0**	44.9 **+0.7**
Median	53.5 **+0.7**	53.5 **+0.7**	51.6 **+0.6**	55.2 **+0.8**	52.0 **+0.6**	50.0 **+0.6**	49.7 **+0.6**	44.7 **+0.5**
Maximum	55.9 **+3.1**	56.6 **+3.9**	54.8 **+3.8**	58.8 **+4.5**	54.6 **+3.2**	52.6 **+3.2**	52.1 **+3.1**	46.4 **+2.3**
Correction by Pareto (type I) replacing of market income + CEQ Method								
Minimum	52.6 **–0.1**	52.7 **–0.1**	50.9 **+0.0**	54.1 **–0.2**	51.2 **–0.1**	49.3 **–0.1**	48.9 **–0.1**	44.0 **–0.1**
Mean	53.8 **+1.1**	53.8 **+1.0**	52.0 **+1.0**	55.7 **+1.4**	52.5 **+1.1**	50.6 **+1.2**	50.1 **+1.1**	45.3 **+1.1**
Median	53.4 **+0.7**	53.5 **+0.7**	51.7 **+0.7**	55.2 **+0.9**	52.1 **+0.8**	50.2 **+0.8**	49.8 **+0.8**	44.9 **+0.7**
Maximum	55.9 **+3.1**	55.7 **+2.9**	53.9 **+3.0**	58.4 **+4.1**	54.6 **+3.3**	52.7 **+3.3**	52.3 **+3.3**	47.4 **+3.2**

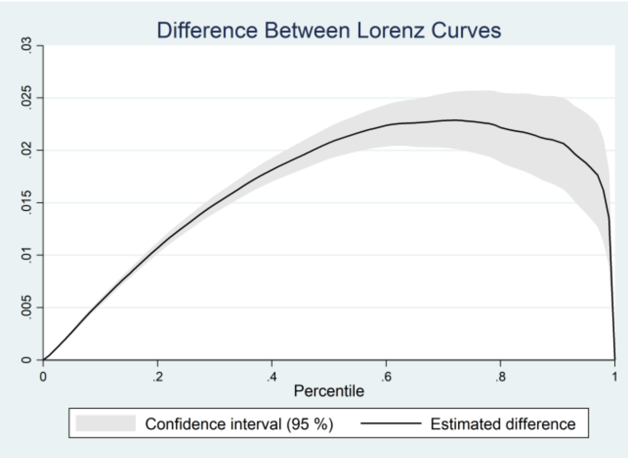
Notes: Pc.pt. differences from uncorrected Ginis in **bold**. These Ginis and differences in them arise from 'Gini (HH-size & sampling weighted data)' in tables 1 & 3, and 'Semiparametric Gini' in table 4. Ginis and percentage point changes are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2014, CEQ database.

Table 7. Redistributive impacts of fiscal tools: high/center/low estimates of impacts on inequality

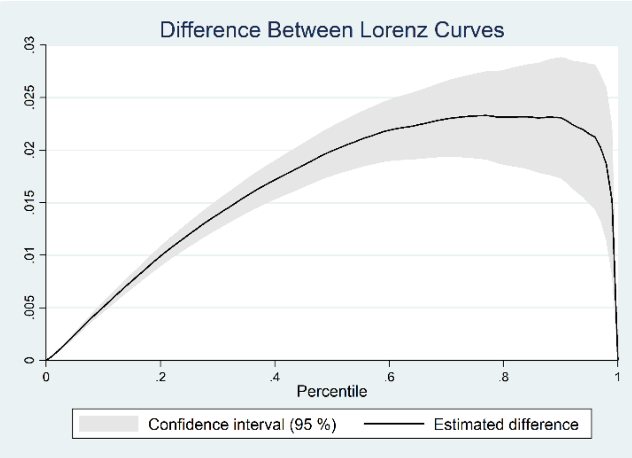
	Market income inequality	+ Net contribut. pensions Market → Market+Pensions	+ Cash-like transfers Market+Pensions → Gross	+ Nontaxable income Taxable → Gross	– Direct taxes Gross → Disposable[i]	– Indirect taxes & subsidies Disposable → Consumable	+ Net in-kind programs Consumable → Final	Final income inequality
Gini coefficient: pc.pt. change								
Uncorrected	52.75	+0.04	–1.80	–3.32	–1.56	–0.43	–4.83	44.17
Income distrib. corrected for nonresponse by reweighting								
High	60.17	+0.00	–1.74	–3.32	–1.58	–0.44	–4.78	51.89
Center	56.41	–0.12	–1.67	–3.69	–1.52	–0.45	–4.76	47.89
Low	53.78	–0.24	–1.59	–4.42	–1.27	–0.47	–4.70	45.24
Each income concept corrected for top income mismeasurement by replacing								
High	55.86	+0.05	–1.84	–3.36	–1.53	–0.47	–4.87	46.42
Center	54.10	+0.15	–1.93	–3.46	–1.65	–0.56	–5.19	44.92
Low	53.03	+0.78	–1.87	–4.00	–2.15	–0.52	–5.68	44.37
Market income corrected for top income biases by replacing + CEQ Method								
High	55.85	+0.00	–1.79	–3.73	–1.41	–0.45	–4.85	47.36
Center	54.00	+0.04	–1.80	–3.42	–1.51	–0.44	–4.84	45.50
Low	53.15	–0.14	–1.77	–4.50	–1.21	–0.48	–4.89	44.61
Top 10 percent income share: pc.pt. change								
Uncorrected	42.10	+0.04	–0.85	–2.28	–1.29	–0.35	–3.13	36.52
Income distrib. corrected for nonresponse by reweighting								
High	51.12	–0.03	–0.83	–2.39	–1.33	–0.37	–3.12	45.08
Center	46.48	–0.17	–0.84	–3.00	–1.29	–0.40	–3.24	40.54
Low	43.30	–0.33	–0.84	–4.08	–1.05	–0.42	–3.40	37.62
Each income concept corrected for top income mismeasurement by replacing								
High	46.57	+0.28	–0.92	–2.19	–0.85	–0.48	–2.61	39.80
Center	43.72	+0.18	–1.05	–1.99	–1.40	–0.52	–3.57	37.36
Low	40.28	+1.22	–0.94	–2.03	–2.18	–0.43	–4.44	35.70
Market income corrected for top income biases by replacing + CEQ Method								
High	46.17	+0.04	–0.86	–2.41	–1.23	–0.37	–3.15	40.48
Center	43.63	–0.05	–0.86	–2.86	–1.13	–0.38	–3.20	38.01
Low	42.60	–0.13	–0.88	–3.80	–0.89	–0.43	–3.36	37.03
Top 1 percent income share: pc.pt. change								
Uncorrected	13.59	–0.38	–0.30	–1.43	–0.69	–0.15	–1.17	10.90
Income distrib. corrected for nonresponse by reweighting								
High	23.70	–0.39	–0.30	–1.56	–0.73	–0.18	–1.20	20.19
Center	18.01	–0.49	–0.35	–2.50	–0.67	–0.22	–1.45	14.83
Low	14.56	–0.65	–0.42	–4.27	–0.34	–0.28	–1.82	11.76
Each income concept corrected for top income mismeasurement by replacing								
High	17.48	–0.49	–0.28	–1.45	–0.65	–0.21	–1.24	13.08
Center	15.89	–0.60	–0.37	–1.46	–0.98	–0.25	–1.75	11.95
Low	14.09	–0.05	–0.33	–1.64	–1.50	–0.22	–2.30	11.22
Market income corrected for top income biases by replacing + CEQ Method								
High	20.42	–0.39	–0.35	–2.36	–0.42	–0.22	–1.40	17.37
Center	16.10	–0.36	–0.32	–1.64	–0.61	–0.17	–1.23	13.32
Low	14.31	–0.47	–0.43	–4.04	–0.05	–0.31	–1.79	11.62

[i] Alternatively, this can be obtained as 'market+pensions → net market' for estimates within 0.1 pc.pt. of those above.
Notes: These Ginis are comparable to 'Gini (HH-size & sampling weighted data)' in tables 1 & 3, and 'Semiparametric Gini' in table 4. Ginis and percentage point changes are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2014, CEQ database.

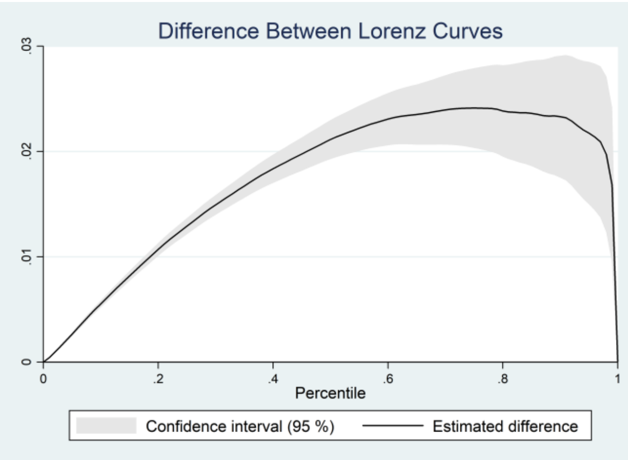
Figure 2. Lorenz curve: market vs. disposable income per capita, uncorrected versus corrected income distributions



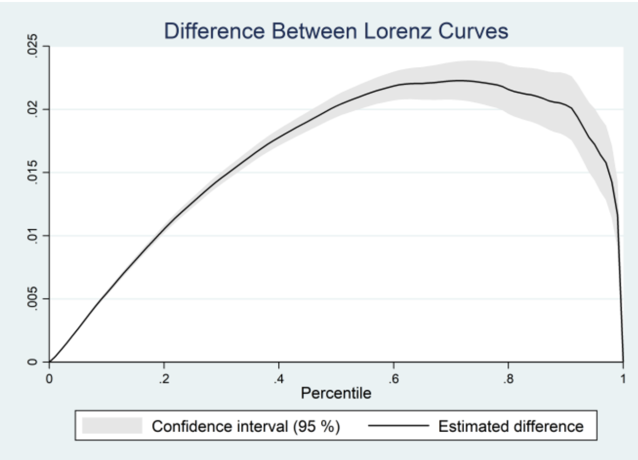
i. ENIGH sampling weights uncorrected for nonresponse



ii. Weights corrected for unit nonresponse (model 1)



iii. Top 10% of incomes of each income concept replaced with Pareto (I) values



iv. Top 10% of net market incomes replaced with Pareto (I) values, other income concepts imputed by CEQ Method

Notes: Positive values indicate that disposable income Lorenz curve dominates, and shows less inequality than market income Lorenz curve. Distributions account for sampling weights and household size.
Source: Own analysis of ENIGH 2014, CEQ database.

Appendix 1. Additional Summary Statistics and Estimation Results for ENIGH 2014

ENIGH 2014 covers 73,508 members of 19,479 households (domestic servants and guests excluded). The 19,479 fully interviewed households represent 88.2 percent of the 21,786 households (21,427 distinct dwellings) selected for participation under the sampling design. Another 838 households could not be fully interviewed even though they were contacted (type A nonresponse), 1,312 households could not be contacted because they were listed in unoccupied dwellings (type B), and 157 households were listed in what turned out to be invalid dwellings (type C).[16] INEGI does not conduct proxy interviews in place of units failing to respond to ENIGH, and does not perform imputation or top-coding of incomes.

The dataset available from the CEQ Institute includes information on household demographics and residence; level of current school attendance; employment and contract status; all core income concepts; various income sources, transfers and contributions; national poverty lines; indicators for participation in various transfer programs; sampling stratum and primary sampling unit (PSU) identifiers; selected household durable assets and utilities, financial asset indicators; and sampling weights.

Sampling weights correct for sampling bias and unit non-response bias at the level of PSUs, and inflate to the national population of 119.9 million. Item-nonresponse in the ENIGH was mitigated by telephone re-interviewing of households during the editing process. Incomes were checked for potential data entry errors.

ENIGH 2014 survey sample was obtained from a stratified multi-phase sample with 2,626 PSUs, that is, blocks of dwellings selected from state geostatistical areas stratified according to four geographic and socio-economic criteria (200 strata). Sampling frame was based on demographic and geographic information from the 2010 National Census (Censo de Población y Vivienda). Households were interviewed between August and November of the survey year.

Table A1-1. Population and sample sizes, non-response rates, and mean incomes by state

State code	State	Represented population	Strata	PSUs	Fully interviewed households (individuals)	Type A non-responding HHs (%)	Mean market income per cap.	Mean gross income per cap.	Mean disposable income per cap.	Mean final income per cap.
01	Aguascalientes	1,273,449	5	90	553 (2,169)	30 (5.1%)	46,163	49,773	45,660	48,793
02	Baja California	3,445,408	5	96	547 (1,880)	4 (0.7%)	57,134	59,291	53,989	56,879
03	Baja California Sur	747,857	7	63	543 (1,916)	5 (0.9%)	54,577	57,688	52,122	56,378
04	Campeche	897,279	7	75	539 (2,100)	34 (5.9%)	46,698	49,368	45,192	48,956
05	Coahuila de Zarag.	2,940,002	5	96	515 (1,938)	47 (8.4%)	58,962	61,824	56,936	59,651
06	Colima	712,419	5	81	558 (1,987)	14 (2.4%)	50,086	53,440	49,476	52,661
07	Chiapas	5,182,656	6	48	565 (2,480)	12 (2.1%)	18,293	20,817	19,641	23,550
08	Chihuahua	3,681,549	7	93	548 (1,955)	18 (3.2%)	44,455	46,791	42,959	46,061
09	Distrito Federal	8,799,243	4	154	685 (2,423)	53 (7.2%)	81,471	86,106	77,703	79,085
10	Durango	1,752,061	6	78	549 (2,085)	24 (4.2%)	34,153	37,135	34,518	38,551
11	Guanajuato	5,783,856	6	71	548 (2,192)	30 (5.2%)	35,197	37,270	34,118	37,603
12	Guerrero	3,553,259	6	60	496 (2,042)	49 (9.0%)	25,258	28,055	26,554	30,612
13	Hidalgo	2,850,418	6	48	544 (2,004)	17 (3.0%)	32,455	35,382	33,268	37,190
14	Jalisco	7,855,551	6	92	601 (2,342)	60 (9.1%)	49,736	51,559	47,083	49,407
15	México	16,680,605	7	130	713 (2,878)	38 (5.1%)	42,718	44,711	39,859	43,012
16	Mich. de Ocampo	4,573,656	7	51	602 (2,361)	4 (0.7%)	29,104	31,866	29,983	33,632
17	Morelos	1,901,086	7	75	528 (1,873)	28 (5.0%)	39,324	43,812	41,834	45,010
18	Nayarit	1,207,885	7	63	550 (1,978)	14 (2.5%)	41,608	44,730	41,580	45,409
19	Nuevo León	5,024,096	6	116	616 (2,239)	45 (6.8%)	66,530	69,825	62,289	64,970

[16] Source: Instituto Nacional de Estadística y Geografía (2016), Encuesta Nacional de Ingresos y Gastos de los Hogares ENIGH 2014: Operative de Campo, México: INEGI.

20	Oaxaca	3,988,647	6	45	569 (2,306)	13 (2.2%)	20,430	23,733	22,662	26,656
21	Puebla	6,137,672	8	77	641 (2,592)	34 (5.0%)	30,072	31,955	30,129	33,717
22	Querétaro	1,980,758	6	75	573 (2,208)	5 (0.9%)	50,522	52,967	48,065	50,479
23	Quintana Roo	1,540,017	6	92	517 (1,922)	25 (4.6%)	51,554	52,882	47,841	51,202
24	San Luis Potosí	2,728,705	6	69	517 (2,025)	53 (9.3%)	34,697	37,769	35,244	39,613
25	Sinaloa	2,960,736	6	78	585 (2,114)	1 (0.2%)	48,545	52,576	48,319	52,940
26	Sonora	2,902,867	8	84	547 (1,930)	14 (2.5%)	54,927	58,384	53,468	57,705
27	Tabasco	2,359,988	8	152	1,885 (7,069)	47 (2.4%)	37,257	39,919	37,282	41,197
28	Tamaulipas	3,511,188	5	93	536 (1,885)	31 (5.5%)	49,401	53,077	48,380	51,934
29	Tlaxcala	1,266,058	5	78	533 (2,228)	26 (4.7%)	27,577	29,600	28,030	32,308
30	Veracruz de Ignacio	8,004,171	7	74	679 (2,417)	8 (1.2%)	29,721	32,790	30,841	34,110
31	Yucatán	2,096,010	7	78	539 (2,067)	36 (6.3%)	40,605	44,357	41,072	44,234
32	Zacatecas	1,567,160	7	51	558 (1,903)	19 (3.3%)	32,298	35,994	34,220	37,823
	Nationwide	119,906,312	200	2,626	19,479 (73,508)	838 (4.1%)	43,178	45,985	42,117	45,374

Note: Calculation of nonresponse rates omits type B and C nonresponses. Mean incomes account for sampling weights but are computed only among responding households, and may not be representative of underlying population.
Source: Instituto Nacional de Estadística y Geografía (2016), Encuesta Nacional de Ingresos y Gastos de los Hogares ENIGH 2014: Operative de Campo, México: INEGI; own analysis of ENIGH 2014, CEQ database.

Table A1-2. Estimation results for selected multivariate models of response probability

Specification of *g(x)*	$\widehat{\theta_0}$ (s.e.)	$\widehat{\theta_1}$ (s.e.)	$\widehat{\theta_2}$ (s.e.)	$\widehat{\theta_3}$ (s.e.)	$\widehat{\theta_4}$ (s.e.)	Sum of squared wghted. errors	Factor of propor-tionality (σ^2)	AIC SIC	Gini (s.e.): Market inc. pc. Weighted data	HH-size wted, no sampling wghts	Equal HH weights= unwghtd
Uncorrected									52.75	50.63	52.84
									(0.82)	(0.45)	(0.49)
$\theta_0+\theta_1 10^{-6}$mkt. inc.	3.274	-1.196	48.914			37,501	0.449	232.12	61.68	57.94	59.07
$+\theta_2 10^{-15}$mkt. inc.^{2}	(0.003)	(0.006)	(0.370)					229.42	(5.91)	(4.80)	(3.62)
$\theta_0+\theta_1 10^{-6}$mkt. inc.pc	3.220	-2.888	0.341			39,284	0.463	233.61	60.16	57.13	60.50
$+\theta_2 10^{-12}$mkt. inc.pc^2	(0.003	(0.027)	(0.006)					230.91	(5.33)	(4.16)	(3.87)
$\theta_0+\theta_1$log(mkt. inc.)	133.821	-0.147	-150.00	-370.00	290.24	27,883	0.300	226.64	52.58	50.52	52.69
$+\theta_2$urban/10$+\theta_3$age$+\theta_4$age^2	(38·10^3)	(0.007)	(38·10^4)	(13.78)	(11.09)			224.69	(0.81)	(0.44)	(0.49)
$\theta_0+\theta_1$log(mkt. inc.)	11.896	-0.651	-0.295	-1.842	-2.262	34,261	0.431	233.23	54.05	51.56	53.76
$+\theta_2$sch.attend$+\theta_3$age$+\theta_4$age^2	(0.141)	(0.008)	(0.013)	(0.640)	(0.686)			231.28	(1.07)	(0.65)	(0.63)
$\theta_0+\theta_1$log(mkt. inc.)	24.037	-0.535	-13.963	-28.471		31,849	0.371	228.90	53.17	50.83	53.07
$+\theta_2$urban$+\theta_3$HH size	(14·10^3)	(0.010)	(14·10^3)	(0.428)				226.44	(0.99)	(0.58)	(0.56)
$\theta_0+\theta_1$log(mkt. inc.)	207.895	-1.088	-3.229	-7·10^3	6.3·10^3	24,833	0.321	222.93	51.88	49.88	52.31
$+\theta_2$urb$+\theta_3$hsize$+\theta_4$10hsize^2	(19·10^3)	(0.006)	(0.307)	(.7·10^6)	(.7·10^6)			220.98	(0.80)	(0.45)	(0.48)
$\theta_0+\theta_1$log(mkt. inc.)	16.995	-1.102	0.050	-23.743		33,022	0.426	230.05	57.09	53.64	55.46
$+\theta_2$male$+\theta_3$hsize	(0.102)	(0.007)	(0.042)	(0.456)				227.60	(2.60)	(1.94)	(1.52)
$\theta_0+\theta_1$log(mkt. inc.)	24.206	-0.427	0.316	-16.641		34,241	0.400	231.21	53.36	51.03	53.22
$+\theta_2$male$+\theta_3$urban	(12·10^4)	(0.009)	(0.041)	(12·10^4)				228.76	(0.94)	(0.53)	(0.54)
$\theta_0+\theta_1$log(mkt. inc.)	10.741	-0.641	-2.490			34,702	0.445	229.64	54.13	51.67	53.87
$+\theta_2$postsecondary.attend	(0.105)	(0.009)	(0.033)					226.94	(1.03)	(0.62)	(0.62)
$\theta_0+\theta_1$log(mkt. inc.)	3.264	-0.149	-1.218	-0.795		84,202	1.347	260.01	54.20	51.64	53.92
$+\theta_2$postsec.attend$+\theta_3$formal	(0.055)	(0.005)	(0.038)	(0.010)				257.55	(1.06)	(0.61)	(0.62)

Note: Standard errors on Gini coefficients are jackknife estimates. Variables are normalized: age=(years-12)/100; HH size=(#-1)/100. Measures of fit are not entirely comparable across models with different controls, because of different sample sizes.

