ifo WORKING PAPERS

301
2019

May 2019

Lower and Upper Bounds of Inequality of Opportunity in Emerging Economies

Paul Hufe, Andreas Peichl, Daniel Weishaar

ifo INSTITUTE

Leibniz Institute for Economic Research
at the University of Munich

Imprint:

ifo Working Papers
Publisher and distributor: ifo Institute – Leibniz Institute for Economic Research at the University of Munich
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49(0)89 9224 0, Telefax +49(0)89 985369, email ifo@ifo.de
www.ifo.de

An electronic version of the paper may be downloaded from the ifo website:
www.ifo.de

ifo Working Paper No. 301

Lower and Upper Bounds of Inequality of Opportunity in Emerging Economies*

Abstract

Equality of opportunity is an important normative ideal that concerns politicians and the larger public alike. In spite of its wide acceptance, current estimation approaches in the literature suffer from severe data restrictions that lead to biased estimates of inequality of opportunity. These shortcomings are particularly pronounced for emerging economies in which comprehensive household survey data often is unavailable. In this paper, we address these issues by estimating lower and upper bounds of inequality of opportunity for a set of emerging economies. Thereby, we address recent critiques that worry about the prevalence of lower bound estimates and the ensuing scope for downplaying the normative significance of inequality.

JEL Code: D31, D63, I32

Keywords: Equality of opportunity; inequality; emerging economies

Paul Hufe
ifo Institute – Leibniz Institute for Economic Research
at the University of Munich,
University of Munich
Poschingerstr. 5,
81679 Munich, Germany
hufe@ifo.de

Andreas Peichl**
ifo Institute – Leibniz Institute for Economic Research
at the University of Munich,
University of Munich, IZA, CESifo
Poschingerstr. 5,
81679 Munich, Germany
peichl@ifo.de

Daniel Weishaar
University of Munich
daniel.weishaar@campus.lmu.de

* We are grateful to F. Andreoli, P. Brunori, F. Ferreira, R. Kanbur, V. Peragine, J. Roemer, M. Stoeckli and A. Trannoy as well as participants at seminars and conferences in Canazei, Luxembourg and Munich for helpful suggestions on earlier drafts. We gratefully acknowledge funding from Deutsche Forschungsgemeinschaft (DFG) through NORFACE project \IMCHILD: The impact of childhood circumstances on individual outcomes over the life-course" (PE 1675/5-1).

** Corresponding author.

1 Introduction

Inequality has recently received increasing prominence in debates among economists, in policy circles and among the wider public.[1] While some inequality is tolerable and might even be desirable, substantial differences based on pre-determined characteristics such as ethnicity, sex and race run counter to most theories of justice that view such inequalities as inherently unfair. This issue is especially important in developing and emerging countries. Yet, previous attempts of measuring inequality of opportunity (IOp) in these countries are encumbered by data constraints that usually lead to an underestimation of IOp. To overcome this issue, this is the first paper to provide both lower and upper bound estimates of IOp in emerging economies.

IOp is an ideal of distributive justice that garners wide-spread support in the general public (Faravelli 2007; Cappelen et al. 2007; Alesina et al. 2018). Opportunity egalitarians distinguish ethically justifiable (fair) inequalities from unjustifiable (unfair) inequalities by reference to the concepts of circumstances and effort.[2] Circumstances are defined as all factors affecting an outcome which are not under the control of the individual, while effort variables are (at least partially) under control. Whereas inequalities based on exogenous circumstances are considered unfair, all inequalities that are the result of individual effort exertion are deemed fair causes of inequality. Hence, while for example outcome differences due to gender, parental background or the birthplace of an individual are considered as unfair inequalities, working hours and educational decisions are (partially) under the control of an individual and the ensuing outcome differences are therefore considered as (partially) fair.

Empirical estimates of IOp face two major empirical challenges. First, many circumstance factors are unobserved in the data which leads to an underestimation of their aggregate impact on individual life outcomes (Ferreira and Gignoux 2011; Balcázar 2015; Hufe et al. 2017). Second, if the ratio between the number of parameters to be estimated and the available degrees of freedom becomes large, the ensuing noise in the parameter estimates will artificially inflate the impact of observed circumstances on individual life outcomes (Brunori et al. 2018a; Brunori et al. 2018b). The extant literature refers to these phenomena as upward and downward biases in IOp estimates. In particular the first concern has led researchers to question the usefulness of IOp estimates for policy evaluation (Kanbur and Wagstaff 2016; Andreoli et al. 2019).[3] The difference between the true and the lower bound estimate may be particularly pronounced in emerging economies due to a lack of comprehensive data that enable the researcher to construct finely-grained partitions of the population into circumstance types.

To address this issue, we calculate lower and upper bounds of IOp for a set of twelve emerging economies for which we can draw on longitudinal household surveys (Argentina, China, Chile, Ethiopia, Indonesia, Malawi, Mexico, Peru, Russia, South Africa, Thailand and Tanzania). For

[1] Much of the literature has focused on income inequality, but there are many dimensions of inequality. See Klasen et al. (2018) for a recent overview.

[2] This separation was formulated in the works of Roemer (1993, 1998), Van de gaer (1993), and Fleurbaey (1995). For recent reviews of the literature see Roemer and Trannoy (2015) and Ramos and Van de gaer (2016).

[3] For illustrative purposes, assume that the true estimate for IOp in country A amounts to 50%, i.e. half of total inequality can be explained by individual differences in circumstances. However, due to partial observability of circumstances, the researcher computes a lower bound estimate of 10%. Kanbur and Wagstaff (2016) worry that policymakers take the lower bound estimates of IOp as a reference point and consequently downplay the need for inequality-reducing policy interventions.

each country we calculate standard lower bound measures of IOp. In addition, we leverage the panel dimension of the data to calculate upper bound estimates of IOp based on the fixed effect estimator proposed in Niehues and Peichl (2014). As a consequence, we can determine reasonable bounds for IOp in these countries and thus address the concern of misleading reference points for policymakers. To the best of our knowledge, this is the first paper to conduct such a bounding exercise for emerging countries. Moreover, by analyzing a large set of emerging economies with broad geographical coverage and in different stages of development, we also contribute to the emerging literature on IOp in developing countries.[4]

2 Conceptual Framework for Estimating IOp

Important life outcomes are determined by an extensive vector of personal characteristics including cognitive ability, gender, parental background, health, socio-emotional skills, educational attainment, occupational choices and working hours. All of these factors can be subsumed by a binary classification. If they are completely beyond the realm of individual control, they are called circumstances. To the contrary, if they can be at least partially controlled by individuals, they are called effort. The more circumstances explain the distribution of outcomes, the stronger the violation of the opportunity egalitarian ideal.[5]

Following Ferreira and Gignoux (2011) this idea can be formulated as follows. We assume a finite population indexed by $i \in \{1, ..., N\}$, where each individual in period s is characterized by the tuple $\{y_{is}, \mathbf{C_i}, \mathbf{e_{is}}\}$. y_{is} constitutes the outcome of interest, $\mathbf{C_i}$ the vector of time-invariant and individual-specific circumstances, and $\mathbf{e_{is}}$ period-specific effort exertion. We can construct a partition of disjunct types $\Pi = \{T_1, ..., T_P\}$, such that all members of a type are homogeneous in circumstances $\mathbf{C_i}$. Equality of opportunity is achieved if mean advantage levels are equalized across types, i.e. if $\mu_k(y) = \mu_l(y) \ \forall \ l, k \mid T_k, T_l \in \Pi$.[6] Computing inequality in the distribution of type-means, $\mathrm{I}(\mu_{is}^k)$, now gives a scalar measure of IOp that reflects differences due to circumstances but is invariant to differential effort exertion within circumstance types.

Lower Bound Estimation. To the extent that we do not observe all circumstances, we only can construct type partition $\Pi' = \{T_1, ..., T_Q\}$ with $Q < P$. Ferreira and Gignoux (2011) show that the resulting estimate of IOp is weakly smaller than the true estimate, i.e. $\mathrm{I}(\mu_{is}^k)' \leq \mathrm{I}(\mu_{is}^k)$. In line with the extant literature, we construct the counterfactual distribution of type means in a two-step procedure. Assuming a log-linear relationship between the outcome of interest and circumstances, and allocating the correlation between $\mathbf{C_i}$ and $\mathbf{e_{is}}$ to the unfair part of inequality

[4]See Brunori et al. (2015) and Alesina et al. (2019) for work on Africa, Ferreira and Gignoux (2011) for work on Latin America, and Son (2013) for work on Asia. Furthermore, the Equal Chances project provides an internationally comparable database on IOp and intergenerational mobility.

[5]Note that we follow the normative ideal of Roemer (1998) who proposes that outcome differences due to a correlation between circumstances and effort constitute a violation of equality of opportunity. For example, although occupational choices may be partially under the control of individuals, income differences following from occupational differences across males and females violate the ideal of equal of opportunities This is a particular normative stance that can be easily relaxed. See Jusot et al. (2013) for a discussion.