Source: Own analysis of ENIGH 2014, CEQ database.

Table A1-3. Distribution of top incomes, various income concepts: Pareto (type I) estimates

	(1) Market income per cap.	(2) Market inc + pensions per cap.	(3) Gross income per cap.	(4) Taxable income per cap.	(5) Net market income per cap.	(6) Disposable income per cap.	(7) Consumable income per cap.	(8) Final income per cap.
Pareto (type I) coefficient α								
Top 25%	1.62 (.03)	1.59 (.03)	1.60 (.03)	1.59 (.03)	1.65 (.03)	1.66 (.03)	1.67 (.03)	1.78 (.03)
Top 10%	1.76 (.05)	1.77 (.05)	1.79 (.06)	1.77 (.06)	1.82 (.05)	1.83 (.06)	1.85 (.06)	1.94 (.06)
Top 5%	1.84 (.09)	1.89 (.09)	1.91 (.09)	1.84 (.09)	1.93 (.09)	1.94 (.09)	1.94 (.09)	2.00 (.09)
Top 1%	1.99 (.22)	2.03 (.22)	2.06 (.22)	1.97 (.20)	2.04 (.20)	2.04 (.20)	2.06 (.21)	2.09 (.21)
Top 0.1%	3.61 (1.22)	2.73 (1.05)	2.73 (1.05)	3.42 (2.23)	3.33 (2.04)	3.32 (2.04)	3.37 (2.02)	3.38 (2.07)
Inverted Pareto coefficient $\beta = \alpha/[\alpha - 1]$								
Top 25%	2.60	2.69	2.66	2.70	2.54	2.51	2.49	2.29
Top 10%	2.31	2.30	2.27	2.29	2.22	2.21	2.18	2.06
Top 5%	2.19	2.13	2.10	2.19	2.08	2.07	2.06	2.00
Top 1%	2.01	1.97	1.95	2.03	1.96	1.96	1.94	1.92
Top 0.1%	1.38	1.58	1.58	1.41	1.43	1.43	1.42	1.42
Top income share (%)								
Top 25%	65.57	66.24	65.03	67.96	64.59	63.31	62.9	58.64
Top 10%	43.72	43.9	42.85	44.84	42.4	41.45	40.93	37.36
Top 5%	31.37	31.13	30.4	32.4	30	29.27	29.03	26.29
Top 1%	14.09	13.6	13.32	14.77	12.97	12.67	12.46	11.22
Top 0.1%	3.33	3.25	3.17	3.16	2.64	2.57	2.52	2.26
Gini among top incomes (×100)								
Top 25%	44.48	45.78	45.37	45.92	43.53	43.14	42.70	39.17
Top 10%	39.76	39.37	38.70	39.28	37.94	37.65	37.02	34.73
Top 5%	37.22	35.85	35.35	37.40	34.95	34.74	34.86	33.44
Top 1%	33.65	33.44	32.95	33.97	32.72	32.80	32.20	31.56
Top 0.1%	19.21	20.80	20.84	16.53	16.60	16.63	16.73	16.47
Half coefficient of variation squared[i]								
Top 10%	--	--	--	--	--	--	--	--
Top 5%	--	--	--	--	--	--	--	--
Top 1%	--	3.66	2.14	--	2.95	3.28	1.81	1.24
Top 0.1%	0.024	0.092	0.093	0.030	0.034	0.034	0.032	0.032
Minimum income								
Top 25%	46,477	48,042	48,594	32,983	44,590	45,144	43,769	48,405
Top 10%	83,714	88,416	89,443	60,492	80,362	81,033	78,427	83,243
Top 5%	125,755	133,943	135,000	90,492	120,148	120,867	115,837	120,053
Top 1%	304,486	311,864	311,864	221,505	273,162	273,162	262,714	265,543
Top 0.1%	1,052,184	1,051,861	1,051,861	776,087	918,303	918,303	885,516	887,620
Mean income								
Top 25%	120,970	129,222	129,213	89,079	113,341	113,362	108,838	110,782
Top 10%	193,608	203,344	202,739	138,770	178,530	178,900	170,679	171,801
Top 5%	275,295	284,737	283,212	198,524	249,688	249,705	238,568	240,687
Top 1%	611,169	613,757	607,382	450,317	535,633	536,631	509,466	508,646
Top 0.1%	1,455,471	1,659,742	1,661,022	1,096,199	1,313,212	1,314,199	1,259,601	1,261,269

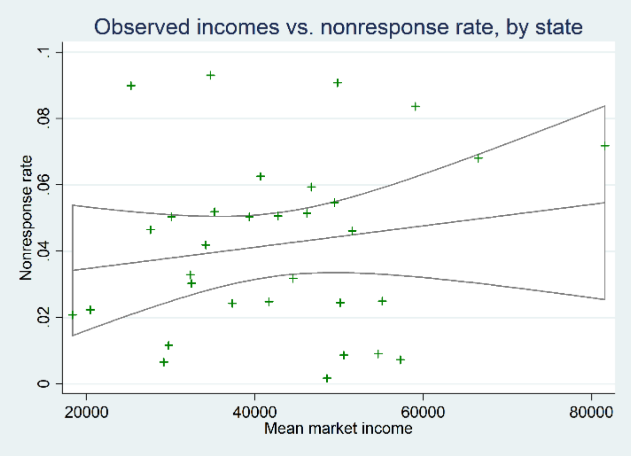
[i] Unable to calculate for other top income groups.
Source: Own analysis of ENIGH 2014, CEQ database.

Table A1-4. Generalized Pareto (type II) results, various income concepts (individual sampling-weighted sample)

	Market income per cap.	Market income + pensions per cap.	Gross income per cap.	Taxable income per cap.	Net market income per cap.	Disposable income per cap.	Consumable income per cap.	Final income per cap.
Pareto (type II) shape coefficient ξ								
Top 25%	0.50 (.02)	0.48 (.02)	0.48 (.02)	0.48 (.02)	0.48 (.02)	0.48 (.02)	0.48 (.02)	0.45 (.02)
Top 10%	0.48 (.03)	0.45 (.03)	0.46 (.03)	0.51 (.03)	0.45 (.03)	0.45 (.03)	0.45 (.03)	0.46 (.03)
Top 5%	0.48 (.04)	0.45 (.04)	0.46 (.04)	0.51 (.04)	0.44 (.04)	0.44 (.04)	0.44 (.04)	0.44 (.04)
Top 1%	0.52 (.08)	0.53 (.09)	0.55 (.09)	0.46 (.08)	0.41 (.08)	0.41 (.08)	0.43 (.08)	0.44 (.08)
Top 0.1%	0.78 (.43)	0.38 (.13)	0.38 (.13)	1.59 (.34)	1.11 (.34)	1.12 (.34)	1.04 (.26)	1.04 (.26)
Pareto (type II) scale coefficient, log(σ)								
Top 25%	10.37 (.02)	10.46 (.02)	10.46 (.02)	10.09 (.02)	10.33 (.02)	10.33 (.02)	10.29 (.02)	10.32 (.02)
Top 10%	10.85 (.03)	10.93 (.03)	10.92 (.03)	10.49 (.03)	10.80 (.03)	10.80 (.03)	10.74 (.03)	10.73 (.03)
Top 5%	11.19 (.04)	11.25 (.04)	11.23 (.04)	10.84 (.04)	11.12 (.04)	11.12 (.04)	11.07 (.04)	11.07 (.04)
Top 1%	11.92 (.13)	11.91 (.14)	11.87 (.14)	11.68 (.12)	11.88 (.12)	11.89 (.11)	11.81 (.12)	11.78 (.12)
Top 0.1%	12.01 (.54)	12.85 (.19)	12.85 (.20)	10.39 (.21)	11.49 (.39)	11.49 (.39)	11.53 (.25)	11.52 (.25)
1/ξ								
Top 25%	1.99	2.07	2.07	2.10	2.08	2.07	2.10	2.20
Top 10%	2.08	2.21	2.18	1.97	2.23	2.23	2.20	2.20
Top 5%	2.07	2.22	2.17	1.96	2.28	2.28	2.29	2.29
Top 1%	1.93	1.90	1.83	2.17	2.42	2.44	2.31	2.25
Top 0.1%	1.29	2.61	2.62	0.63	0.90	0.90	0.97	0.96
Inverted Pareto coefficient								
Top 25%	2.01	1.94	1.93	1.91	1.93	1.93	1.91	1.83
Top 10%	1.92	1.83	1.85	2.03	1.81	1.81	1.83	1.84
Top 5%	1.94	1.82	1.86	2.04	1.78	1.78	1.77	1.78
Top 1%	2.08	2.11	2.21	1.85	1.71	1.69	1.76	1.80
Top 0.1%	4.50	1.62	1.62	-1.70	-8.72	-8.56	-28.18	-25.81
Gini among top incomes								
Top 25%	38.76	38.45	38.25	38.19	37.50	37.34	36.91	34.56
Top 10%	35.75	34.59	34.56	36.67	33.79	33.66	33.54	32.43
Top 5%	34.79	32.94	33.01	35.88	32.03	31.89	31.74	31.14
Top 1%	34.17	33.93	34.57	32.27	29.87	29.81	30.19	30.26
Top 0.1%	33.82	22.78	22.78	-2.86	41.97	41.45	73.52	71.53
Log pseudo-likelihood ($LL/10^6$)								
Top 25%	-356.00	-358.00	-358.00	-347.00	-354.00	-354.00	-353.00	-353.00
Top 10%	-148.00	-148.00	-148.00	-144.00	-147.00	-147.00	-146.00	-146.00
Top 5%	-76.00	-76.00	-76.10	-74.00	-75.20	-75.30	-75.20	-74.90
Top 1%	-16.10	-16.20	-16.40	-15.70	-16.00	-16.00	-15.90	-16.00
Top 0.1%	-1.86	-1.52	-1.52	-1.36	-1.35	-1.35	-1.33	-1.33

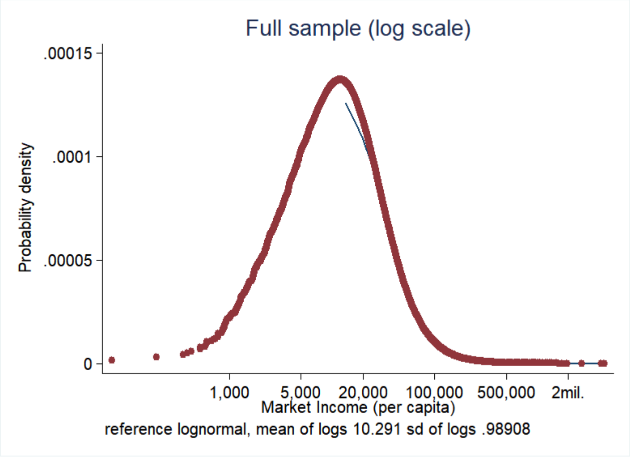
Source: Own analysis of ENIGH 2014, CEQ database.

Figure A1-1. Mean observed market income among respondents, and nonresponse rate, by state

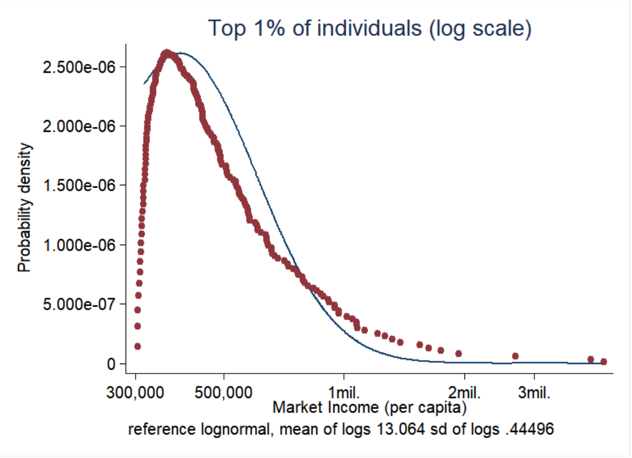


Notes: 95% confidence interval around linear fitted line is shown.
Source: Own analysis of ENIGH 2014, CEQ database.

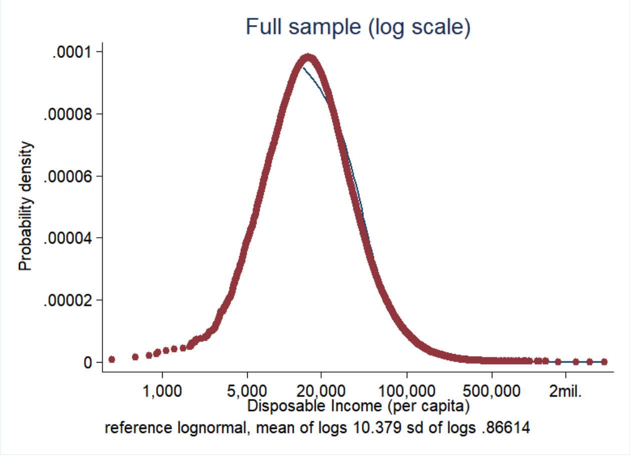
Figure A1-2. Density function of income per capita, with reference lognormal density



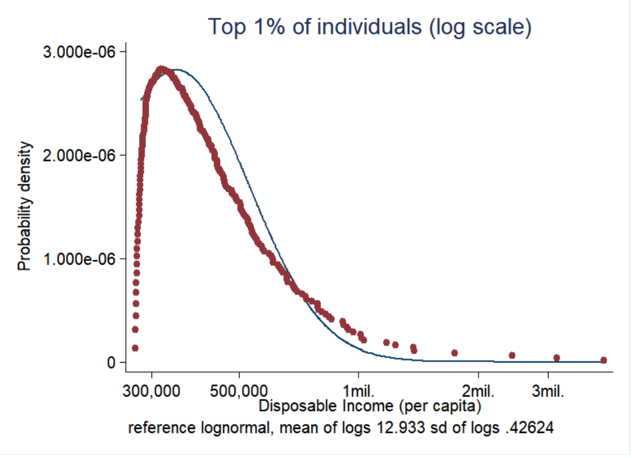
i. Market income, full sample



i. Market income, top 1%



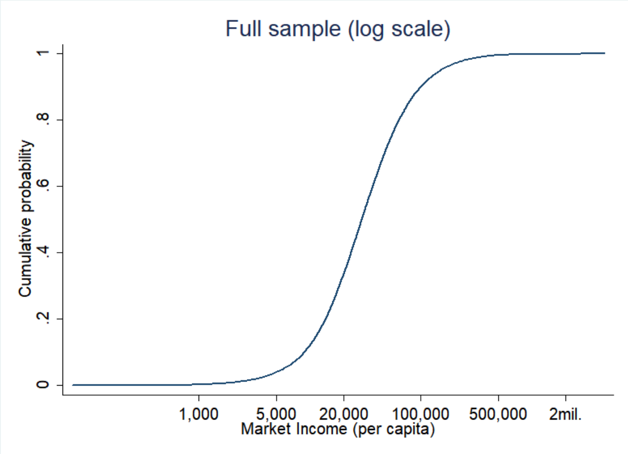
iii. Disposable income, full sample



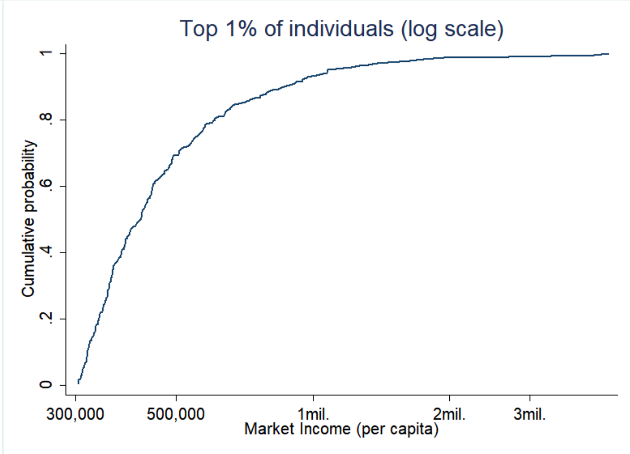
i. Disposable income, top 1%

Notes: Non-positive incomes omitted.
Source: Own analysis of ENIGH 2014, CEQ database.

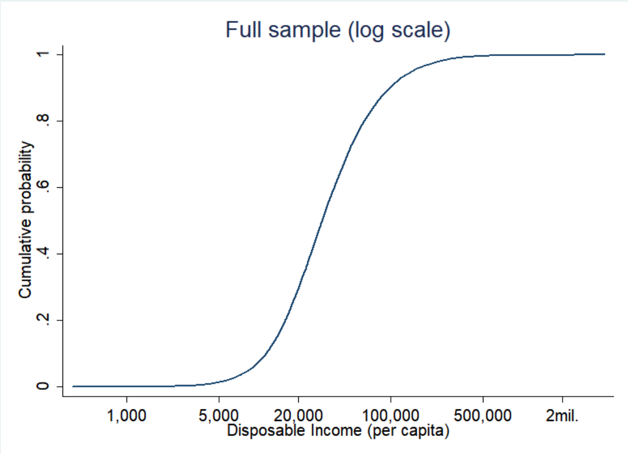
Figure A1-3. Cumulative density function of disposable income per capita (log scale)



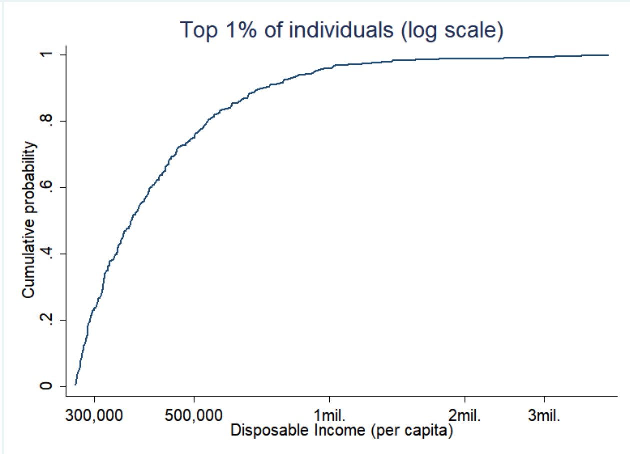
i. Market income, full sample



i. Market income, top 1%



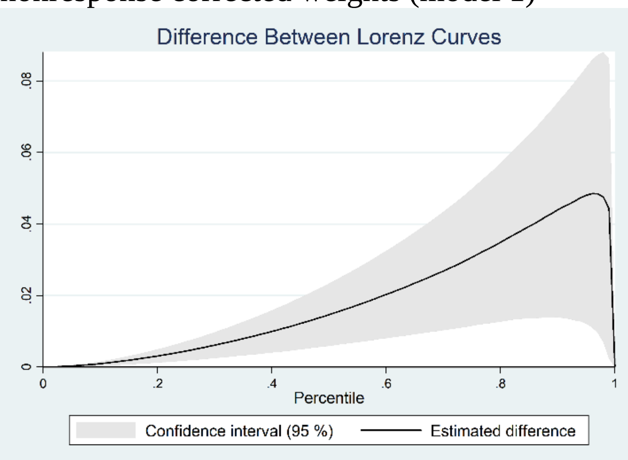
iii. Disposable income, full sample



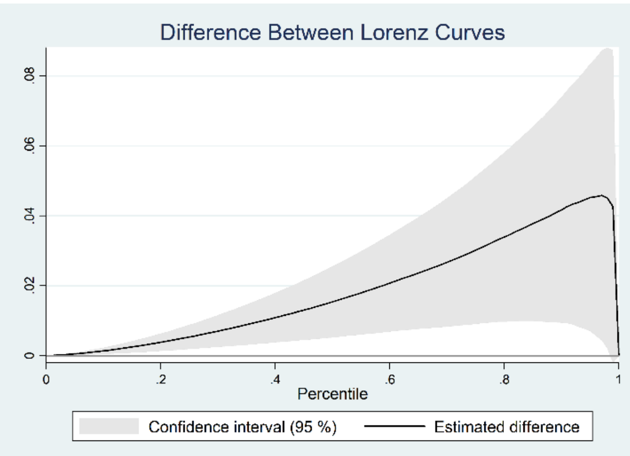
i. Disposable income, top 1%

Source: Own analysis of ENIGH 2014, CEQ database.

Figure A1-4. Lorenz curve: market and disposable income per capita, uncorrected versus unit-nonresponse corrected weights (model 1)



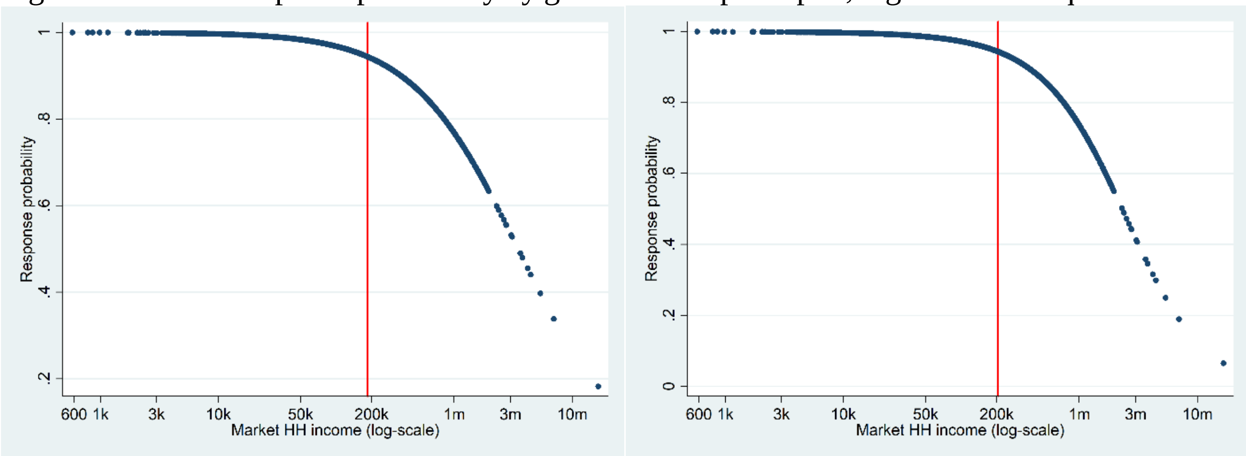
i. Market income per capita



i. Disposable income per capita

Notes: Positive values indicate that the Lorenz curve uncorrected for unit-nonresponse dominates, and shows less inequality than the corrected Lorenz curve. Distributions account for sampling weights and household size.