[6]The literature refers to this as the ex-ante utilitarian approach (Ramos and Van de gaer 2016). For the sake of brevity and in view of its prevalence in empirical applications we forego a discussion of alternative approaches.

we estimate

$$\ln y_{is} = \alpha + \boldsymbol{\beta} * \mathbf{C_i} + \epsilon_{is}, \tag{1}$$

and then use the vector of estimated parameters $\hat{\boldsymbol{\beta}}$ to construct the distribution of type means:

$$\tilde{\mu}_{is}^{LB} = \exp\left\{\hat{\alpha} + \hat{\boldsymbol{\beta}} * \mathbf{C_i} + \frac{\sigma^2}{2}\right\}.^{7} \tag{2}$$

Recent contributions have argued that this lower bound estimate may be upward biased due to sampling variance in the distribution of type means (Brunori et al. 2018a; Brunori et al. 2018b). The sampling variance increases as the number of parameters increases in comparison to the available degrees of freedom. We therefore also provide lower bound estimates based on cross-validated lasso estimations (Tibshirani 2011), which select the relevant circumstance parameters in a way that minimizes the out-of-sample variance of the estimate.

Upper Bound Estimation. Since unobserved circumstances are time invariant by definition, they can be conceived as unobserved individual heterogeneity that is captured by individual fixed effects. We construct the distribution of type means in a three-step procedure. In a first step, using observations from all periods $t \neq s$, we estimate

$$\ln y_{it} = c_i + u_t + \epsilon_{it}, \tag{3}$$

where c_i represents the individual fixed effect. u_t captures year-specific effects such as macroeconomic conditions affecting all individuals equally. Next, we regress the individual outcome in period s on the obtained individual-specific effects

$$\ln y_{is} = \Psi * \hat{c}_i + \epsilon_{is}, \tag{4}$$

and subsequently use the vector of parameters $\hat{\Psi}$ to construct the distribution of type means:

$$\tilde{\mu}_{is}^{UB} = \exp\left\{\hat{\Psi} * \hat{c}_i + \frac{\sigma^2}{2}\right\} \tag{5}$$

Note that this estimator would yield the true estimate of IOp if c_i captured time-invariant circumstances only. However, to the extent that effort exertion is time-invariant (e.g. long-term motivation, ambition), these effort components are absorbed by the individual fixed effect. Therefore, this estimator delivers an upper bound because it refers to the maximum possible amount of variation that can be explained by circumstances.

3 Data

We estimate IOp in income and consumption expenditures for twelve emerging economies in different geographical areas of the world ranging from Africa (Ethiopia, Malawi, South Africa, Tanzania), Central and South America (Argentina, Chile, Mexico, Peru), Europe and Central

[7] In the estimation of predicted values, $\frac{\sigma^2}{2}$ corrects for differences in the marginal impact of circumstances due to the log-transformation (Blackburn 2007).

Asia (Russia), to East and South-East Asia (China, Indonesia, Thailand). The sample selection is guided by the availability of household panel data surveys with a sufficient number of observations in the longitudinal dimension. According to the World Bank (2018) classification, the country sample covers low income economies (Ethiopia, Malawi, Tanzania), lower-middle income economies (Indonesia), upper-middle income economies (China, Mexico, Peru, Russia, South Africa, Thailand) and high-income economies (Argentina, Chile), and avails data in the time range from 1988 to 2017.[8] For each country, we provide estimates for the most recent wave.[9] The earliest estimates refer to the year 2009 (Chile, Ethiopia, Mexico) while the latest estimates are for the year 2017 (Russia, South Africa, Thailand). Table A.1 provides an overview of the underlying data sources.

We consider two possible outcome dimensions of interest. First, we calculate IOp in individual incomes, where income is measured by gross or net incomes. Gross incomes are defined as market incomes before taxation or transfers. Net incomes are defined as total net income after taxes and transfers.[10] To account for resource sharing at the household level, we also calculate IOp in incomes at the household level. This is particularly relevant, given that female participation in formal labor markets often is low in emerging economies (Cubas 2016). To account for differential household composition, we calculate equivalized household incomes based on the modified OECD equivalence scale. Second, to derive an even more direct measure of IOp in material well-being, we also consider household consumption expenditures as an outcome of interest. As in the case of household incomes, we deflate expenditures by the modified OECD equivalence scale. Furthermore, all income and expenditure values are inflation-adjusted via Consumer Price Index (CPI) data to the country-specific year of analysis (Federal Reserve Bank of St. Louis 2018). As shown in Table A.1, data availability imposes restrictions on cross-country comparability since not all outcomes are available for all countries in our sample. However, for each outcome variable, we have information on at least 9 out of 12 countries.

In addition to the availability of the outcome variables, the considered data sets vary in the availability of circumstance indicators. In order to compare IOp across countries based on the same circumstances, we start with the smallest common denominator of circumstances. Unfortunately, we have only two variables in all datasets for all countries in our sample: year of birth and gender. In column 3 of Table 1 we display additional circumstance variables available for each country. These include birthplace, education of parents, information on ethnic, religious or linguistic background. Further circumstance variables include parents' endowment with wealth and land, geographic characteristics of the birthplace as well as individual's body height.

To ensure the consistency of intra-country comparisons, we only retain those units of observation for which we observe all circumstance variables and a minimum number of observations for each outcome variable available in the particular country data set. Specifically, we focus on individuals with positive outcomes in at least three periods of observation. Furthermore, we

[8] Note that throughout this paper, we refer to outcome (income/expenditure) and not survey years.

[9] An exception to this rule is Peru, for which it was not possible to construct a representative sample for the most recent available year (2011). For Peru, we thus use 2010 as the year of interest in our baseline specification. In addition, note that we also check for the robustness of the results regarding the year of interest in Section 4.

[10] In some countries market income comprises labor market earnings, only. These countries are indicated in Table A.1 accordingly.

focus on individuals aged between 25 and 55, i.e. the prime working age defined by the OECD (2018). On average, these restrictions imply the utilization of 3.0 (Argentina) to 14.6 (Thailand) years per individual (Column 7 of Table 1).[11]

Once we have constructed the respective distribution of type means, we follow the literature and use the mean log deviation (MLD) to provide scalar measures of IOp.

Table 1: Baseline Specification, Lower & Upper Bounds, by Country

	Circumstances		Fixed Effect Estimator			
	LB1	LB2 & LB3	Start	End	Min. Year	∅ Years
Argentina	gender, year of birth	birthplace	2013	2014	3	3.00
Chile	gender, year of birth	birthplace, education of father/mother, ethnicity, labor force status of father/mother, chronic disease	2006	2008	4	4.00
China	gender, year of birth	ethnicity	1988	2010	3	3.95
Ethiopia	gender, year of birth	education of father/mother, ethnicity, religion	1994	2004	3	4.63
Indonesia	gender, year of birth	education of father/mother, ethnicity, religion, language	1992	2006	3	3.59
Malawi	gender, year of birth	education of father/mother, religion	2004	2008	3	3.54
Mexico	gender, year of birth	language	1999	2004	5	6.43
Peru	gender, year of birth	birthplace, language, chronic disease	1998	2011	3	3.68
Russia	gender, year of birth	birthplace, urbanity of birthplace, education of father/mother, labor force status of father/mother, height	1994	2016	5	10.16
South Africa	gender, year of birth	birthplace, education of father/mother, ethnicity	2008	2015	4	4.48
Tanzania	gender, year of birth	birthplace, ethnicity, religion, height	1991	2004	3	4.71
Thailand	gender, year of birth	education of father/mother, wealth of parents, land of parents	1997	2016	3	14.64

Notes: Column 2 describes the internationally comparable set of circumstance characteristics that is used to estimate LB1. Column 3 describes the country-specific circumstances that are added to LB1 to estimate LB2 and LB3. All year specifications refer to the year in which the outcome (income/expenditure) was realized. The last two columns describe the distribution of data points per unit of observation. The figures include the year of analysis which, however, is not used in the estimation of the unit fixed effect.
Source: Own calculations based on the panel survey data described in Table A.1.

[11] Note that the number of waves used to calculate the fixed effect according to equation 3 differs both across countries (see Table 1) and across individuals within countries. In Section 4, we show that this heterogeneity does not systematically bias our results for the upper bound.