Figure A1-5. Unit response probability by gross income per capita, logarithmic vs. quadratic model

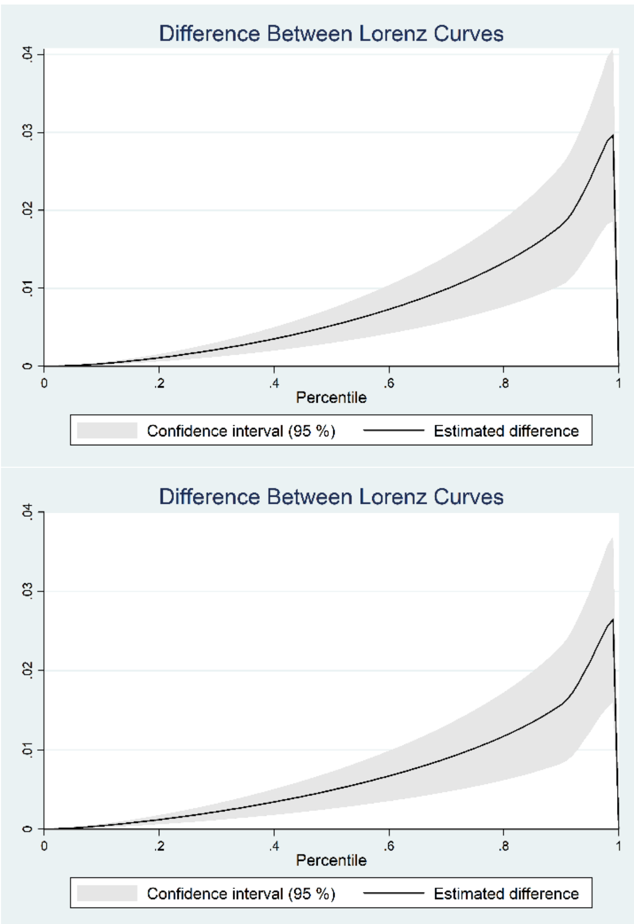


i. Logarithmic model of market income (model 1) $\theta_0+\theta_1\log(\text{market inc.})$

ii. Quadratic logarithmic model of market income (Model 7) $\theta_0+\theta_1\log(\text{market inc.})^2$

Notes: Red line shows mean market household income in the corrected income distribution: i) 184,505 and ii) 203,382.
Source: Own analysis of ENIGH 2014, CEQ database.

Figure A1-6. Lorenz curve: market and disposable income per capita, top 10% of incomes replaced with Pareto I values

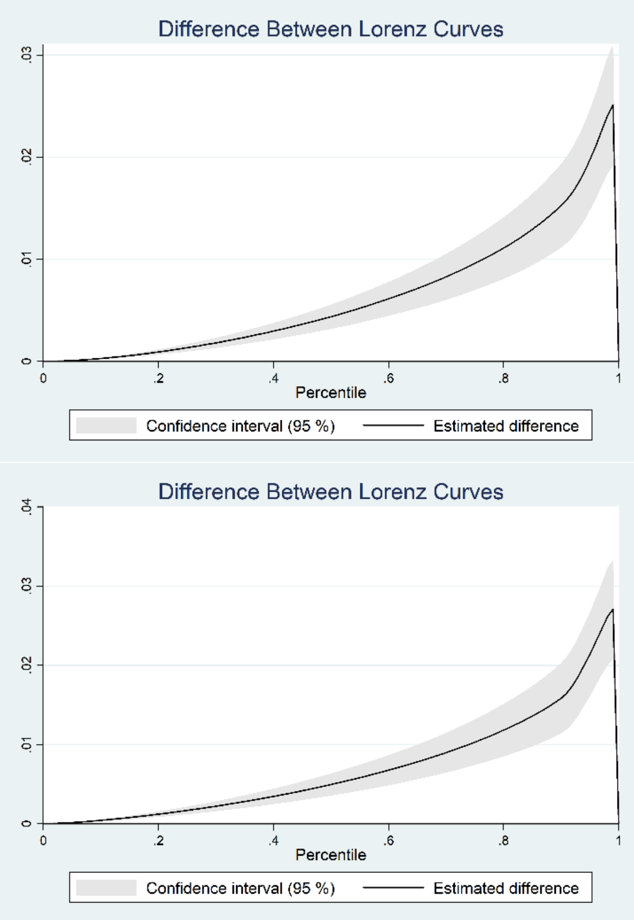


i. Market income per capita

i. Disposable income per capita

Notes: Pareto replacing is performed on own income concept. Positive values indicate that the uncorrected Lorenz curve dominates, and shows less inequality than the corrected Lorenz curve. Distributions account for sampling weights and household size.

Figure A1-7. Lorenz curve: market and disposable income per capita, top 10% of net market incomes replaced with Pareto I values, other income concepts imputed using CEQ Method

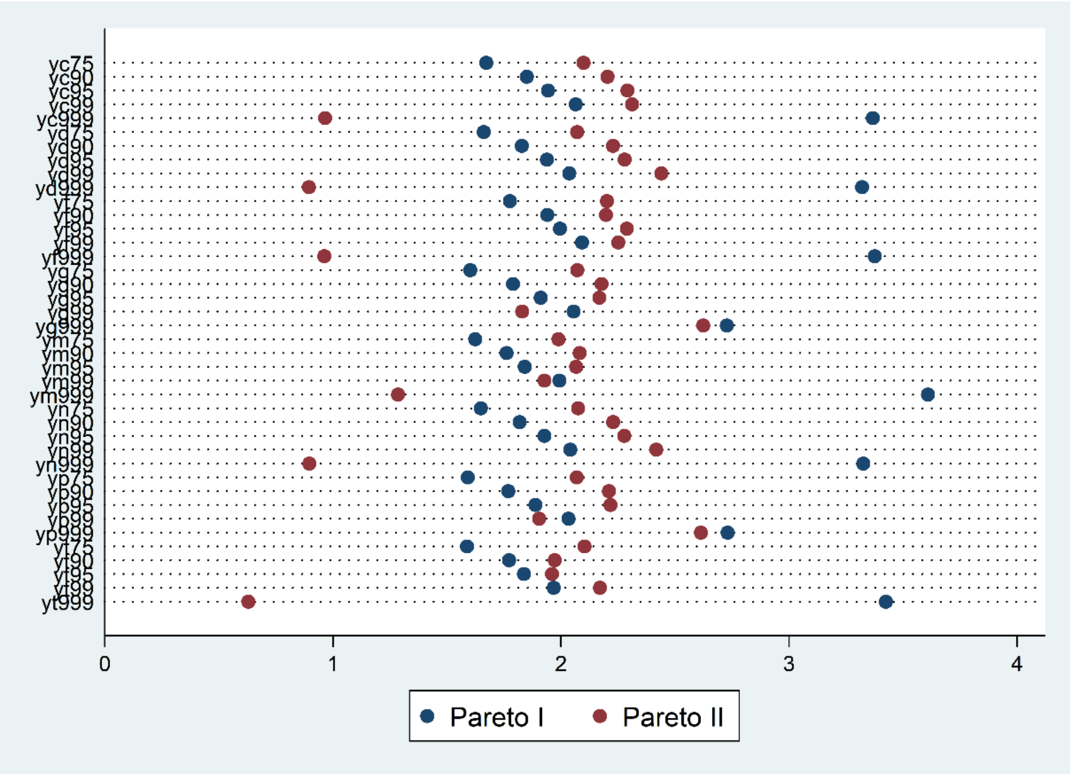


i. Market income per capita

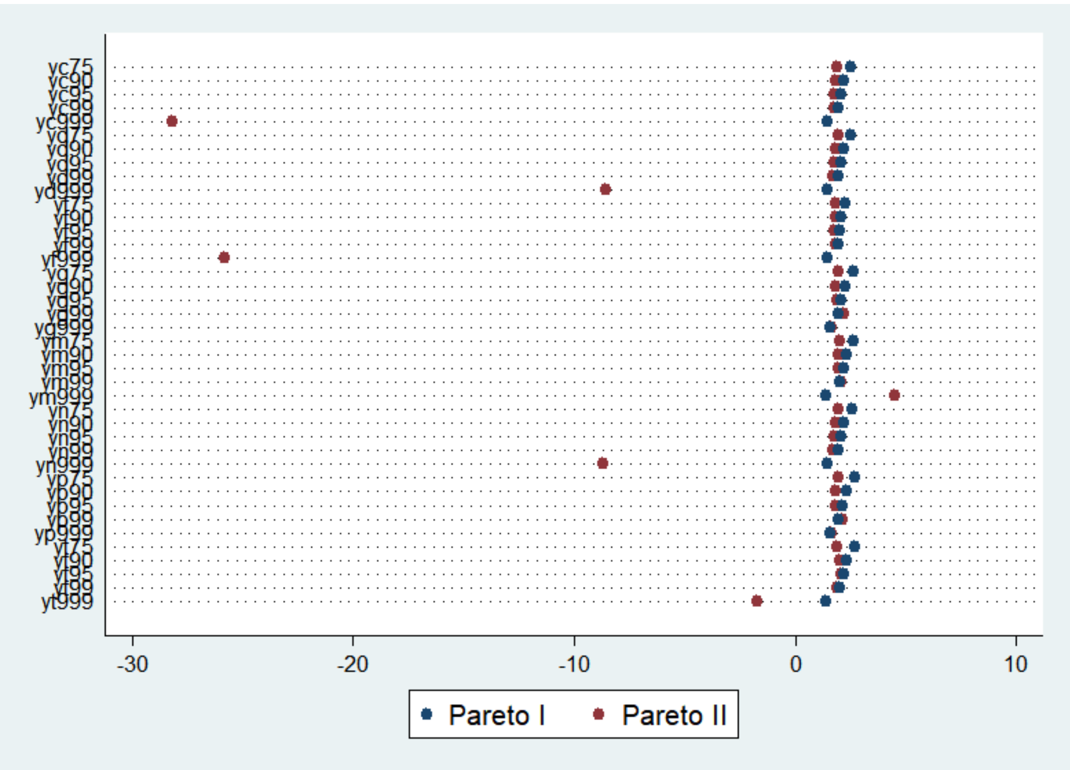
i. Disposable income per capita

Notes: Positive values indicate that the uncorrected Lorenz curve dominates, and shows less inequality than the corrected Lorenz curve. Distributions account for sampling weights and household size.

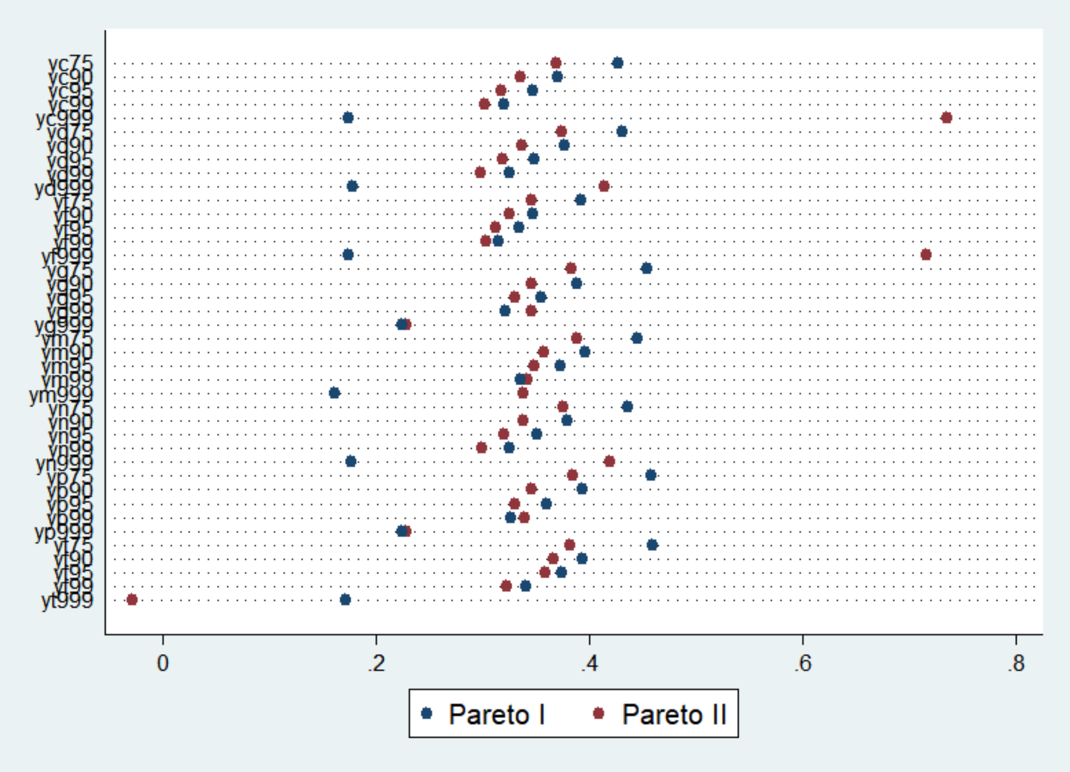
Figure A1-8. Comparison of Pareto I and Pareto II models, various income concepts and top income cutoffs



i. Pareto coefficient



ii. Inverted Pareto coefficient



iii. Gini coefficient

Notes: 'yp' is market income plus pensions, 'ym' market income, 'yt' taxable income, 'yg' gross income, 'yn' net market income, 'yd' disposable income, 'yc' consumable income, 'yf' final income.
Source: Own analysis of ENIGH 2014, CEQ database.

Appendix 2. Summary Statistics and Estimation Results for ENIGH 2012

Survey sample in ENIGH 2012 was obtained from a stratified multi-phase sample with 1,111 PSUs, that is, blocks of dwellings selected from state geostatistical areas stratified according to four geographic and socio-economic criteria (123 strata). Sampling frame was based on demographic and geographic information from the 2010 National Census (Censo de Población y Vivienda).[17] Item-nonresponse was mitigated by telephone re-interviewing households during the editing process. Incomes were checked for potential data entry errors. INEGI does not perform imputation or top-coding of incomes.

ENIGH 2012 covers 33,694 members of 9,002 households (plus 32 domestic servants and guests). The 9,002 fully interviewed households represent 88.2 percent of the 10,210 households (10,062 distinct dwellings) selected for participation under the sampling design. Another 252 households could not be fully interviewed even though they were contacted (type A nonresponse), 839 households could not be contacted because they were listed in unoccupied dwellings (type B), and 117 households were listed in what turned out to be invalid dwellings (type C).[18] INEGI does not use replacement units for those failing to respond to ENIGH. Sampling weights correct for sampling bias and unit non-response bias at the level of primary sampling units (PSUs), and inflate to the national population of 117.3 million (source: own analysis of ENIGH 2012). Tables A2-1 through A2-3 provide basic descriptive statistics of the Mexican ENIGH 2012 sample.

[17] Own analysis of ENIGH 2012, and Instituto Nacional de Estadística y Geografía (2013), Diseño muestral: Formación de las unidades primariasde muestreo para el levantamiento, Encuesta Nacional de Ingresos y Gastos de los Hogares ENIGH 2012, México: INEGI. http://catalog.ihsn.org/index.php/catalog/3971/download/54031 (accessed 8 August 2018).

[18] Source: Resultados Definitivos por Código de Resultados de Entrevista de Campo ENIGH – 2012.

Table A2-1a. Income summary statistics, various income concepts (sampling-weighted sample)

	Market income per cap.	Market inc +pensions per cap.	Gross income per cap.	Taxable income per cap.	Net market income per cap.	Disposable income per cap.	Consumable income per cap.	Final income per cap.
99.9th %ile	610,435	642,581	642,581	515,538	536,234	536,234	524,765	526,396
99th %ile	274,420	294,574	295,241	203,607	264,017	264,017	260,020	261,362
95th %ile	128,837	137,814	139,344	92,348	123,945	124,248	120,696	125,845
90th %ile	84,011	88,878	90,000	59,957	82,148	82,279	80,844	84,746
75th %ile	45,967	48,508	49,088	32,772	44,992	45,367	44,792	49,133
Mean	41,290	43,428	44,327	28,850	40,126	41,025	40,410	44,712
Median	25,249	26,253	27,012	17,648	25,196	25,768	25,706	29,934
25th %ile	13,882	14,432	15,741	8,317	14,148	15,367	15,512	19,866
10th %ile	7,384	7,775	9,419	2,316	7,723	9,295	9,358	13,902
5th %ile	4,567	4,765	6,804	522	4,754	6,781	6,843	11,010
1st %ile	1,935	2,003	3,609	0	1,987	3,609	3,717	7,257
Std. dev.	66,078	68,727	68,527	52,742	61,667	61,478	58,611	58,456
Skewness	14.15	13.16	13.25	20.73	15.26	15.38	14.23	14.25
Kurtosis	471.08	411.09	415.24	979.16	575.96	582.24	506.46	508.50
Sample	33,694	33,694	33,694	32,296	33,694	33,694	33,694	33,694
Gini (HH-size & sampling weighted)	52.09	52.39	50.77	54.32	50.98	49.29	48.63	43.99
	(0.76)	(0.72)	(0.73)	(0.81)	(0.69)	(0.70)	(0.68)	(0.67)
Top 0.1% inc. share	2.89	2.85	2.79	3.38	2.71	2.65	2.54	2.30
0.1-1% inc. share	8.44	8.36	8.20	8.96	7.83	7.67	7.48	6.82
1-5% inc. share	17.05	17.27	16.96	17.66	16.94	16.62	16.44	15.18
5-10% inc. share	12.51	12.64	12.45	12.70	12.40	12.20	12.13	11.45
Mean log dev. (GE0)	0.502	0.509	0.453	0.606	0.479	0.424	0.412	0.323
	(0.010)	(0.010)	(0.009)	(0.012)	(0.009)	(0.009)	(0.008)	(0.007)
Theil index (GE1)	0.544	0.547	0.518	0.598	0.515	0.485	0.470	0.390
	(0.022)	(0.020)	(0.020)	(0.029)	(0.021)	(0.020)	(0.019)	(0.017)
Half coef. of var. squared (GE2)	1.281	1.252	1.195	1.671	1.181	1.123	1.052	0.855
	(0.216)	(0.195)	(0.188)	(0.419)	(0.224)	(0.214)	(0.186)	(0.153)

Notes: MXN0 incomes (3 household observations for market income, 714 for taxable income) are omitted in computations of the Gini. Gini standard errors are jack-knife estimates on household-level data (recognizing that household-member incomes are copies of one another), accounting for household size. Ginis and standard errors are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-1b. Income summary statistics, various income concepts (sampling-weighted sample)

	Market income, household	Taxable income, household	Gross income, household	Disposable income, household
99.9th %ile	1,106,022	1,066,304	1,106,022	1,106,022
99th %ile	357,572	254,035	378,584	324,101
95th %ile	161,906	116,378	177,564	160,459
90th %ile	107,522	74,361	118,779	108,861
75th %ile	55,715	37,738	59,881	56,430
Mean	51,771	34,830	56,358	52,298
Median	29,526	18,880	32,093	30,789
25th %ile	16,147	7,918	18,316	17,760
10th %ile	8,806	1,076	10,944	10,811
5th %ile	5,596	0	7,979	7,956
1st %ile	2,335	0	4,150	4,150
Std. dev.	95,550	80,035	98,428	89,995
Skewness	14.37	19.97	13.41	15.46
Kurtosis	371.22	668.12	331.30	443.55

Sample	8,999	8,288	9,002	9,002
Gini (HH-size & sampling weighted)	50.07	52.36	47.97	46.29
Top 0.1% inc. share	2.71	2.65	2.54	2.30
0.1-1% inc. share	7.83	7.67	7.48	6.82
1-5% inc. share	16.94	16.62	16.44	15.18
5-10% inc. share	12.40	12.20	12.13	11.45
Mean log dev.	0.465	0.594	0.404	0.374
(GE0)	(0.014)	(0.017)	(0.013)	(0.011)
Theil index (GE1)	0.475	0.523	0.437	0.404
	(0.022)	(0.025)	(0.020)	(0.018)
Half coef. of var.	0.855	1.076	0.784	0.704
squared (GE2)	(0.084)	(0.122)	(0.073)	(0.067)

Notes: MXN0 incomes (3 household observations for market income, 714 for taxable income) are omitted in computations of the Gini. Gini standard errors are jack-knife estimates on household-level data (recognizing that household-member incomes are copies of one another), accounting for household size. Ginis and standard errors are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-2. National Gini coefficient, various income concepts and weights