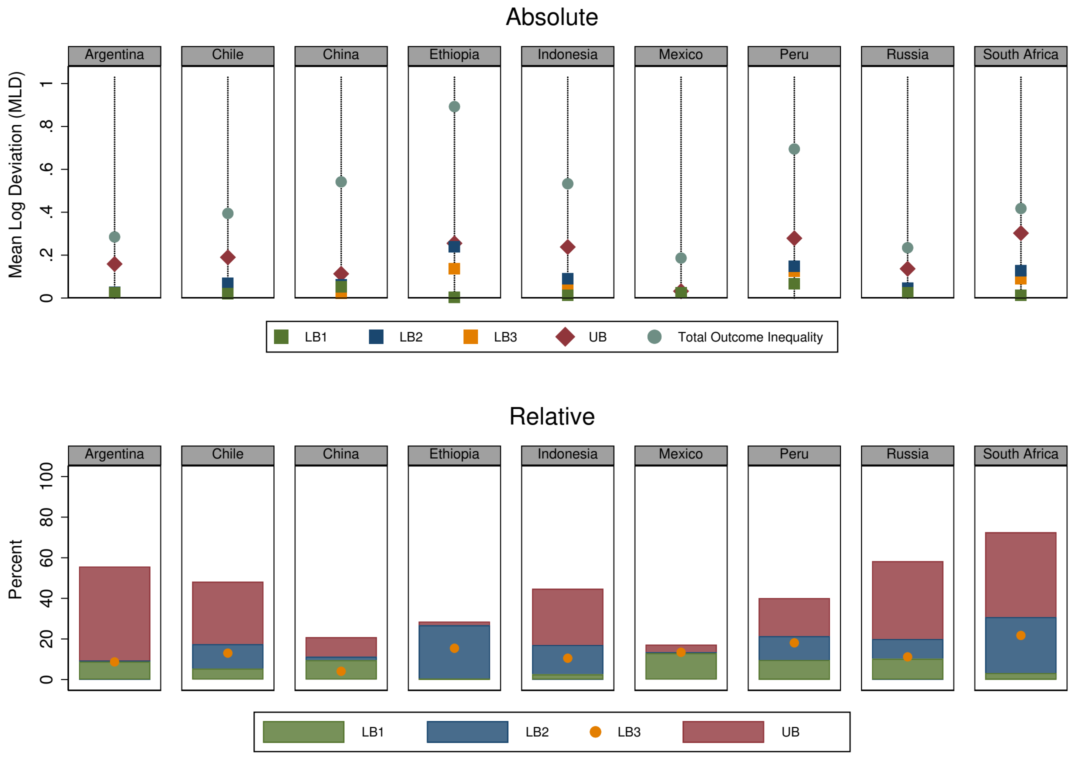
4 Empirical Results

Figures 1-3 show IOp estimates for individual incomes, household incomes and household expenditures, respectively. In each figure the upper panel shows an absolute measure where the MLD is applied to the distribution of circumstance type means. The lower panel shows a relative measure where the absolute measure of IOp is scaled by total outcome inequality. Hence, the relative measure expresses in percentage terms how much of total inequality can be attributed to the influence of circumstances. An overview table including all results is presented in Appendix Table B.2.

Individual Income. The upper panel of Figure 1 shows substantial variation in individual income inequality across the countries under consideration. Total outcome inequality figures based on the MLD range between 0.2 (Russia, Mexico) and 0.9 (Ethiopia). Data points for the first lower bound estimate (LB1) indicate IOp based on the circumstance variables gender and year of birth, only. Data points for the second lower bound estimate (LB2) indicate IOp based on all circumstances available in the particular country data set. Data points for the third lower bound estimate (LB3) indicate IOp based on the same set of circumstances. However, in contrast to LB2 we take account of potential upward biases due to sampling variation by applying a lasso estimation in which the relevant circumstances are chosen by means of 5-fold cross-validation. Data points for the upper bound estimate (UB) indicate IOp based on the fixed effect estimation procedure outlined in Section 2.

Estimates of LB1 indicate that gender and year of birth can explain only a very limited share of total outcome inequality. In relative terms, these variables explain only between 0.3% (Ethiopia) and 13.0% (Mexico) of total outcome inequality. Such low estimates reflect the concern of Kanbur and Wagstaff (2016) that internationally comparable estimates of IOp that are based on a common set of circumstance variables will further magnify the limited capacity of lower bound estimators to detect unfairness in a given outcome distribution. Estimates of LB2 show that the integration of additional country-specific circumstance variables substantially increases IOp estimates. Estimates now range between 9.4% (Argentina) to 30.7% (South Africa). This remains true even after accounting for the sampling variation to address the potential upward bias of LB2: According to the lasso-based estimates of LB3, between 4.1% (China) and 21.7% (South Africa) of outcome inequality can be considered unfair. However, IOp estimates based on *observed* circumstances on average only represent 7.3% (LB1), 18.9% (LB2) and 12.9% (LB3) of total inequality. If policy makers indeed take such estimates as reference points to evaluate the need for inequality-reducing policy reforms, one may be reasonably concerned about the real-world consequences of lower bound IOp estimates. Therefore, we take account of *unobserved* circumstances through the fixed effects estimation procedure that determines the UB estimates of IOp. The UB estimates vary between 17.2% (Mexico) and 72.5% (South Africa) and therefore show a significant upward correction of IOp in comparison to LB1-3. The unweighted average across countries yields a relative measure of 42.4%, i.e. almost half of observed inequality can be attributed to circumstance characteristics.

Figure 1: Inequality of Opportunity, Individual Income



Notes: The upper panel shows absolute estimates of inequality of opportunity. The lower panel shows estimates of inequality as a share of total inequality. LB1 and LB2 indicate estimates based on the circumstance sets indicated in Table 1. LB3 uses the same set of circumstances as LB2 but employs a lasso estimation to account for sampling variance. UB indicates the upper bound estimate based on the fixed effect specification.
Source: Own calculations based on the panel survey data described in Table A.1.

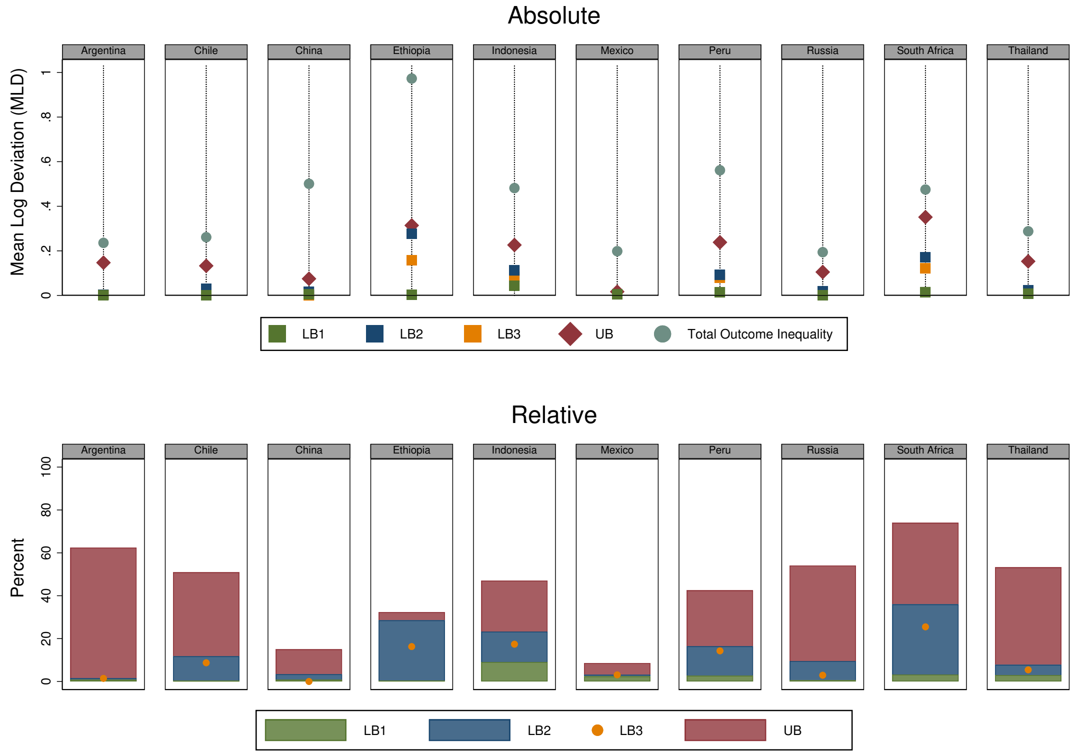
Household Income. Results concerning individual incomes neither consider resource sharing at the household level nor heterogeneity in household structures. In the following, we thus provide absolute and relative IOp estimates for equivalized household incomes.

Figure 2 shows that resource sharing at the household level reduces outcome inequality in most countries in our sample. Exceptions are Ethiopia, Mexico and South Africa, for which household representative income is more unequally distributed than individual income. The unusual pattern in these countries is mostly driven by assortative matching. As a consequence, inequality in individual incomes for households with a given number of members is lower than the respective inequality in household incomes. In addition, in South Africa and Mexico individual incomes are negatively correlated with the number of children. Adjusting for the household structure thus further exacerbates the existing inequality in individual incomes.

Since some part of gender-based differences are wiped-out through household resource sharing, LB1 decreases substantially for all countries except Ethiopia, Indonesia and South Africa. On average, gender and year of birth only account for 2.5% of total inequality in equivalized household incomes for the countries in our sample. Similarly, LB2 (LB3) tends to decrease for the vast majority of countries and now lies between 1.6%, Argentina (0%, China) and 36% (25.5%, South Africa). In spite of sizable reductions, a substantial share of IOp remains even after controlling for household structure and resource sharing. Across countries, estimates for

LB2 (LB3) suggest that observed circumstances can explain on average 14.4% (9.5%) of total inequality in household incomes. In contrast to the lower bound estimates, the UB estimates for IOp in household incomes are rather constant when comparing them to their individual income analogues: Average IOp slightly increases to 43.2% in relative terms.[12]