Household income concept	Unweighted sample	Sampling weighted sample	HH size & sampling weighted sample (for HH-level income)
Market income, per capita	52.46 (0.49)	52.09 (0.76)	--
Taxable income, per capita	55.17 (0.52)	54.32 (0.81)	--
Gross income, per capita	50.41 (0.48)	50.77 (0.73)	--
Disposable income, per capita	48.95 (0.47)	49.29 (0.69)	--
Market income, household	50.41 (0.47)	50.07 (0.76)	48.39 (0.89)
Taxable income, household	53.27 (0.51)	52.36 (0.80)	50.50 (0.88)
Gross income, household	47.74 (0.47)	47.97 (0.74)	46.44 (0.87)
Disposable income, household	46.08 (0.47)	46.29 (0.70)	44.79 (0.77)

Notes: Ginis are estimated on household-level data, accounting for household size in calculations for income per capita, to recognize uncertainty due to the limited number of household-level datapoints on incomes. Different weighting schemes are used to reflect inequality at the household, household-size adjusted household, or individual level. 0-incomes (3 for market income, and 714 for taxable income) are omitted in the respective computations. Jackknife estimates of standard errors are provided. Ginis and standard errors are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-3. ENIGH 2012: Sample sizes, non-response rates, and mean incomes by state

State code	State	Represented population	Strata	PSUs	Fully interviewed households (individuals)	Type A non-responding HHs (%)	Mean market income per cap.	Mean gross income per cap.	Mean disposable income per cap.	Mean final income per cap.
01	Aguascalientes	1,239,249	4	36	266 (1,055)	8 (2.9%)	43,121	45,665	41,994	45,634
02	Baja California	3,341,179	4	45	267 (945)	1 (0.4%)	61,188	66,897	60,904	66,070
03	Baja California Sur	702,275	4	30	264 (903)	7 (2.6%)	61,467	65,027	58,178	63,088
04	Campeche	869,325	4	33	287 (1,047)	2 (0.7%)	45,662	48,497	44,286	50,268
05	Coahuila de Zarag.	2,860,330	3	42	266 (918)	10 (3.6%)	40,954	43,494	39,925	43,856
06	Colima	689,225	4	39	275 (986)	5 (1.8%)	43,082	46,149	42,623	46,962
07	Chiapas	5,057,829	4	21	296 (1,287)	2 (0.7%)	18,202	20,951	20,012	24,079
08	Chihuahua	3,610,120	3	42	266 (884)	3 (1.1%)	36,084	40,658	37,697	41,842
09	Distrito Federal	8,838,219	2	80	365 (1,222)	3 (0.8%)	73,955	79,655	72,716	74,362
10	Durango	1,715,230	4	30	272 (1,058)	8 (2.9%)	30,629	33,712	31,577	36,041
11	Guanajuato	5,683,512	4	27	266 (1,109)	8 (2.9%)	36,004	38,308	35,831	39,380
12	Guerrero	3,506,806	4	24	254 (1,016)	3 (1.2%)	21,879	24,675	23,619	28,016

13	Hidalgo	2,769,342	4	21	280 (1,077)	5 (1.8%)	28,920	31,392	29,781	33,921
14	Jalisco	7,662,548	4	47	334 (1,343)	27 (7.5%)	46,244	48,307	44,788	47,940
15	México	16,182,659	4	53	372 (1,379)	7 (1.8%)	43,320	45,164	41,396	44,246
16	Mich. de Ocampo	4,482,196	4	24	288 (1,199)	2 (0.7%)	27,813	29,964	28,572	32,425
17	Morelos	1,857,630	4	30	265 (909)	5 (1.9%)	37,404	40,748	38,301	41,497
18	Nayarit	1,162,206	4	27	281 (1,086)	1 (0.4%)	37,065	40,318	37,330	41,742
19	Nuevo León	4,889,962	4	47	243 (849)	24 (9.0%)	65,982	70,759	64,115	67,019
20	Oaxaca	3,930,798	4	18	280 (1,079)	5 (1.8%)	22,618	24,758	23,116	27,369
21	Puebla	6,015,009	4	27	275 (1,035)	7 (2.5%)	30,263	31,825	29,829	33,645
22	Querétaro	1,913,297	4	30	280 (936)	5 (1.8%)	59,096	66,816	62,658	65,282
23	Quintana Roo	1,451,607	4	45	277 (1,012)	3 (1.1%)	52,303	53,583	49,146	53,515
24	San Luis Potosí	2,683,229	4	30	263 (970)	12 (4.4%)	34,245	37,024	34,634	38,327
25	Sinaloa	2,913,641	4	33	275 (1,018)	4 (1.4%)	40,037	43,588	40,350	45,532
26	Sonora	2,822,119	4	42	261 (900)	17 (6.1%)	53,640	56,726	51,956	56,896
27	Tabasco	2,313,323	4	21	263 (984)	7 (2.6%)	36,428	38,443	36,156	40,525
28	Tamaulipas	3,431,876	4	45	243 (869)	25 (9.3%)	45,601	49,697	45,499	50,477
29	Tlaxcala	1,229,949	3	36	273 (1,115)	2 (0.7%)	26,362	27,857	26,681	30,776
30	Veracruz de Ign.	7,877,528	4	32	360 (1,370)	9 (2.4%)	30,720	33,657	32,115	36,300
31	Yucatán	2,041,572	4	30	267 (1,118)	11 (4.0%)	33,963	36,972	34,587	38,400
32	Zacatecas	1,540,639	4	24	278 (1,016)	11 (3.8%)	29,534	32,413	29,862	34,547
	Nationwide	117,284,429	123	1,111	9,002 (33,694)	249 (2.7%)	41,290	44,327	41,025	44,712

Note: Mean incomes account for sampling weights but are computed only among responding households, and may not be representative of underlying population.
Source: Instituto Nacional de Estadística y Geografía (INEGI, 2013), Encuesta Nacional de Ingresos y Gastos de los Hogares 2012: Diseño muestral; Resultados Definitivos por Código de Resultados de Entrevista de Campo ENIGH – 2012; own analysis of ENIGH 2012, CEQ database.

Table A2-4. Estimation results for various univariate logistic models of response probability

Specification of $g(x_i, \theta)$	$\widehat{\theta_0}$ (s.e.)	$\widehat{\theta_1}$ (s.e.)	Sum of Squared Weighted Errors	Factor of Proportio-nality (σ^2)	AIC SIC	Per-capita market income, unweighted data, Gini (s.e.)	Per-capita market income, weighted data, Gini (s.e.)	HH market income, weighted data, Gini (s.e.)
Uncorrected						52.46 (0.49)	52.09 (0.76)	48.39 (0.89)
1: $\theta_0+\theta_1\log$(market inc.)	10.593	-0.611	295,667	0.161	296.20	52.78	52.50	55.02
	(.020)	(.002)			293.59	(0.52)	(0.81)	(1.24)
2: $\theta_0+\theta_1\log$(taxable inc.)	5.668	-0.316	636,310	0.424	320.73	55.27[a]	54.57[a]	58.34[a]
	(0.008)	(0.001)			318.11	(0.56)	(0.88)	(1.54)
3: $\theta_0+\theta_1\log$(gross inc.)	10.592	-0.605	292,641	0.160	295.87	52.76	52.47	54.98
	(.020)	(.002)			293.26	(0.51)	(0.80)	(1.23)
4: $\theta_0+\theta_1\log$(mkt. inc.pc)	8.175	-0.450	303,484	0.163	297.04	52.72	52.41	54.99
	(0.015)	(0.001)			294.42	(0.51)	(0.79)	(1.25)
5: $\theta_0+\theta_1\log$(taxbl. inc.pc)	4.578	-0.246	650,032	0.428	321.41	55.33[a]	54.61[a]	58.44[a]
	(0.006)	(0.001)			318.80	(0.55)	(0.87)	(1.56)
6: $\theta_0+\theta_1\log$(gross inc.pc)	8.150	-0.441	299,257	0.161	296.59	52.70	52.38	54.95
	(0.016)	(0.002)			293.97	(0.51)	(0.79)	(1.24)
7: $\theta_0+\theta_1\log(\text{gross inc.})^2$	6.915	-0.025	293,452	0.160	295.96	52.80	52.53	55.04
	(0.010)	(0.000)			293.35	(0.52)	(0.81)	(1.25)
8: $\theta_0+\theta_1 10^{-3}$ gross inc.	93.088	-0.030	532,100	0.284	315.00	55.84	58.75	68.28
	(0.468)	(0.000)			312.39	(3.14)	(6.33)	(13.61)
9: $\theta_0+\theta_1 10^{-12}\text{gross inc.}^2$	3.714	-0.584	285,710	0.153	295.10	54.74	55.95	61.38
	(0.001)	(0.000)			292.49	(1.41)	(2.59)	(5.73)
10: $\theta_0+\theta_1(10^{-3}\text{gross inc.})^{½}$	4.501	-0.079	296,436	0.160	296.28	53.30	53.25	55.97
	(0.002)	(0.000)			293.67	(0.60)	(0.96)	(1.62)

Note: Standard errors on Gini coefficients are jackknife estimates. [a] Because this sample omits observations with missing taxable income, Gini of market income would not be comparable to table 2. Gini of taxable income is reported instead.
Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-5. Estimation results for selected multivariate models of response probability

Specification of *g(x)*	$\widehat{\theta_0}$ (s.e.)	$\widehat{\theta_1}$ (s.e.)	$\widehat{\theta_2}$ (s.e.)	$\widehat{\theta_3}$ (s.e.)	$\widehat{\theta_4}$ (s.e.)	Sum of Squared Weighted Errors	Factor of Proportio -nality (σ^2)	AIC	Mkt. inc. per cap. weighted data, Gini (s.e.)
Uncorrected									52.09 (0.76)
$\theta_0+\theta_1 10^{-6}$gross inc.	3.685	0.231	-0.657			285,673	0.153	297.10	55.98
$+\theta_2 10^{-12}$gross inc.^{2}	(0.002)	(0.012)	(0.004)						(2.66)
$\theta_0+\theta_1 10^{-6}$gross inc.pc	3.699	-0.688	-0.440			228,694	0.126	289.98	56.48
$+\theta_2 10^{-12}$gross inc.pc^2	(0.001)	(0.024)	(0.007)						(4.57)
$\theta_0+\theta_1$log(gross inc.pc)	51.704	-1.647	2.162	-87.016	55.082	231,404	0.158	294.36	52.66
$+\theta_2$urban$+\theta_3$age$+\theta_4$age^2	(0.179)	(0.002)	(0.003)	(0.534)	(0.388)				(0.88)
$\theta_0+\theta_1$log(gross inc.pc)	255.093	-0.838	-13.148	-620	395.035	250,534	0.198	296.90	51.90
$+\theta_2$sch.attend$+\theta_3$age$+\theta_4$age^2	(0.868)	(0.001)	(0.036)	(2.269)	(1.482)				(0.75)
$\theta_0+\theta_1$log(gross inc.pc)	7.398	-0.158	-1.887	-2.180		289,632	0.153	299.54	51.98
$+\theta_2$urban$+\theta_3$HH size	(0.028)	(0.003)	(0.015)	(0.007)					(0.77)
$\theta_0+\theta_1$log(gross inc.pc)	24.272	-0.273	-15.935	-76.587	252.898	277,454	0.152	300.17	51.85
$+\theta_2$urb$+\theta_3$hsize$+\theta_4$hsize^2	($11\cdot10^3$)	(0.003)	($11\cdot10^3$)	(0.364)	(2.049)				(0.78)
$\theta_0+\theta_1$log(gross inc.pc)	26.633	-0.683	-15.496	-23.586		278,514	0.151	298.29	52.65
$+\theta_2$male$+\theta_3$hsize	($23\cdot10^3$)	(0.002)	($23\cdot10^3$)	(0.064)					(0.85)
$\theta_0+\theta_1$log(gross inc.pc)	32.360	-1.194	-17.268	1.728		275,568	0.167	297.95	54.41
$+\theta_2$male$+\theta_3$urban	($27\cdot10^3$)	(0.001)	($27\cdot10^3$)	(0.003)					(1.26)
$\theta_0+\theta_1$log(gross inc.pc)	8.123	-0.438	-0.265			299,251	0.161	298.59	52.38
$+\theta_2$postsecondary.attend	(0.016)	(0.002)	(0.040)						(0.79)
$\theta_0+\theta_1$log(gross inc.pc)	20.778	-0.738	-4.585	-16.566		265,835	0.167	296.80	52.12
$+\theta_2$postsecond.attend$+\theta_3$age	(0.027)	(0.002)	(0.018)	(0.027)					(0.77)

Note: Standard errors on Gini coefficients are jackknife estimates. Variables are normalized: age=(years-12)/100 ; HH size=(#-1)/100.
Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-6. ENIGH 2012: Income summary statistics for various income concepts, nonresponse corrected weights

	Market income per cap.	Market income + pensions per cap.	Gross income per cap.	Taxable income per cap.	Net market income per cap.	Disposable income per cap.	Consumable income per cap.	Final income per cap.
99.9th %ile	656,327	642,581	642,581	515,538	536,234	536,234	524,765	526,396
99th %ile	281,576	315,599	315,599	209,259	272,727	272,727	260,226	262,465
95th %ile	133,423	141,873	142,222	95,663	127,003	127,170	124,157	127,759
90th %ile	86,077	90,777	91,450	61,967	83,688	84,098	82,499	86,346
75th %ile	46,670	49,318	49,738	33,197	45,473	45,843	45,485	50,035
Mean	42,309	44,432	45,322	29,618	40,992	41,882	41,234	45,536
Median	25,552	26,631	27,349	17,935	25,466	25,967	25,944	30,219
25th %ile	14,094	14,614	15,927	8,476	14,313	15,536	15,636	20,035
10th %ile	7,506	7,881	9,467	2,400	7,820	9,358	9,402	13,957
5th %ile	4,652	4,812	6,898	538	4,806	6,876	6,952	11,061
1st %ile	1,984	2,003	3,703	0	2,003	3,703	3,785	7,340
Std. dev.	69,207	71,747	71,550	55,500	64,366	64,180	61,114	60,950
Skewness	14.27	13.32	13.40	20.83	15.52	15.63	14.49	14.51
Kurtosis	462.14	406.97	410.82	947.45	572.59	578.43	504.98	507.11
Sample	33,686	33,686	33,686	33,686	33,686	33,686	33,686	33,686
Top 0.1% inc. share	3.02	2.97	2.91	3.53	2.84	2.78	2.66	2.41
0.1-1% inc. share	8.60	8.47	8.30	9.13	7.90	7.74	7.56	6.90
1-5% inc. share	17.17	17.41	17.11	17.77	17.08	16.77	16.58	15.33
5-10% inc. share	12.56	12.70	12.52	12.73	12.44	12.25	12.19	11.51
Gini (HH-size & sampling weighted data)	52.50	52.76	51.19	54.70	51.33	49.68	49.02	44.42
	(0.81)	(0.77)	(0.78)	(0.87)	(0.74)	(0.74)	(0.73)	(0.72)
Gini (HH-size wted, no sampling wght)	52.78	53.07	50.74	55.44	51.70	49.27	48.64	43.50
	(0.52)	(0.51)	(0.51)	(0.56)	(0.50)	(0.50)	(0.49)	(0.48)
Gini (equal HH weights = unweighted)	54.89	55.02	53.02	58.59	53.72	51.64	50.93	46.88
	(0.74)	(0.71)	(0.72)	(0.84)	(0.70)	(0.71)	(0.69)	(0.69)
Mean log dev. (GE0)	0.510	0.516	0.461	0.613	0.485	0.431	0.419	0.329
	(0.010)	(0.010)	(0.010)	(0.013)	(0.010)	(0.010)	(0.009)	(0.008)
Theil index (GE1)	0.556	0.558	0.529	0.611	0.525	0.496	0.480	0.399
	(0.024)	(0.023)	(0.022)	(0.033)	(0.023)	(0.023)	(0.021)	(0.019)
Half coef. of var. squared (GE2)	1.338	1.304	1.246	1.756	1.233	1.174	1.098	0.896
	(0.242)	(0.219)	(0.211)	(0.470)	(0.253)	(0.243)	(0.211)	(0.174)

Notes: Statistics are based on non-response correction weights estimated in the logarithmic model of market income (model 1). These statistics exclude 3 household observations with market income of 0. The statistics are still comparable to those in table 1, which are extremely robust to this exclusion (changing by 0.01 at most.) (Another 714 household observations are omitted in computations of the Gini for taxable income.) Gini standard errors are jack-knife estimates on household-level data (recognizing that household-member incomes are copies of one another), accounting for household size except in last row. Ginis and standard errors are multiplied by 100 for clarity of presentation.

Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-7. Distribution of top incomes, various income concepts: Pareto (type I) estimates

Cutoff percentile	Market income per cap.	Taxable income per cap.	Gross income per cap.	Disposable income per cap.	Market income, household	Taxable income, household	Gross income, household	Disposable income, household
Pareto (type I) coefficient α								
Top 25%	1.63 (.04)	1.61 (.04)	1.64 (.04)	1.67 (.04)	1.87 (.06)	1.87 (.06)	1.84 (.06)	1.92 (.06)
Top 10%	1.83 (.07)	1.81 (.07)	1.84 (.07)	1.89 (.07)	2.10 (.11)	2.11 (.11)	2.20 (.12)	2.21 (.11)
Top 5%	2.14 (.13)	2.11 (.13)	2.22 (.13)	2.18 (.12)	2.31 (.20)	2.37 (.21)	2.39 (.22)	2.54 (.22)
Top 1%	2.31 (.23)	2.35 (.27)	2.53 (.27)	2.74 (.40)	2.44 (.40)	1.96 (.25)	2.28 (.38)	2.46 (.36)
Top 0.1%	1.74 (.43)	1.76 (.44)	1.52 (.26)	1.67 (.35)	11.58 (6.0)	5.57 (3.41)	12.42 (4.9)	4.48 (1.41)
Inverted Pareto coefficient $\beta = \alpha/[\alpha - 1]$								
Top 25%	2.59	2.64	2.56	2.49	2.15	2.15	2.19	2.09
Top 10%	2.20	2.23	2.19	2.12	1.91	1.90	1.83	1.83
Top 5%	1.88	1.90	1.82	1.85	1.76	1.73	1.72	1.65
Top 1%	1.76	1.74	1.65	1.57	1.69	2.04	1.78	1.68
Top 0.1%	2.35	2.32	2.92	2.49	1.09	1.22	1.09	1.29
Top income share (%)								
Top 25%	65.96	68.35	65.04	63.81	61.12	63.53	60.49	58.76
Top 10%	43.15	44.75	42.66	41.16	37.57	38.92	36.49	35.32
Top 5%	28.99	30.43	28.45	27.61	24.95	25.79	24.10	22.78
Top 1%	11.46	12.25	11.05	10.03	9.89	11.34	9.57	8.85
Top 0.1%	3.39	3.62	3.51	3.03	0.52	0.63	0.59	0.87
Gini among top incomes (×100)								
Top 25%	44.30	45.18	43.71	42.91	36.54	36.55	37.39	35.28
Top 10%	37.62	38.21	37.19	36.00	31.32	30.99	29.33	29.18
Top 5%	30.49	31.14	29.07	29.79	27.56	26.78	26.44	24.55
Top 1%	27.59	26.99	24.65	22.27	25.82	34.19	28.01	25.57
Top 0.1%	40.30	39.54	49.08	42.88	4.51	9.86	4.19	12.57
Half coefficient of variation squared (GE2)[i]								
Top 10%	--	--	--	--	1.44	0.94	0.51	0.46
Top 5%	0.84	1.41	0.48	0.56	0.33	0.24	0.21	0.13
Top 1%	0.22	0.26	0.16	0.09	0.20	--	0.24	0.17
Top 0.1%	--	--	--	--	0.0000	0.0002	0.0014	0.0004
Minimum income								
Top 25%	45,967	32,772	49,088	45,367	194,623	144,590	201,836	188,979
Top 10%	84,011	59,957	90,000	82,279	324,495	243,256	349,393	315,991
Top 5%	128,837	92,348	139,344	124,248	462,288	349,826	491,296	447,836
Top 1%	274,420	203,607	295,241	264,017	955,536	661,967	951,337	850,260
Top 0.1%	610,435	515,538	642,581	536,234	2,682,735	1,933,700	2,650,090	1,992,925
Mean income								
Top 25%	118,759	86,809	125,432	113,404	417,055	310,181	443,812	395,176
Top 10%	185,464	133,920	196,722	174,445	624,850	460,674	639,892	574,836
Top 5%	242,700	177,718	254,243	229,045	821,342	606,089	838,281	731,199
Top 1%	470,864	354,848	491,538	417,642	1,626,169	1375,461	1,645,154	1,424,930
Top 0.1%	1,410,079	1,194,779	1,453,842	1,194,596	2,766,433	2,077,948	3,036,647	2,186,562

[i] Unable to calculate for other top income groups.
Notes: Pareto robust standard errors reported.
Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-8. Replacement of top incomes with Pareto I distribution: uncorrected vs. corrected Ginis

Cutoff percentile	Market income per cap.	Market inc + pensions per cap.	Gross income per cap.	Taxable income per cap.	Net market income per cap.	Disposable income per cap.	Consumable income per cap.	Final income per cap.
Nonparametric Gini among bottom incomes								
Top 25%	31.55 (0.34)	31.85 (0.35)	29.25 (0.31)	34.95 (0.41)	31.06 (0.34)	28.43 (0.31)	28.06 (0.30)	23.56 (0.25)
Top 10%	37.53 (0.33)	37.83 (0.33)	35.73 (0.31)	40.16 (0.38)	36.86 (0.33)	34.73 (0.30)	34.27 (0.30)	29.86 (0.27)
Top 5%	41.63 (0.37)	41.98 (0.38)	40.09 (0.37)	43.80 (0.39)	40.85 (0.37)	38.92 (0.35)	38.41 (0.35)	33.92 (0.33)
Top 1%	47.58 (0.46)	47.97 (0.45)	46.25 (0.45)	49.56 (0.47)	46.81 (0.45)	45.03 (0.44)	44.63 (0.48)	39.90 (0.42)
Top 0.1%	50.83 (0.66)	51.26 (0.64)	49.62 (0.64)	52.97 (0.68)	49.84 (0.59)	48.13 (0.60)	47.52 (0.59)	42.89 (0.58)
Nonparametric Gini among top incomes								
Top 25%	35.02 (1.22)	35.10 (1.15)	34.96 (1.15)	35.80 (1.37)	34.07 (1.10)	33.94 (1.10)	33.45 (1.06)	31.80 (1.05)
Top 10%	30.17 (1.64)	29.80 (1.53)	29.69 (1.52)	31.62 (1.87)	29.05 (1.51)	28.92 (1.51)	28.46 (1.46)	27.60 (1.44)
Top 5%	27.72 (2.04)	27.09 (1.92)	27.03 (1.92)	29.33 (2.40)	26.33 (2.00)	26.23 (2.00)	25.77 (1.93)	25.27 (1.91)
Top 1%	24.48 (4.31)	23.23 (4.12)	23.21 (4.11)	27.37 (4.98)	24.38 (4.31)	24.33 (4.30)	24.26 (4.22)	23.95 (4.09)
Top 0.1%	24.51 (9.14)	20.32 (9.00)	20.32 (9.00)	27.24 (12.77)	25.41 (10.50)	25.40 (10.50)	24.78 (10.11)	24.71 (10.09)
Pareto (type I) coefficient								
Top 25%	1.63 (.04)	1.63 (0.04)	1.64 (0.04)	1.61 (0.04)	1.66 (0.04)	1.67 (0.04)	1.68 (0.04)	1.78 (0.05)
Top 10%	1.83 (.07)	1.82 (0.07)	1.84 (0.07)	1.81 (0.07)	1.90 (0.07)	1.89 (0.07)	1.91 (0.07)	1.97 (0.07)
Top 5%	2.14 (.13)	2.17 (0.13)	2.22 (0.13)	2.11 (0.13)	2.19 (0.12)	2.18 (0.12)	2.21 (0.12)	2.30 (0.12)
Top 1%	2.31 (.23)	2.49 (0.27)	2.53 (0.27)	2.35 (0.27)	2.74 (0.40)	2.74 (0.40)	2.73 (0.38)	2.91 (0.45)
Top 0.1%	1.74 (.43)	1.52 (0.26)	1.52 (0.26)	1.76 (0.44)	1.67 (0.35)	1.67 (0.35)	1.61 (0.32)	1.62 (0.32)
Parametric Gini among top incomes								
Top 25%	44.30	44.26	43.71	45.18	43.15	42.91	42.27	39.05
Top 10%	37.62	37.99	37.19	38.21	35.72	36.00	35.37	33.92
Top 5%	30.49	29.97	29.07	31.14	29.64	29.79	29.28	27.79
Top 1%	27.59	25.12	24.65	26.99	22.32	22.27	22.39	20.73
Top 0.1%	40.30	49.08	49.08	39.54	42.87	42.88	45.03	44.79
Semiparametric Gini								
Top 25%	56.32	56.58	54.83	59.39	55.07	53.38	52.63	47.22
Top 10%	53.99	54.47	52.69	56.45	52.62	51.04	50.32	45.50
Top 5%	52.53	52.87	51.11	54.86	51.47	49.81	49.14	44.35
Top 1%	52.23	52.45	50.81	54.34	50.87	49.18	48.55	43.85
Top 0.1%	52.35	52.88	51.27	54.54	51.27	49.57	48.97	44.32
Uncorrected	52.09	52.39	50.77	54.32	50.98	49.29	48.63	43.99
Gini correction								
Top 25%	+4.23	+2.26	+4.06	+10.10	+5.00	+1.02	+4.66	+0.93
Top 10%	+1.90	+0.15	+1.92	+7.16	+2.55	–1.32	+2.35	–0.79
Top 5%	+0.44	–1.45	+0.34	+5.57	+1.40	–2.55	+1.17	–1.94
Top 1%	+0.14	–1.87	+0.04	+5.05	+0.80	–3.18	+0.58	–2.44
Top 0.1%	+0.26	–1.44	+0.50	+5.25	+1.20	–2.79	+1.00	–1.97

Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-9. Income summary statistics: Replacing top net market incomes with Pareto I estimates, and imputing other income concepts by CEQ Method

	Market income per cap.	Market inc +pensions per cap.	Gross income per cap.	Taxable income per cap.	Net market income per cap.	Disposable income per cap.	Consumable income per cap.	Final income per cap.
99.9th %ile	1,018,660	1,018,660	1,018,660	805,485	917,787	917,787	865,873	868,318
99th %ile	297,023	329,833	329,833	225,028	292,123	292,123	288,573	288,573
95th %ile	125,060	129,735	130,308	89,033	115,243	115,936	113,154	118,098
90th %ile	83,247	88,376	88,825	59,825	82148	82,293	80,812	84,701
75th %ile	45,803	48,508	49,088	32,772	44992	45,367	44,792	49,133
Mean	42,591	44,730	45,628	30,187	41,427	42,326	41,619	45,922
Median	25,232	26,253	27,012	17,667	25,196	25,768	25,706	29,934
25th %ile	13,865	14,432	15,741	8,354	14,148	15,367	15,512	19,866
10th %ile	7,364	7,775	9,419	2,313	7,723	9,295	9,358	13,902
5th %ile	4,552	4,765	6,804	492	4,754	6,781	6,843	11,010
1st %ile	1,935	2,003	3,609	0	1,987	3,609	3,713	7,257
Std. dev.	102,090	104,859	104,718	91,988	99,416	99,288	92,713	92,575
Skewness	43.11	40.50	40.64	55.99	46.28	46.44	43.84	43.94
Kurtosis	3,191.63	2,881.43	2,895.64	4,732.05	3,548.63	3,565.36	3,270.63	3,282.45
Sample	33,694	33,694	33,694	32,296	33,694	33,694	33,694	33,694
Gini (HH-size &	53.68	53.85	52.25	56.37	52.60	50.93	50.20	45.54
sampling weighted)	(1.07)	(1.03)	(1.04)	(1.27)	(1.06)	(1.07)	(1.03)	(1.02)
Top 0.1% inc. share	4.61	4.53	4.45	5.72	4.53	4.43	4.22	3.83
0.1-1% inc. share	10.22	10.20	10.005	11.32	9.86	9.65	9.39	8.56
1-5% inc. share	16.14	16.23	15.96	16.35	15.80	15.53	15.40	14.28
5-10% inc. share	11.84	11.87	11.70	11.89	11.58	11.40	11.37	10.77
Mean log dev.	0.535	0.539	0.483	0.651	0.512	0.457	0.442	0.350
(GE0)	(0.018)	(0.017)	(0.017)	(0.024)	(0.018)	(0.017)	(0.016)	(0.014)
Theil index (GE1)	0.645	0.644	0.613	0.741	0.621	0.589	0.567	0.477
	(0.057)	(0.053)	(0.053)	(0.082)	(0.058)	(0.057)	(0.053)	(0.048)
Half coef. of var.	2.873	2.748	2.633	4.643	2.879	2.751	2.481	2.032
squared (GE2)	(1.320)	(0.196)	(1.151)	(2.578)	(1.396)	(1.339)	(1.158)	(0.956)

Notes: MXN0 incomes (3 household observations for market income, 714 for taxable income) are omitted in computations of the Gini. Gini standard errors are jack-knife estimates on household-level data (recognizing that household-member incomes are copies of one another), accounting for household size. Ginis and standard errors are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-10. Summary results of correction methods: corrected Ginis for all income concepts

	Market income per cap.	Market inc.+ pensions per cap.	Gross income per cap.	Taxable income per cap.	Net market income per cap.	Disposable income per cap.	Consumable income per cap.	Final income per cap.
Uncorrected	52.1	52.4	50.8	54.3	51.0	49.3	48.6	44.0
Correction by reweighting (models in table A2-4)								
Minimum	52.4 **+0.3**	52.7 **+0.3**	51.1 **+0.3**	54.6 **+0.3**	51.1 **+0.1**	49.6 **+0.3**	48.9 **+0.3**	44.3 **+0.3**
Mean	53.2 **+1.1**	53.5 **+1.1**	51.9 **+1.1**	55.7 **+1.4**	52.1 **+1.1**	50.5 **+1.2**	49.8 **+1.2**	45.2 **+1.2**
Median	52.4 **+0.3**	52.7 **+0.3**	51.2 **+0.4**	54.7 **+0.4**	51.3 **+0.3**	49.6 **+0.3**	49.0 **+0.4**	44.4 **+0.4**
Max	58.7 **+6.6**	58.7 **+6.3**	57.2 **+6.4**	62.7 **+8.4**	57.9 **+6.9**	56.3 **+7.0**	55.3 **+6.7**	50.7 **+6.7**
Correction by Pareto (type I) replacing of own income concept								
Minimum	52.2 **+0.1**	52.5 **+0.1**	50.8 **+0.0**	54.3 **+0.0**	50.9 **–0.1**	49.2 **–0.1**	48.6 **+0.0**	43.9 **–0.1**
Mean	53.5 **+1.4**	53.9 **+1.5**	52.1 **+1.4**	55.9 **+1.6**	52.3 **+1.3**	50.6 **+1.3**	49.9 **+1.3**	45.0 **+1.0**
Median	52.5 **+0.4**	52.9 **+0.5**	51.3 **+0.5**	54.9 **+0.6**	51.5 **+0.5**	49.8 **+0.5**	49.1 **+0.5**	44.4 **+0.4**
Max	56.3 **+4.2**	56.6 **+4.2**	54.8 **+4.1**	59.4 **+5.1**	55.1 **+4.1**	53.4 **+4.1**	52.6 **+4.0**	47.2 **+3.2**
Correction by Pareto (type I) replacing of market income + CEQ Method								
Minimum	52.0 **–0.1**	52.3 **–0.1**	50.6 **–0.1**	54.2 **–0.1**	50.8 **–0.1**	49.1 **–0.2**	48.5 **–0.2**	43.8 **–0.1**
Mean	53.3 **+1.2**	53.5 **+1.1**	51.9 **+1.1**	55.9 **+1.6**	52.2 **+1.2**	50.5 **+1.2**	49.8 **+1.2**	45.1 **+1.1**
Median	52.5 **+0.4**	52.8 **+0.4**	51.2 **+0.4**	54.9 **+0.6**	51.4 **+0.5**	49.7 **+0.5**	49.1 **+0.4**	44.4 **+0.4**
Max	55.9 **+3.8**	55.9 **+3.5**	54.3 **+3.5**	59.4 **+5.0**	54.8 **+3.8**	53.1 **+3.9**	52.2 **+3.6**	47.5 **+3.5**

Notes: These Ginis are comparable to 'Gini (HH-size & sampling weighted data)' in tables A2-1, A2-6 and A2-9, and 'Semiparametric Gini' in table A2-8. Ginis, and differences in them, are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2012, CEQ database.

Table A2-11. Redistributive impacts of fiscal tools: high/center/low estimates

	Market income inequality	+ Net contribut. pensions Market → Market+Pensions	+ Cash-like transfers Market+Pensions → Gross	+ Nontaxable income Taxable → Gross	– Direct taxes Gross → Disposable[i]	– Indirect taxes & subsidies Disposable → Consumable	+ Net in-kind programs Consumable → Final	Final income inequality
Gini coefficient: pc.pt. change								
Uncorrected	52.09	+0.18	–1.60	–4.12	–1.32	–0.72	–4.66	43.09
Income distrib. corrected for nonresponse by reweighting								
High	52.57	+0.26	–1.58	–3.53	–1.51	–0.66	–4.60	44.48
Center	52.50	+0.26	–1.57	–3.51	–1.51	–0.66	–4.60	44.42
Low	52.41	+0.28	–1.59	–3.53	–1.49	–0.66	–4.60	44.34
Each income concept corrected for top income mismeasurement by replacing								
High	56.32	+0.26	–1.75	–4.56	–1.45	–0.75	–5.41	47.22
Center	53.99	+0.48	–1.78	–3.76	–1.65	–0.72	–4.82	45.54
Low	52.23	+0.22	–1.64	–3.53	–1.63	–0.63	–4.70	43.85
Market income corrected for top income biases by replacing + CEQ Method								
High	55.89	–0.03	–1.59	–5.08	–1.13	–0.90	–4.71	47.54
Center	53.68	+0.18	–1.60	–4.12	–1.32	–0.72	–4.66	45.54
Low	51.96	+0.30	–1.62	–3.53	–1.50	–0.66	–4.64	43.84
Top 10 percent income share: pc.pt. change								
Uncorrected	40.89	+0.23	–0.72	–2.30	–1.26	–0.55	–2.84	35.75
Income distrib. corrected for nonresponse by reweighting								
High	41.43	+0.19	–0.70	–2.32	–1.31	–0.55	–2.85	36.21
Center	41.35	+0.20	–0.71	–2.32	–1.30	–0.55	–2.84	36.15
Low	41.25	+0.20	–0.70	–2.33	–1.29	–0.54	–2.85	36.07
Each income concept corrected for top income mismeasurement by replacing								
High	47.13	+0.23	–0.84	–2.43	–1.34	–0.65	–4.31	40.22
Center	43.15	+0.46	–0.95	–2.08	–1.51	–0.61	–3.15	37.41
Low	37.75	–1.01	–0.69	–2.60	–1.93	–0.19	–2.94	30.99
Market income corrected for top income biases by replacing + CEQ Method								
High	45.93	–0.06	–0.75	–4.29	–0.85	–0.86	–3.12	40.29
Center	42.81	+0.02	–0.71	–3.17	–1.11	–0.63	–2.94	37.44
Low	40.74	+0.21	–0.70	–2.25	–1.29	–0.55	–2.85	35.56
Top 1 percent income share: pc.pt. change								
Uncorrected	11.33	–0.12	–0.22	–1.35	–0.67	–0.30	–0.90	9.12
Income distrib. corrected for nonresponse by reweighting								
High	11.68	–0.19	–0.22	–1.45	–0.70	–0.31	–0.91	9.35
Center	11.62	–0.18	–0.23	–1.45	–0.69	–0.30	–0.91	9.31
Low	11.59	–0.17	–0.22	–1.43	–0.69	–0.30	–0.91	9.30
Each income concept corrected for top income mismeasurement by replacing								
High	15.99	+0.45	–0.35	–1.44	–1.03	–1.14	–1.34	12.58
Center	13.95	+0.52	–0.37	–1.13	–1.01	–0.99	–0.75	11.33
Low	11.46	–0.19	–0.21	–1.20	–1.03	–0.65	–0.62	8.77
Market income corrected for top income biases by replacing + CEQ Method								
High	19.89	–0.43	–0.36	–4.88	–0.02	–0.84	–1.60	16.64
Center	14.83	–0.10	–0.27	–2.59	–0.38	–0.47	–1.22	12.39
Low	11.16	–0.20	–0.21	–1.40	–0.70	–0.30	–0.89	8.86

[i] Alternatively, this can be obtained as 'market+pensions → net market' for estimates within 0.1 pc.pt. of those above.
Notes: These Ginis are comparable to 'Gini (HH-size & sampling weighted data)' in tables 1 & 3, and 'Semiparametric Gini' in table 4.
Source: Own analysis of ENIGH 2012, CEQ database.

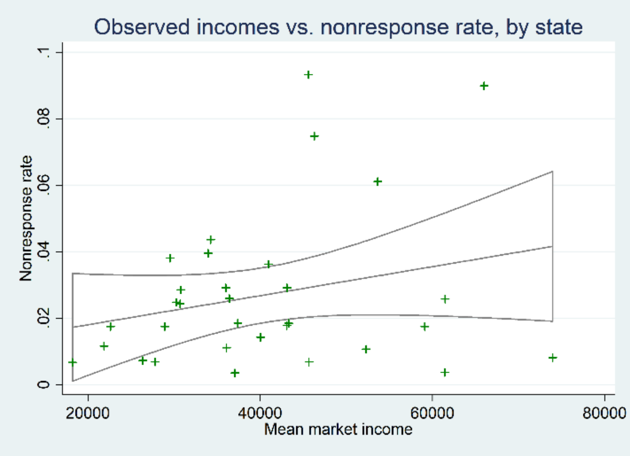
Table A2-12. Generalized Pareto (type II) results, various income concepts (individual sampling-weighted sample)

Cutoff percentile	Market income per cap.	Taxable income per cap.	Gross income per cap.	Disposable income per cap.	Market income, household	Taxable income, household	Gross income, household	Disposable income, household
Pareto (type II) shape coefficient ξ								
Top 25%	0.44 (.02)	0.43 (.02)	0.45 (.02)	0.42 (.02)	0.36 (.05)	0.37 (.05)	0.33 (.05)	0.32 (.05)
Top 10%	0.33 (.04)	0.37 (.04)	0.32 (.04)	0.31 (.04)	0.34 (.07)	0.38 (.07)	0.36 (.07)	0.31 (.06)
Top 5%	0.37 (.05)	0.41 (.05)	0.39 (.06)	0.32 (.05)	0.41 (.11)	0.47 (.09)	0.45 (.12)	0.48 (.11)
Top 1%	0.34 (.11)	0.44 (.12)	0.32 (.09)	0.45 (.12)	0.01 (.33)	-0.15 (.15)	-0.01 (.45)	-0.05 (.28)
Top 0.1%	0.10 (.22)	0.25 (.25)	0.07 (.21)	0.27 (.34)	--[i]	--	--	--
Pareto (type II) scale coefficient, log(σ)								
Top 25%	10.42 (.03)	10.10 (.03)	10.45 (.03)	10.39 (.03)	11.71 (.05)	11.42 (.06)	11.81 (.05)	11.69 (.05)
Top 10%	10.94 (.04)	10.58 (.04)	11.01 (.04)	10.89 (.04)	12.09 (.08)	11.75 (.08)	12.06 (.08)	12.00 (.08)
Top 5%	11.10 (.06)	10.77 (.07)	11.11 (.07)	11.08 (.06)	12.24 (.16)	11.86 (.13)	12.19 (.18)	11.97 (.16)
Top 1%	11.72 (.11)	11.36 (.15)	11.76 (.09)	11.40 (.15)	13.26 (.50)	13.33 (.28)	13.32 (.61)	13.17 (.42)
Top 0.1%	13.17 (.33)	12.88 (.32)	13.23 (.33)	12.86 (.43)	--	--	--	--
1/ξ								
Top 25%	2.29	2.30	2.20	2.39	2.74	2.73	3.03	3.13
Top 10%	3.05	2.71	3.14	3.26	2.96	2.66	2.76	3.28
Top 5%	2.67	2.44	2.56	3.13	2.43	2.13	2.22	2.10
Top 1%	2.92	2.28	3.16	2.24	67.59	-6.71	-99.93	-19.90
Top 0.1%	9.77	4.06	15.38	3.72	--	--	--	--
Inverted Pareto coefficient								
Top 25%	1.77	1.77	1.83	1.72	1.57	1.58	1.49	1.47
Top 10%	1.49	1.58	1.47	1.44	1.51	1.60	1.57	1.44
Top 5%	1.60	1.70	1.64	1.47	1.70	1.88	1.82	1.91
Top 1%	1.52	1.78	1.46	1.80	1.02	0.87	0.99	0.95
Top 0.1%	1.11	1.33	1.07	1.37	--	--	--	--
Gini among top incomes								
Top 25%	36.02	36.31	36.50	34.83	30.37	30.51	29.89	28.66
Top 10%	29.92	31.30	29.49	28.62	27.34	27.92	26.74	25.19
Top 5%	27.77	29.31	27.40	25.86	27.21	28.22	27.13	26.46
Top 1%	24.47	27.47	23.02	24.45	19.09	20.78	21.02	18.04
Top 0.1%	25.77	28.69	24.84	28.67	--	--	--	--
Log pseudo-likelihood ($LL/10^6$)								
Top 25%	-348.00	-338.00	-349.00	-346.00	-384.00	-375.00	-385.00	-382.00
Top 10%	-144.00	-140.00	-145.00	-143.00	-157.00	-154.00	-157.00	-156.00
Top 5%	-72.90	-71.40	-73.30	-72.70	-79.80	-77.80	-80.20	-79.40
Top 1%	-15.60	-15.00	-15.30	-14.80	-16.60	-16.40	-17.20	-16.70
Top 0.1%	-1.675	-1.461	-1.677	-1.751	--	--	--	--

[i] Too few observations to fit.