Figure 2: Inequality of Opportunity, Household Representative Income

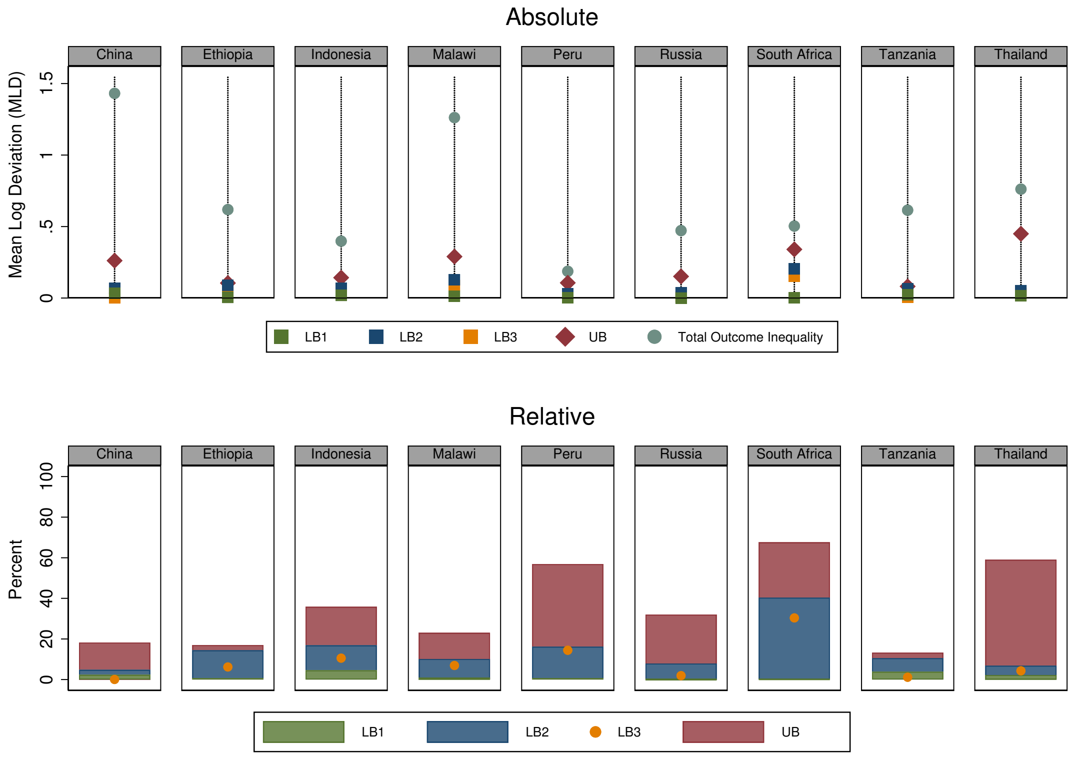


Notes: The upper panel shows absolute estimates of inequality of opportunity. The lower panel shows estimates of inequality as a share of total inequality. LB1 and LB2 indicate estimates based on the circumstance sets indicated in Table 1. LB3 uses the same set of circumstances as LB2 but employs a lasso estimation to account for sampling variance. UB indicates the upper bound estimate based on the fixed effect specification.
Source: Own calculations based on the panel survey data described in Table A.1.

Household Expenditure. Figure 3 presents estimates for IOp in household expenditures. There are more marked cross-country differences in expenditure inequality than for individual- or household based income inequality. We refrain from interpreting these country differences since they are mostly driven by survey-specific questionnaire artifacts. In some countries, total expenditures are queried directly, whereas other surveys provide categorized information on consumption expenditures. Moreover, reference periods vary strongly (weekly, monthly, quarterly)

[12]Two explanations are conceivable. First, household income is the sum of incomes from all household members. Hence, to the extent that household members match based on circumstances that are unobserved in the lower bound estimates, LB1-3 will show a stronger decrease than the UB estimate since the latter also accounts for unobserved circumstance heterogeneity across individuals (Peichl and Ungerer 2016). Second, there may be intra-household substitution in labor market efforts. For example, if households keep working hours fixed but alternate the individual contribution of each spouse, the UB estimate for household income also includes time-constant household-level effort and therefore decreases less than LB1-3 which are net of any time constant effort components. Due to the data limitations that motivate this bounding exercise in the first place, we cannot discriminate between both explanations.

Figure 3: Inequality of Opportunity, Household Representative Expenditure



Notes: The upper panel shows absolute estimates of inequality of opportunity. The lower panel shows estimates of inequality as a share of total inequality. LB1 and LB2 indicate estimates based on the circumstance sets indicated in Table 1. LB3 uses the same set of circumstances as LB2 but employs a lasso estimation to account for sampling variance. UB indicates the upper bound estimate based on the fixed effect specification.
Source: Own calculations based on the panel survey data described in Table A.1.

which may lead to artificial changes through scaling effects. Hence, we focus on the bounds of IOp estimates within countries.