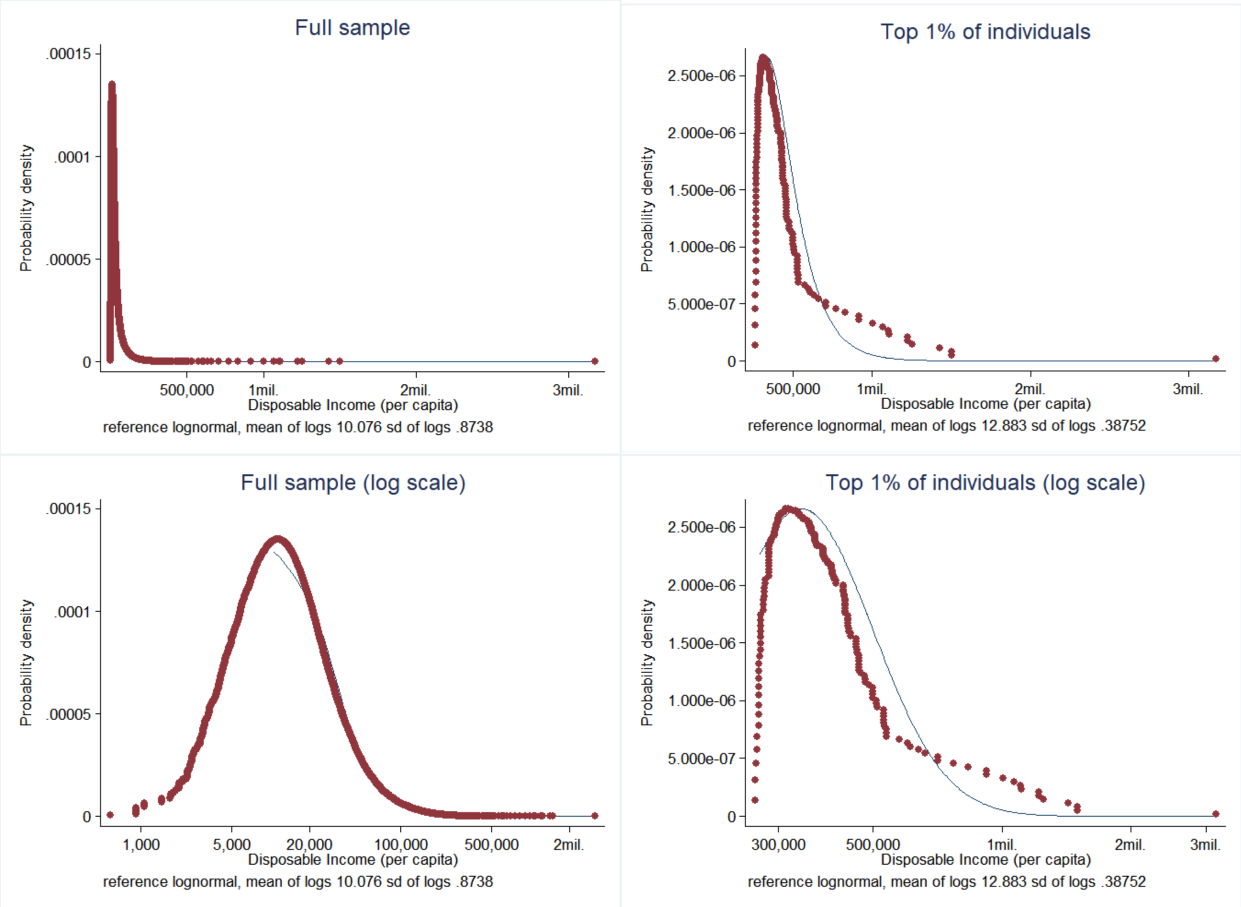
Source: Own analysis of ENIGH 2012, CEQ database.

Figure A2-1. Mean observed market income among respondents and nonresponse rate, by state



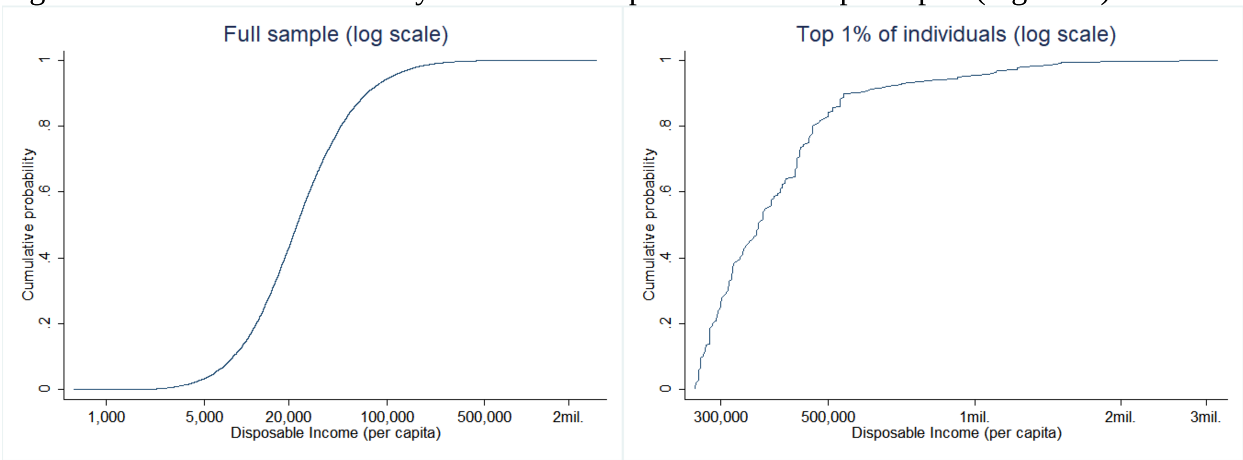
Notes: 95% confidence interval around linear fitted line is shown.
Source: Own analysis of ENIGH 2012, CEQ database.

Figure A2-2. Density function of disposable income per capita, with reference lognormal density



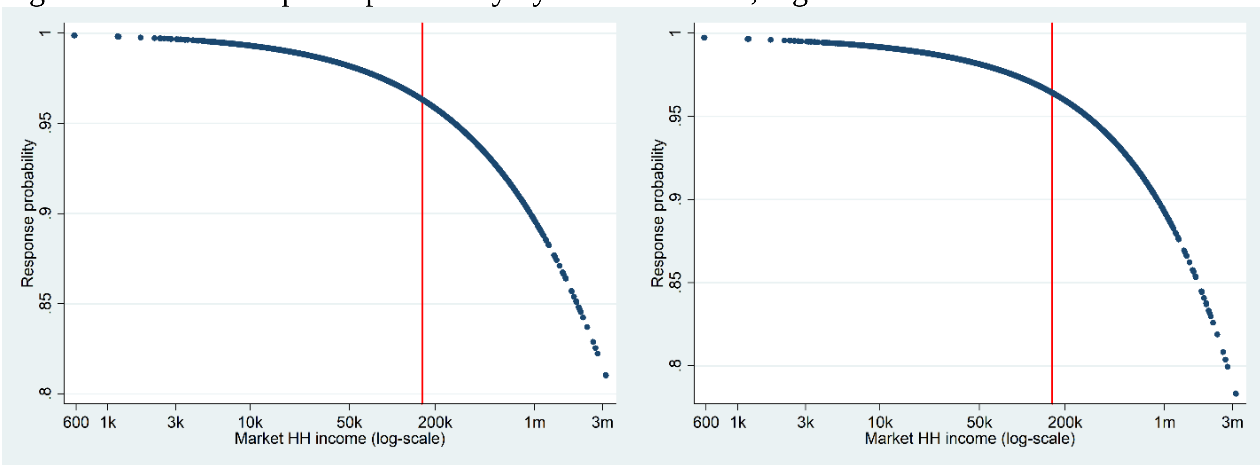
Source: Own analysis of ENIGH 2012, CEQ database.

Figure A2-3. Cumulative density function of disposable income per capita (log scale)



Source: Own analysis of ENIGH 2012, CEQ database.

Figure A2-4. Unit response probability by market income, logarithmic model of market income

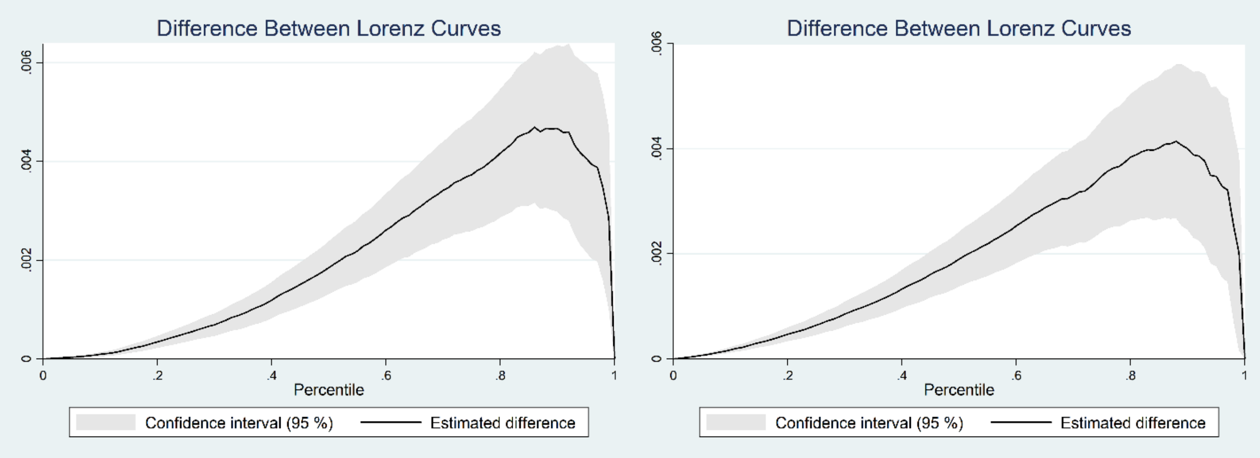


i. Logarithmic model of market income (model 1)
$\theta_0+\theta_1\log(\text{market inc.})$

ii. Quadratic logarithmic model of market income
$\theta_0+\theta_1\log(\text{market inc.})^2$

Notes: Red line shows mean market household income in the corrected income distribution: i) 157,468 and ii) 157,665.
Source: Own analysis of ENIGH 2012, CEQ database.

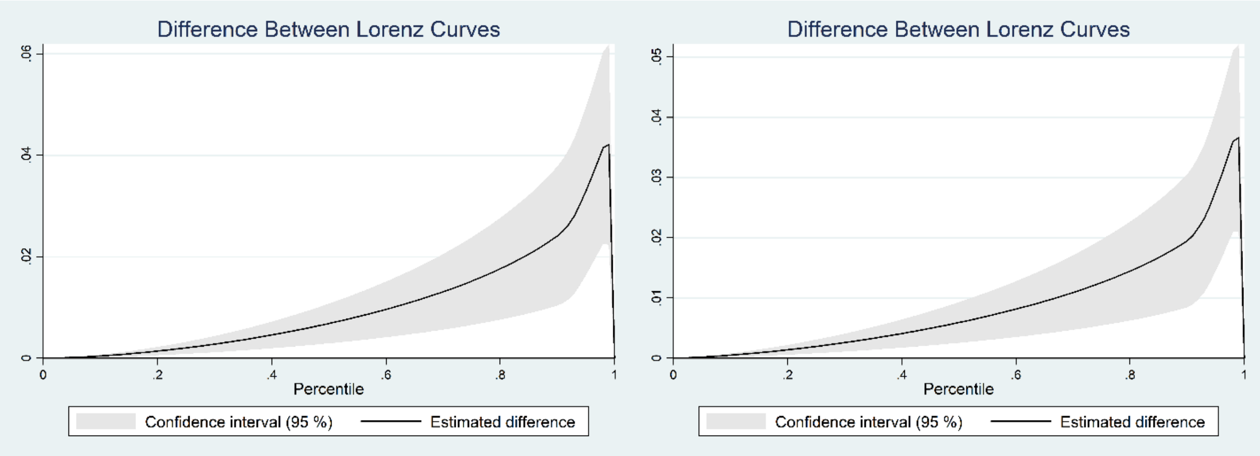
Figure A2-5. Lorenz curve: income per capita, uncorrected versus unit-nonresponse corrected weights (model 1)



i. Market income per capita ii. Disposable income per capita

Notes: Positive values indicate that the Lorenz curve uncorrected for unit-nonresponse dominates, and shows less inequality than the corrected Lorenz curve. Distributions account for sampling weights and household size.
Source: Own analysis of ENIGH 2012, CEQ database.

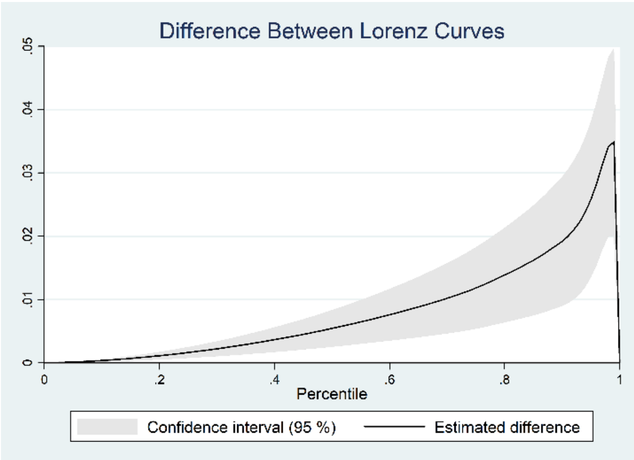
Figure A2-6. Lorenz curve: market and disposable income per capita, top 10% of incomes replaced with Pareto I values

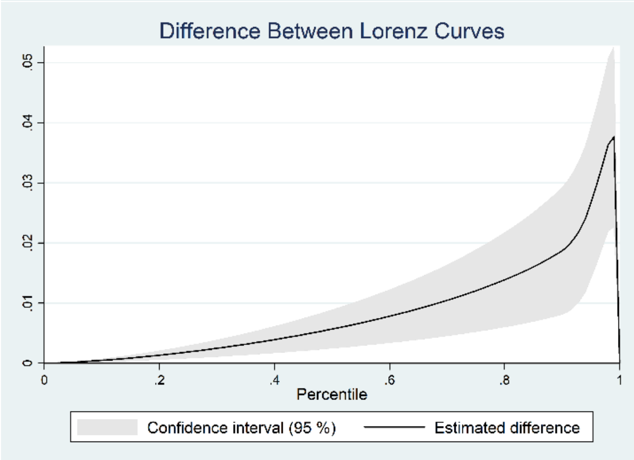


i. Market income per capita i. Disposable income per capita

Notes: Pareto replacing is performed on own income concept. Positive values indicate that the uncorrected Lorenz curve dominates, and shows less inequality than the corrected Lorenz curve. Distributions account for sampling weights and household size.
Source: Own analysis of ENIGH 2012, CEQ database.

Figure A2-7. Lorenz curve: market and disposable income per capita, top 10% of net market incomes replaced with Pareto I values, other income concepts imputed using CEQ Method



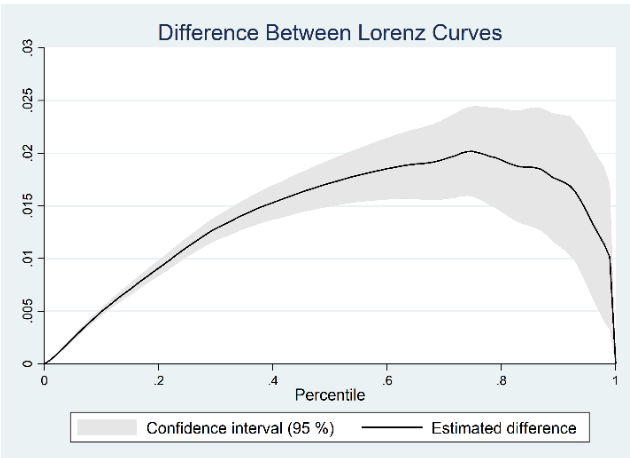


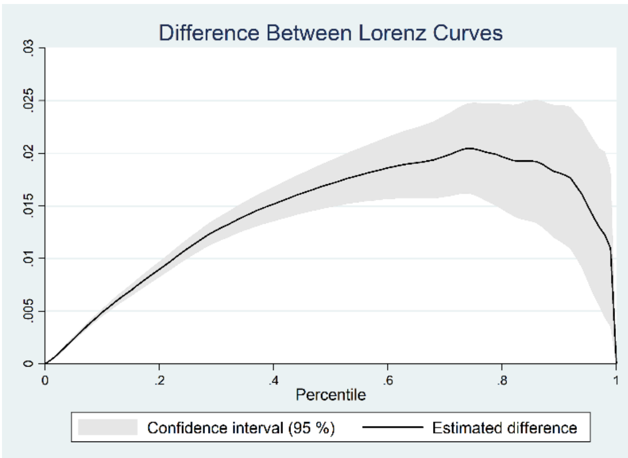
i. Market income per capita

i. Disposable income per capita

Notes: Positive values indicate that the uncorrected Lorenz curve dominates, and shows less inequality than the corrected Lorenz curve. Distributions account for sampling weights and household size.
Source: Own analysis of ENIGH 2012, CEQ database.

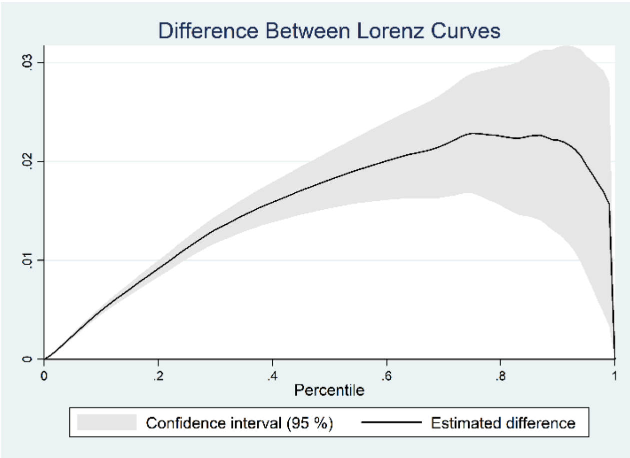
Figure A2-8. Lorenz curve: market vs. disposable income per capita, uncorrected versus corrected income distributions



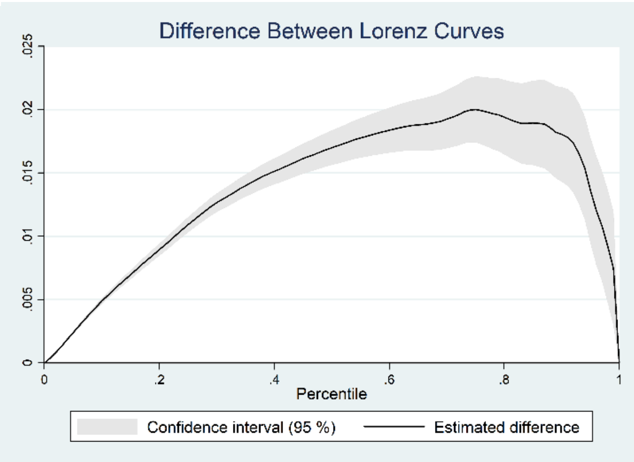


i. ENIGH sampling weights uncorrected for nonresponse

ii. Weights corrected for nonresponse (mdel 1)



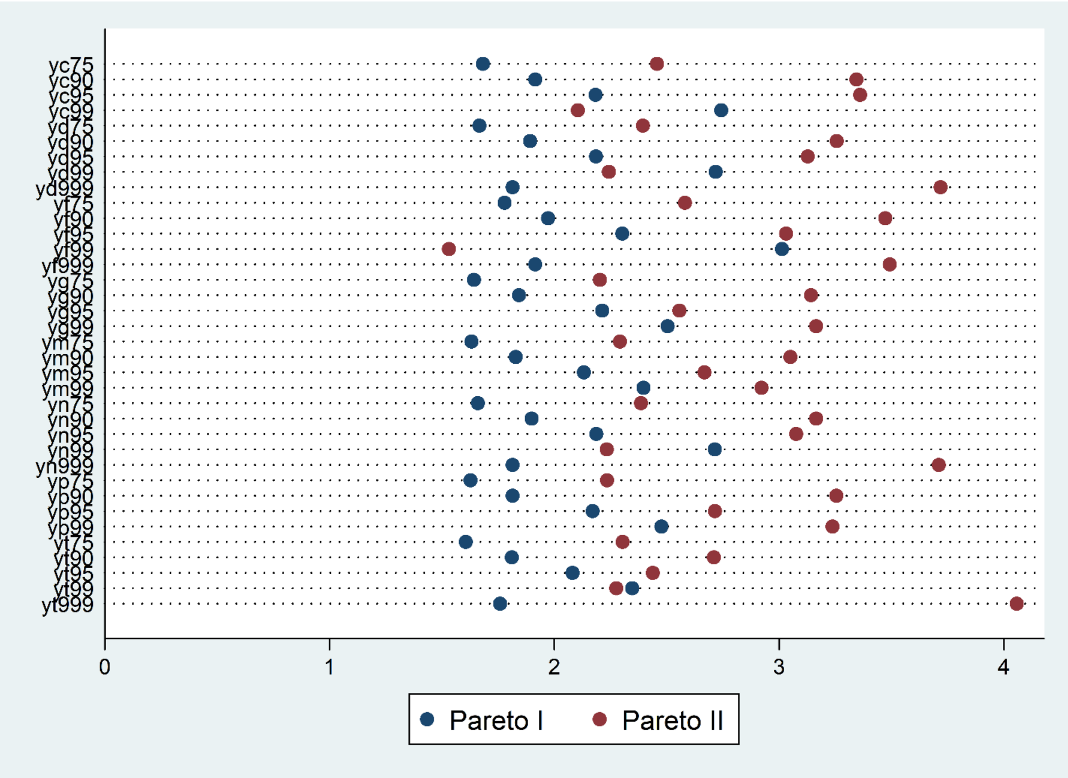
iii. Top 10% of incomes of each income concept replaced with Pareto (I) values



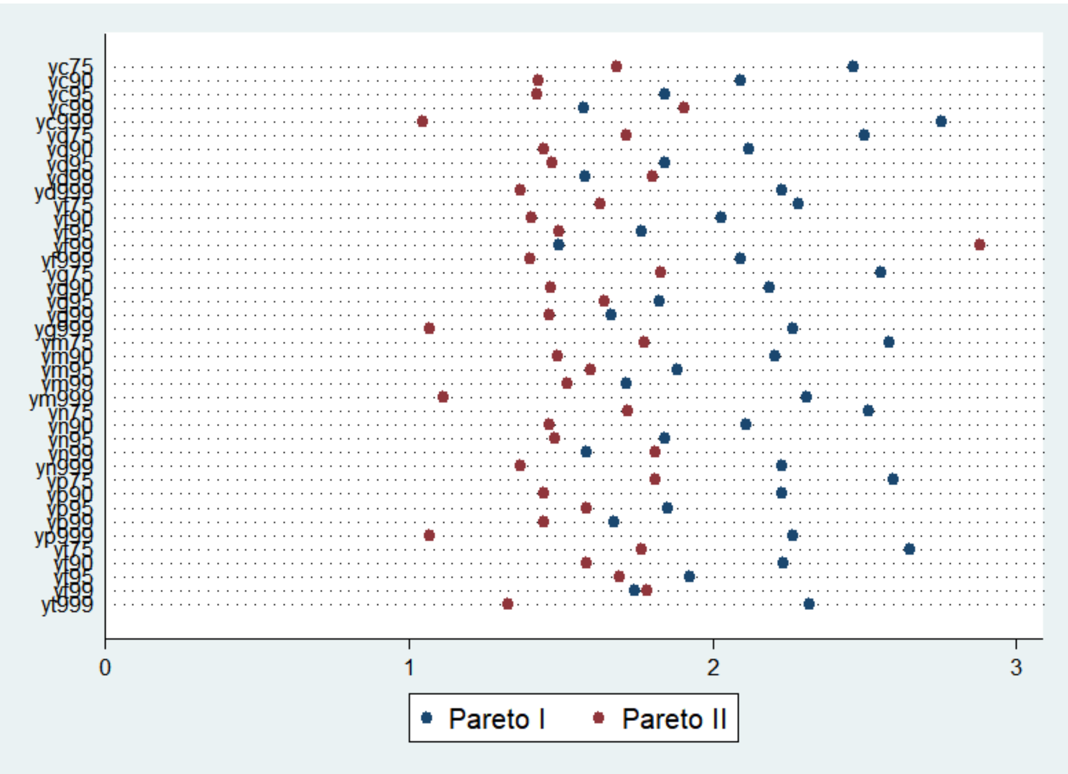
iv. Top 10% of net market incomes replaced with Pareto (I) values, other income concepts imputed by CEQ Method

Notes: Positive values indicate that disposable income Lorenz curve dominates, and shows less inequality than market income Lorenz curve. Distributions account for sampling weights and household size.
Source: Own analysis of ENIGH 2012, CEQ database.

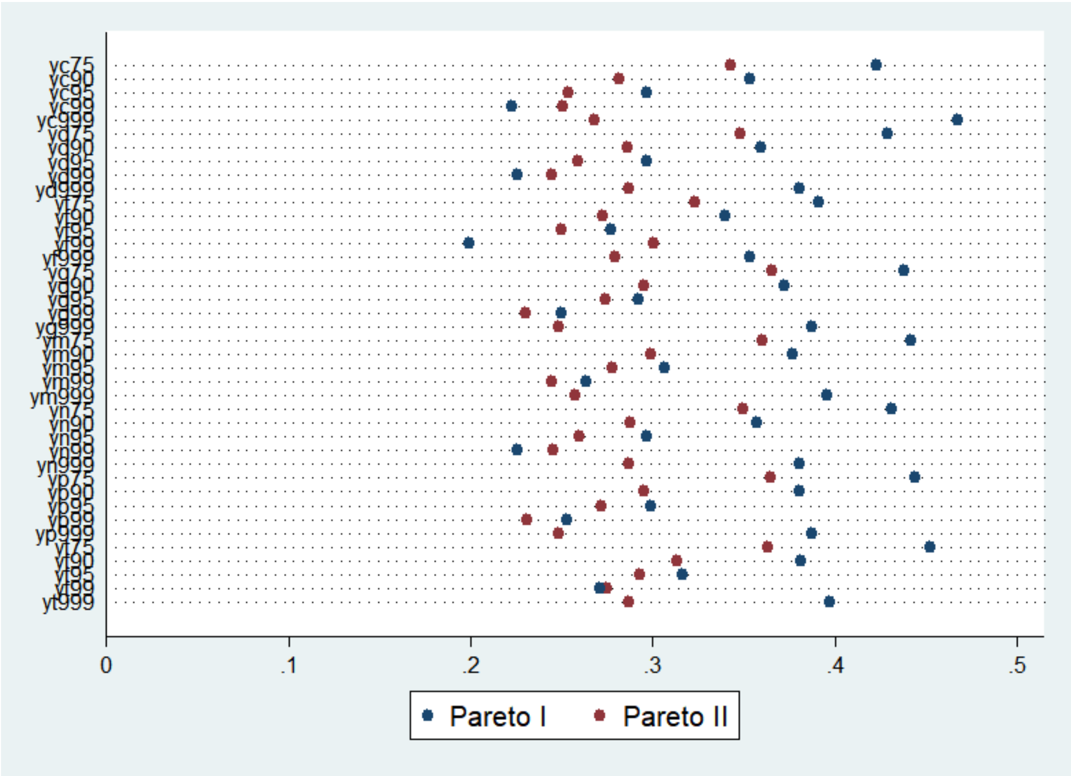
Figure A2-9. Comparison of Pareto I and Pareto II models, various income concepts and top income cutoffs



i. Pareto coefficient



ii. Inverted Pareto coefficient



iii. Gini coefficient

Notes: 'yp' is market income plus pensions, 'ym' market income, 'yt' taxable income, 'yg' gross income, 'yn' net market income, 'yd' disposable income, 'yc' consumable income, 'yf' final income.
Source: Own analysis of ENIGH 2012, CEQ database.

Appendix 3: Summary Statistics and Estimation Results for ENIGH 2010

Table A3-1. Income summary statistics

			Corrected for nonresponse by reweighting	
	Disposable income net of private & social security transfers per cap.	Disposable income per cap.	Disposable income net of private & social security transfers per cap.	Disposable income per cap.
99.9th %ile	545,865	589,642	453,354	483,936
99th %ile	189,013	213,799	173,058	193,413
95th %ile	88,022	101,324	81,463	93,939
90th %ile	61,335	70,977	57,014	65,694
75th %ile	36,295	41,974	33,983	39,407
Mean	30,060	36,050	27,873	33,633
Median	20,248	24,867	18,942	23,447
25th %ile	10,109	14,439	9,055	13,320
10th %ile	2,080	7,545	1,613	6,852
5th %ile	0	4,266	0	3,915
1st %ile	-187	907	-1	1,464
Std. dev.	45,109	48,409	40,476	43,516
Skewness	10.26	9.40	10.73	9.79
Kurtosis	268.15	223.42	296.90	245.85
Sample	27,593	27,593	27,385	27,385
Top 0.1% inc. share	2.70	2.35	2.60	2.27
0.1-1% inc. share	8.09	7.45	7.82	7.17
1-5% inc. share	15.93	15.13	15.72	14.87
5-10% inc. share	12.11	11.60	12.08	11.50
Gini (weighted data)	49.73 (0.39)	47.11 (0.36)	49.88 (0.36)	47.14 (0.33)
Gini (HH-size wted, no sampling wghts)	51.21 (0.31)	48.32 (0.29)	51.49 (0.28)	48.48 (0.26)
Gini (equal HH weights = unweighted)	54.10 (0.41)	50.45 (0.37)	54.16 (0.35)	50.36 (0.31)
Mean log dev. (GE0)	0.555 (0.010)	0.414 (0.007)	0.578 (0.010)	0.424 (0.006)
Theil index (GE1)	0.487 (0.012)	0.434 (0.010)	0.484 (0.010)	0.430 (0.009)
Half coef. of var.^{2} (GE2)	1.126 (0.070)	0.902 (0.052)	1.054 (0.060)	0.837 (0.044)

Notes: MXN0 incomes, 62 household observations, are omitted in computations of the Gini.
Columns 1–2 relies on a sample weighted using ENIGH sampling weights (uncorrected). Columns 3–4 statistics are based on non-response correction weights estimated in the logarithmic model of disposable household income (model 1 in table A3-3). Column 2 statistics exclude 62 household observations with disposable income of 0. The statistics remain comparable to those in column 1, which are extremely robust to this exclusion (changing by 0.01 at the most.) Ginis and standard errors are multiplied by 100 for clarity of presentation. Standard errors on Gini coefficients are jackknife estimates.
Source: Own analysis of ENIGH 2010, LIS database.

Table A3-2. ENIGH 2010: Sample sizes, non-response rates, and mean incomes by state

State code	State	Represented population	Fully interviewed households (individuals)	Type A non-responding HHs (%)	Mean dispos. HH income per cap.
01	Aguascalientes	1,192,473	329 (1,375)	10 (2.9%)	37,332
02	Baja California	3,174,663	504 (1,877)	3 (0.6%)	52,487
03	Baja California Sur	644,489	299 (1,093)	4 (1.3%)	46,743
04	Campeche	824,907	427 (1,620)	12 (2.7%)	37,142
05	Coahuila de Zarag.	2,759,794	564 (2,137)	22 (3.8%)	37,167
06	Colima	653,275	361 (1,267)	23 (6.0%)	45,933
07	Chiapas	4,820,282	2,805 (11,814)	71 (2.5%)	17,185
08	Chihuahua	3,414,766	836 (2,906)	15 (1.8%)	36,741
09	Distrito Federal	8,806,329	2,799 (9,681)	55 (1.9%)	61,666
10	Durango	1,637,472	497 (1,950)	47 (8.6%)	28,842
11	Guanajuato	5,507,127	1,901 (7,810)	38 (2.0%)	30,833
12	Guerrero	3,394,244	868 (3,677)	18 (2.0%)	23,401
13	Hidalgo	2,676,807	513 (2,191)	12 (2.3%)	23,833
14	Jalisco	7,378,707	595 (2,309)	31 (5.0%)	42,208
15	México	15,227,945	2,748 (11,244)	110 (3.8%)	36,887
16	Mich. de Ocampo	4,355,975	700 (2,643)	20 (2.8%)	27,148
17	Morelos	1,781,315	462 (1,744)	17 (3.5%)	35,519
18	Nayarit	1,089,518	408 (1,592)	2 (0.5%)	35,718
19	Nuevo León	4,659,638	411 (1,512)	21 (4.9%)	49,976
20	Oaxaca	3,808,423	1,051 (4,315)	19 (1.8%)	25,468
21	Puebla	5,790,569	662 (2,793)	41 (5.8%)	28,319
22	Querétaro	1,841,607	452 (1,771)	17 (3.6%)	41,294
23	Quintana Roo	1,341,524	350 (1,242)	4 (1.1%)	48,446
24	San Luis Potosí	2,588,544	539 (2,240)	16 (2.9%)	31,307
25	Sinaloa	2,772,382	458 (1,695)	5 (1.1%)	34,806
26	Sonora	2,669,362	649 (2,450)	7 (1.1%)	38,146
27	Tabasco	2,243,345	487 (1,876)	6 (1.2%)	28,160
28	Tamaulipas	3,283,331	558 (1,946)	40 (6.7%)	36,636
29	Tlaxcala	1,175,903	380 (1,551)	3 (0.8%)	28,086
30	Veracruz de Ign.	7,635,224	898 (3,285)	35 (3.8%)	28,373
31	Yucatán	1,956,933	2,719 (10,449)	166 (5.8%)	32,652
32	Zacatecas	1,493,710	425 (1,582)	2 (0.5%)	25,617
	Nationwide	112,572,638	27,655 (107,637)	892 (3.1%)	36,050

Note: Mean incomes account for sampling weights but are computed only among responding households, and may not be representative of underlying population.
Source: Instituto Nacional de Estadística y Geografía (INEGI, 2011), Encuesta Nacional de Ingresos y Gastos de los Hogares 2010: Diseño muestral; Own analysis of ENIGH 2010, LIS database.

Table A3-3. Estimation results for various univariate logistic models of response probability

Specification of $g(x_i,\theta)$	$\widehat{\theta_0}$ (s.e.)	$\widehat{\theta_1}$ (s.e.)	Sum of Squared Weighted Errors	Factor of Proportio -nality (σ^2)	AIC SIC	Gini (s.e.): Dispos. income per capita: Weighted data	HH-size wted, no sampling wghts	Equal HH weights = unweighted
Uncorrected						47.11	48.32	50.45
						(0.36)	(0.29)	(0.37)
1: $\theta_0+\theta_1$log(dispos. inc.)	-2.926	0.162	493.38	16.763	91.54	47.14	48.48	50.36
	(0.544)	(0.049)			88.92	(0.33)	(0.26)	(0.31)
2: $\theta_0+\theta_1$log(dispos. inc.	-2.352	0.105	372.22	16.033	82.52	46.65	47.92	49.96
net of transfers)	(0.414)	(0.039)			79.91	(0.35)	(0.28)	(0.35)
3: $\theta_0+\theta_1$log(disp. inc.pc)	-2.928	0.182	374.71	11.160	82.73	46.98	48.39	50.18
	(0.342)	(0.034)			80.12	(0.33)	(0.25)	(0.30)
4: $\theta_0+\theta_1$log(dispos. inc.)^2	-2.053	0.008	491.37	16.524	91.41	46.96	48.29	50.15
	(0.284)	(0.002)			88.79	(0.32)	(0.25)	(0.31)
5: $\theta_0+\theta_1$log(disp. inc.pc)^2	-2.069	0.010	368.43	10.678	82.19	46.63	48.02	49.78
	(0.178)	(0.002)			79.58	(0.32)	(0.25)	(0.28)
6: $\theta_0+\theta_1 10^{-6}$ dispos. inc.	-1.229	1.223	503.94	16.835	92.22	45.08	46.39	48.29
	(0.052)	(0.475)			89.60	(0.28)	(0.21)	(0.24)
7: $\theta_0+\theta_1 10^{-6}$ disp. inc.pc	-1.301	6.348	354.40	9.161	80.95	43.56	44.93	46.15
	(0.040)	(1.315)			78.34	(0.26)	(0.20)	(0.21)
8: $\theta_0+\theta_1 10^{-12}$dispos. inc.^{2}	-1.229	6.730	458.64	12.692	89.20	43.00	44.42	46.33
	(0.048)	(2.902)			86.59	(0.25)	(0.19)	(0.22)
9: $\theta_0+\theta_1(10^{-3}$dispos. inc.$)^{½}$	-1.479	0.039	491.04	15.258	91.39	45.92	47.24	49.06
	(0.115)	(0.012)			88.77	(0.29)	(0.22)	(0.26)
10: $\theta_0+\theta_1(10^{-3}$disp. inc.pc$)^{½}$	-1.543	0.084	357.27	9.093	81.21	45.03	46.40	47.88
	(0.080)	(0.016)			78.59	(0.29)	(0.22)	(0.23)

Note: Ginis and standard errors are multiplied by 100 for clarity of presentation. Standard errors on Gini coefficients are jackknife estimates.
Source: Own analysis of ENIGH 2010, LIS database.

Table A3-4. Estimation results for selected multivariate models of response probability

Specification of *g(x)*	$\widehat{\theta_0}$ (s.e.)	$\widehat{\theta_1}$ (s.e.)	$\widehat{\theta_2}$ (s.e.)	$\widehat{\theta_3}$ (s.e.)	$\widehat{\theta_4}$ (s.e.)	Sum of squared wghted. errors	Factor of propor-tionality (σ^2)	AIC SIC	Gini (s.e.): Market inc. pc.: Weighted data	HH-size wted, no sampling wghts	Equal HH weights= unwghtd
Uncorrected									47.11	48.32	50.45
									(0.36)	(0.29)	(0.37)
$\theta_0+\theta_1 10^{-6}$dispos. inc.	-1.241	0.304	5.602			458.51	12.768	91.19	43.17	44.58	46.48
$+\theta_2 10^{-12}$dispos. inc.^{2}	(0.149)	(3.802)	(14.77)					88.49	(0.24)	(0.20)	(0.22)
$\theta_0+\theta_1$log(dispos. inc.pc)	-3.343	0.206	-0.062	-1.596	5.735	306.20	7.117	82.27	46.90	48.36	50.20
$+\theta_2$urban/10$+\theta_3$age$+\theta_4$age^2	(1.010)	(0.059)	(0.151)	(6.598)	(10.102)			80.32	(0.33)	(0.25)	(0.29)
$\theta_0+\theta_1$log(dispos. inc.)	-2.959	0.103	0.181	-0.261	4.997	386.77	8.079	89.75	47.27	48.64	50.66
$+\theta_2$urban/10$+\theta_3$age$+\theta_4$age^2	(1.063)	(0.068)	(0.146)	(6.455)	(10.239)			87.79	(0.35)	(0.27)	(0.32)
$\theta_0+\theta_1$log(dispos. inc.pc)	0.474	-0.002	-0.001	-57.058	290.871	8.53	0.334	-32.32	45.56	47.09	49.01
$+\theta_2$sch.yrs$+\theta_3$hsize$+\theta_4$hsize^2	(0.148)	(0.013)	(0.004)	(4.605)	(42.930)			-34.28	(0.39)	(0.31)	(0.30)
$\theta_0+\theta_1$log(dispos. inc.pc) +	0.106	-0.039	0.012	-21.470	-0.203	49.61	1.440	24.03	46.15	48.00	50.04
$\theta_2$sch.yrs$+\theta_3$hsize$+\theta_4$male	(0.295)	(0.028)	(0.008)	(1.157)	(0.137)			22.08	(0.47)	(0.37)	(0.33)
$\theta_0+\theta_1$log(dispos. inc.pc)	-2.006	0.111	0.996	-0.389		318.19	7.664	81.50	45.16	46.71	48.58
$+\theta_2$postsec.attend$+\theta_3$male	(0.461)	(0.038)	(0.702)	(0.320)				79.05	(0.32)	(0.25)	(0.29)
$\theta_0+\theta_1$log(dispos. inc.pc)	-0.912	0.068	1.403	-1.159		288.16	8.507	78.33	44.29	45.91	47.92
$+\theta_2$postsec.attend$+\theta_3$empl.	(0.957)	(0.043)	(0.925)	(0.783)				75.87	(0.31)	(0.25)	(0.33)

Note: Standard errors on Gini coefficients are jackknife estimates. Variables are normalized: age=(years-12)/100; HH size=(#-1)/100. Measures of fit are not entirely comparable across models with different controls, because of different sample sizes.
Source: Own analysis of ENIGH 2010, LIS database.

Table A3-5. Replacement of top incomes with Pareto I distribution: uncorrected vs. corrected Ginis

Cutoff percentile	Disposable income net of transfers per capita	Disposable income per capita
Nonparametric Gini among bottom incomes		
Top 25%	33.17 (0.27)	29.70 (0.22)
Top 10%	37.32(0.24)	34.56 (0.21)
Top 5%	40.34 (0.24)	37.80 (0.22)
Top 1%	45.32 (0.28)	42.92 (0.26)
Top 0.1%	48.52 (0.34)	46.01 (0.31)
Nonparametric Gini among top incomes		
Top 25%	32.04 (0.63)	31.20 (0.58)
Top 10%	29.53 (0.85)	28.42 (0.78)
Top 5%	28.36 (1.06)	27.00 (0.97)
Top 1%	25.85 (1.53)	24.26 (1.44)
Top 0.1%	16.34 (2.51)	15.34 (2.27)
Pareto (type I) coefficient		
Top 25%	1.824 (0.027)	1.851 (0.027)
Top 10%	2.017 (0.044)	2.072 (0.044)
Top 5%	2.146 (0.067)	2.221 (0.066)
Top 1%	2.330 (0.150)	2.476 (0.156)
Top 0.1%	2.870 (0.347)	3.249 (0.427)
Parametric Gini among top incomes		
Top 25%	37.77	37.02
Top 10%	32.97	21.80
Top 5%	30.38	29.06
Top 1%	27.32	25.31
Top 0.1%	21.10	18.19
Semiparametric Gini		
Top 25%	**53.75**	**49.97**
Top 10%	**51.37**	**48.13**
Top 5%	**50.53**	**47.57**
Top 1%	**49.96**	**47.23**
Top 0.1%	**49.80**	**47.14**
Uncorrected	49.73	47.11
Gini correction		
Top 25%	**+4.02**	**+2.86**
Top 10%	**+1.64**	**+1.02**
Top 5%	**+0.80**	**+0.46**
Top 1%	**+0.23**	**+0.12**
Top 0.1%	**+0.07**	**+0.03**

Source: Own analysis of ENIGH 2010, LIS database.

Table A3-6. Summary results of correction methods: corrected Ginis for all income concepts

	Disposable income net of transfers per capita	+ Cash-like private & social security transfers (after-tax)	Disposable income per cap.
	Value **+%pt. correct.**	%pt. Δ	Value **+%pt. correct.**
Gini coefficient			
Uncorrected	49.7	–2.6	47.1
Correction by reweighting (models in table A3-3)			
Minimum	45.9 **–3.8**	–2.9	43.0 **–4.1**
Mean	48.6 **–1.1**	–2.9	45.7 **–1.4**
Median	49.1 **–0.6**	–2.8	46.3 **–0.8**
Maximum	50.6 **+0.9**	–3.5	47.1 **+0.0**
Correction by Pareto (type I) replacing of own income concept			
Minimum	49.8 **+0.1**	–2.7	47.1 **+0.0**
Mean	51.1 **+1.4**	–3.1	48.0 **+0.9**
Median	50.5 **+0.8**	–2.9	47.6 **+0.5**
Maximum	53.8 **+4.1**	–3.8	50.0 **+2.9**
Top 10 percent income share			
Uncorrected	38.8	–2.3	36.5
Correction by reweighting (models in table A3-3)			
Minimum	34.5 **–4.3**	–2.3	32.2 **–4.3**
Mean	37.0 **–1.8**	–2.3	34.7 **–1.8**
Median	37.5 **–1.3**	–2.4	35.1 **–1.4**
Maximum	38.2 **–0.6**	–2.4	35.8 **–0.7**
Correction by Pareto (type I) replacing of own income concept			
Minimum	38.9 **+0.1**	–2.4	36.5 **+0.0**
Mean	39.9 **+1.1**	–2.3	37.6 **+1.1**
Median	39.3 **+0.5**	–2.3	37.0 **+0.5**
Maximum	42.5 **+3.7**	–2.2	40.3 **+3.8**
Top 1 percent income share			
Uncorrected	10.8	–1.0	9.8
Correction by reweighting (models in table A3-3)			
Minimum	8.4 **–2.4**	–0.7	7.7 **–2.1**
Mean	9.7 **–1.1**	–0.9	8.8 **–1.0**
Median	9.9 **–0.9**	–0.9	9.0 **–0.8**
Maximum	10.4 **–0.4**	–1.0	9.4 **–0.4**
Correction by Pareto (type I) replacing of own income concept			
Minimum	10.9 **+0.1**	–1.1	9.8 **+0.0**
Mean	11.9 **+1.1**	–1.1	10.8 **+1.0**
Median	11.7 **+0.9**	–1.0	10.7 **+0.9**
Maximum	13.5 **+2.7**	–1.0	12.5 **+2.7**

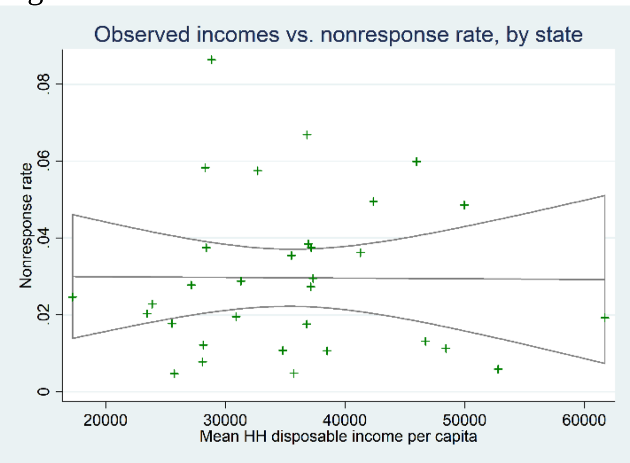
Notes: Pc.pt. differences from uncorrected Ginis in **bold**. These Ginis and differences in them arise from 'Gini (HH-size & sampling weighted data)' in tables A3-1, and 'Semiparametric Gini' in table A3-5. Ginis and percentage point changes are multiplied by 100 for clarity of presentation.
Source: Own analysis of ENIGH 2010, LIS database.

Table A3-7. Generalized Pareto (type II) results, income per capita

Cutoff percentile	Disposable income net of transfers per capita	Disposable income per capita
Pareto (type II) shape coefficient ξ		
Top 25%	0.403 (.020)	0.390 (.019)
Top 10%	0.401 (.030)	0.388 (.029)
Top 5%	0.430 (.043)	0.403 (.040)
Top 1%	0.439 (.082)	0.414 (.074)
Top 0.1%	0.048 (.101)	0.060 (.096)
Pareto (type II) scale coefficient, log(σ)		
Top 25%	10.039 (.029)	10.175 (.029)
Top 10%	10.417 (.043)	10.534 (.042)
Top 5%	10.658 (.063)	10.777 (.058)
Top 1%	11.294 (.134)	11.359 (.123)
Top 0.1%	12.449 (.193)	12.415 (.196)
1/ξ		
Top 25%	2.481	2.566
Top 10%	2.495	2.580
Top 5%	2.327	2.481
Top 1%	2.278	2.415
Top 0.1%	20.867	16.637
Inverted Pareto coefficient		
Top 25%	1.675	1.638
Top 10%	1.669	1.633
Top 5%	1.753	1.675
Top 1%	1.782	1.707
Top 0.1%	1.050	1.064
Gini among top incomes		
Top 25%	32.17	31.42
Top 10%	29.78	28.76
Top 5%	29.21	27.67
Top 1%	27.61	25.63
Top 0.1%	16.86	15.87
Log pseudo-likelihood (LL/10^6)		
Top 25%	-321.00	-325.00
Top 10%	-133.00	-134.00
Top 5%	-67.90	-68.40
Top 1%	-14.30	-14.30
Top 0.1%	-1.497	-1.488

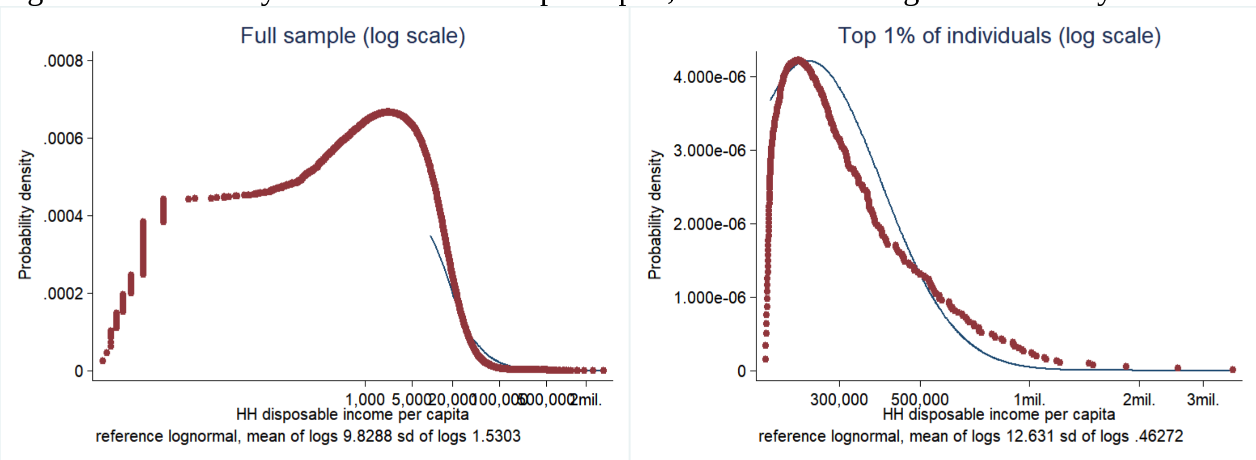
Notes: individual sampling-weighted sample.
Source: Own analysis of ENIGH 2010, LIS database.

Figure A3-1. Mean observed market income among respondents, and nonresponse rate, by state



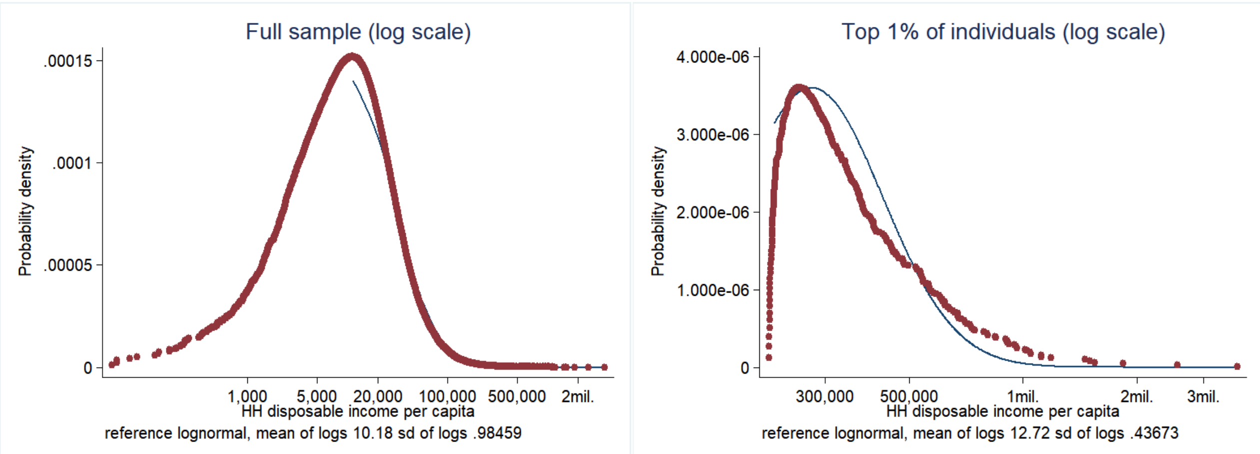
Notes: 95% confidence interval around linear fitted line is shown.
Source: Own analysis of ENIGH 2010, LIS database.

Figure A3-2. Density function of income per capita, with reference lognormal density



i. Disposable income net of transfers, full sample

ii. Disposable income net of transfers, top 1%

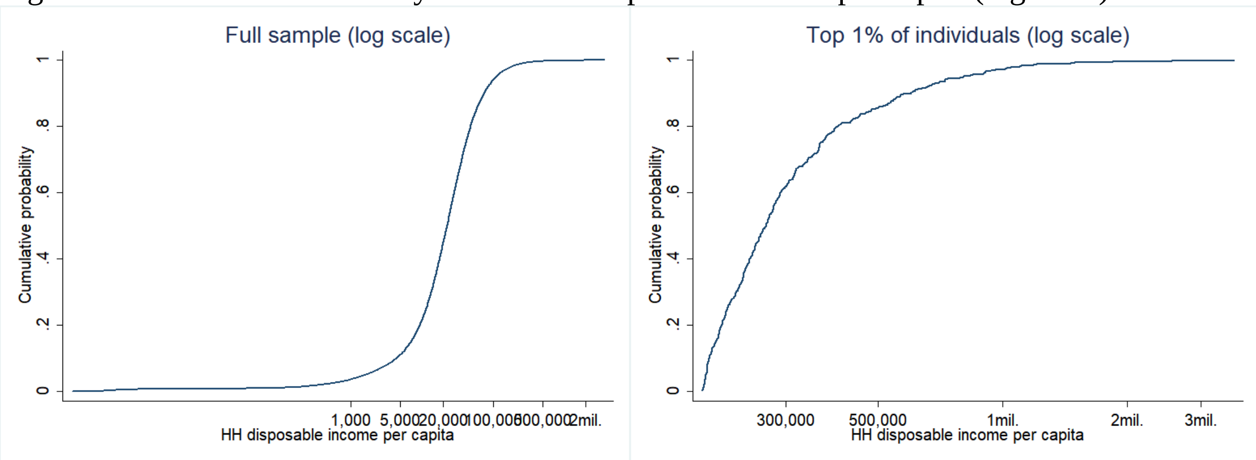


iii. Disposable income, full sample

iv. Disposable income, top 1%

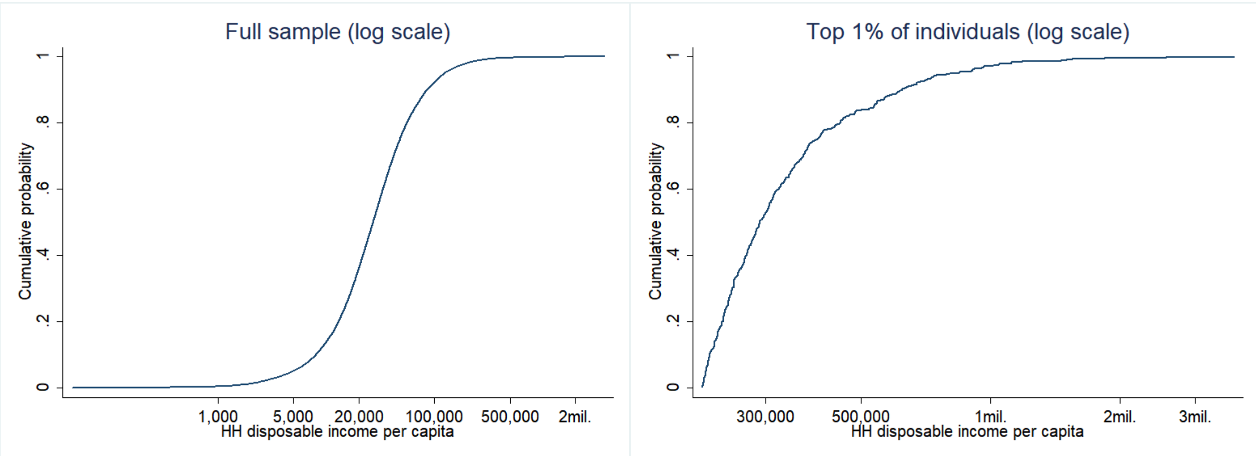
Notes: Non-positive incomes omitted.
Source: Own analysis of ENIGH 2010, LIS database.

Figure A3-3. Cumulative density function of disposable income per capita (log scale)



i. Disposable income net of transfers, full sample

ii. Disposable income net of transfers, top 1%

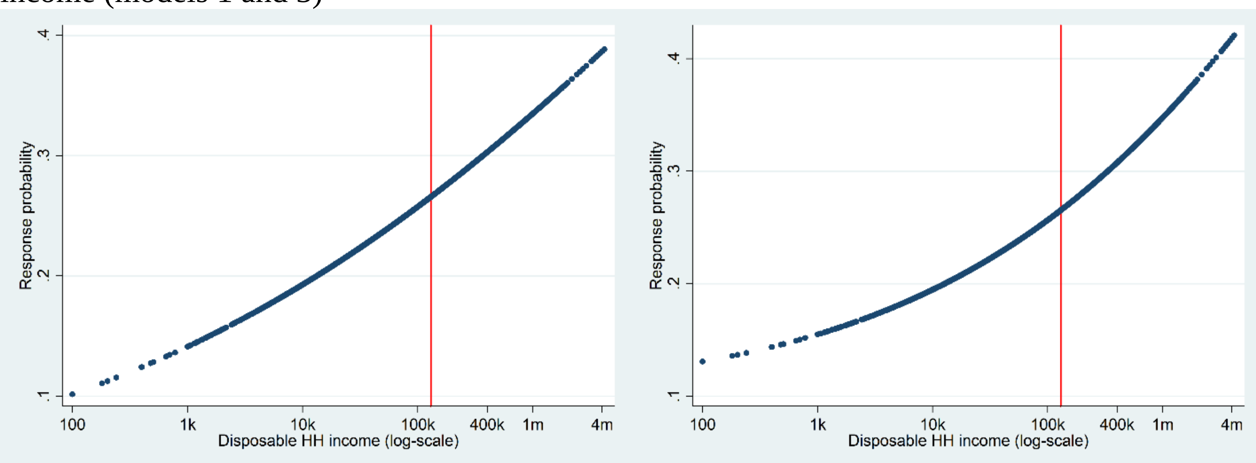


iii. Disposable income, full sample

iv. Disposable income, top 1%

Source: Own analysis of ENIGH 2010, LIS database.

Figure A3-4. Unit response probability by disposable income, logarithmic model of disposable income (models 1 and 3)



i. Logarithmic model of disposable income (model 1) $\theta_0+\theta_1\log(\text{disposable inc.})$

ii. Quadratic logarithmic model of disposable income (Model 3) $\theta_0+\theta_1\log(\text{disposable inc.})^2$

Notes: Red line shows mean disposable HH income in the corrected income distribution: i) 128,482 and ii) 127,674.

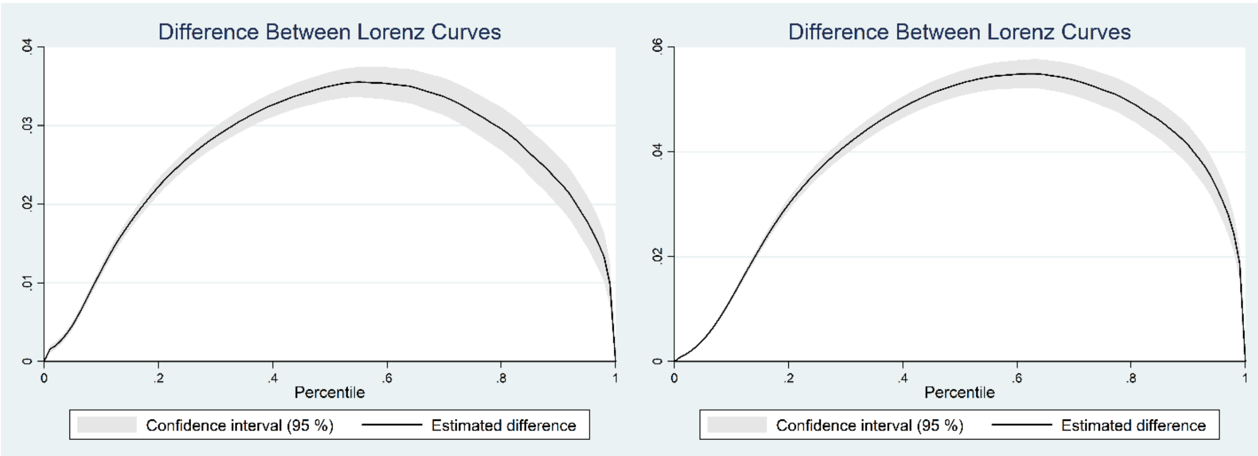
Source: Own analysis of ENIGH 2010, LIS database.

Figure A3-5. Lorenz curve: income per capita, uncorrected versus unit-nonresponse corrected weights (model 1)

i. Disposable income net of transfers, per capita ii. Disposable income per capita

Notes: Positive values indicate that the Lorenz curve uncorrected for unit-nonresponse dominates and shows less inequality than the corrected Lorenz curve, and vice versa. Distributions account for sampling weights and household size.
Source: Own analysis of ENIGH 2010, LIS database.

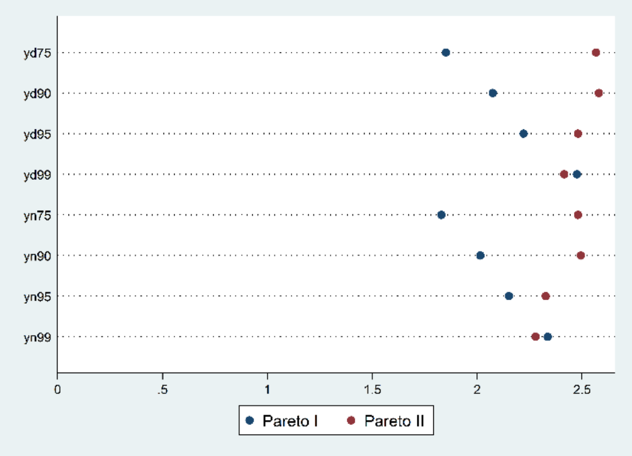
Figure A3-6. Lorenz curve: disposable income vs. disposable income net of transfers, uncorrected versus corrected income distributions



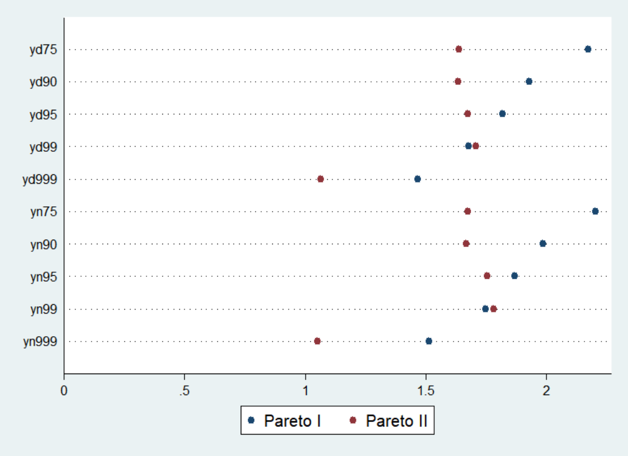
i. ENIGH sampling weights uncorrected for nonresponse ii. Weights corrected for unit nonresponse

Notes: Positive values indicate that disposable income Lorenz curve dominates, and shows less inequality than Lorenz curve of disposable income net of transfers. Distributions account for sampling weights and household size.
Source: Own analysis of ENIGH 2010, LIS database.

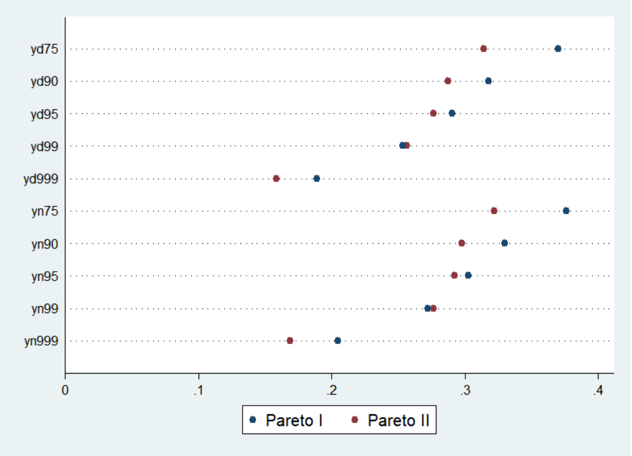
Figure A3-7. Comparison of Pareto I and Pareto II models, various income concepts and top income cutoffs



i. Pareto coefficient



ii. Inverted Pareto coefficient



iii. Gini coefficient

Notes: ‘yn’ is disposable household income net of transfers, ‘yd’ is disposable household income, both per capita.
Source: Own analysis of ENIGH 2010, LIS database.