Independent of the absolute amount of inequality in expenditures, the upper panel of Figure 3 shows that unfair inequality is also manifested in consumption. In line with the results on household incomes, gender and the year of birth are rather negligible determinants of unfair inequality at the household level. However, other observed circumstances matter, which confirms that household level sharing of resources does not eliminate unfair inequality. Referring to LB2 (LB3), between 4.8% (0.1%, China) and 40.4% (30.4%, South Africa) of total inequality in household expenditures can be considered unfair. In analogy to the previously discussed income variables, IOp increases substantially when accounting for unobserved circumstances. In relative terms, the UB estimates lie between 13.3% (Tanzania) and 67.6% (South Africa).

Sensitivity Analysis. We conduct two sensitivity checks. First, our baseline estimates differ in terms of the year for which IOp is estimated. To increase comparability we replicate our analysis for the country-specific data waves in closest proximity to 2009.[13] Given that a society's opportunity structure is shaped by long-run institutional features, we expect that the baseline

[13]See Table C.3 for an overview of the country-specific adjustments. In cases where there are two available data points with the same distance from 2009, the more recent year is chosen.

estimates are similar to those using the harmonized year of interest. Indeed, plotting the latter against the former we see that all estimates nicely group around the 45 degree line (see Figure C.1).

Second, the fixed effects for the upper bound estimators are based on different numbers of individual observations. While we fix a minimum of three data points per individual in order to be eligible for our data sample, the de facto number of observations used for the construction of the individual fixed effect varies across countries (Table 1). To test whether this heterogeneity distorts cross-country comparisons, we provide upper bound estimates in which we restrict the sample to the three most recent observations for each individual. We plot the estimates from this alternative specification against our baseline estimates and again find that all estimates closely align to the 45 degree line (see Figure C.2).

5 Conclusion

Estimates of inequality of opportunity have been heavily criticized for their lower bound property. This concern is particularly relevant for emerging economies whose household surveys tend to avail less comprehensive information on circumstances than data sets in industrialized countries. In this work we address this concern by providing upper and lower bound estimates of IOp for a set of twelve emerging economies.

We find that differences between lower and upper bounds of IOp are substantial across all countries and independent of whether we consider (household) income or expenditure as the outcome of interest. The magnitude of these differences suggests substantial uncertainty with respect to the true estimate of IOp which highlights the need for better data sets and a further efforts to refine the econometric toolkit employed in this literature. In the meantime, however, bounding the range of potential estimates is a useful exercise to limit the scope for downplaying the moral significance of inequality in the countries of interest.

References

Alesina, A., S. Stantcheva, and E. Teso. 2018. "Intergenerational Mobility and Preferences for Redistribution". *American Economic Review* 108 (2): 521–554.

Alesina, A., S. Hohmann, S. Michalopoulos, and E. Papaioannou. 2019. "Intergenerational Mobility in Africa". *mimeo.*

Andreoli, F., T. Havnes, and A. Lefranc. 2019. "Robust inequality of opportunity comparisons: Theory and application to early-childhood policy evaluation". *Review of Economics and Statistics* Forthcoming.

Balcázar, C. F. 2015. "Lower bounds on inequality of opportunity and measurement error". *Economics Letters* 137:102–105.

Blackburn, M. L. 2007. "Estimating wage differentials without logarithms". *Labour Economics* 14 (1): 73–98.

Brunori, P., F. Palmisano, and V. Peragine. 2015. "Inequality of Opportunity in Sub-Saharan Africa". *ECINEQ Working Paper* 2016-390.

Brunori, P., P. Hufe, and D. Mahler. 2018a. "The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees". *ifo Working Papers* 252.

Brunori, P., V. Peragine, and L. Serlenga. 2018b. "Upward and downward bias when measuring inequality of opportunity". *Social Choice and Welfare* Forthcoming.

Cappelen, A. W., A. D. Hole, E. Ø. Sørensen, and B. Tungodden. 2007. "The Pluralism of Fairness Ideals: An Experimental Approach". *American Economic Review* 97 (3): 818–827.

Cubas, G. 2016. "Distortions, infrastructure, and female labor supply in developing countries". *European Economic Review* 87:194 –215.

Faravelli, M. 2007. "How context matters: A survey based experiment on distributive justice". *Journal of Public Economics* 91 (7–8): 1399–1422.

Federal Reserve Bank of St. Louis. 2018. "Consumer Price Indexes (CPI and PCE)". Visited on 05/12/2018. https://fred.stlouisfed.org/categories/9.

Ferreira, F. H. G., and J. Gignoux. 2011. "The Measurement of Inequality of Opportunity: Theory and an Application to Latin America". *Review of Income and Wealth* 57 (4): 622–657.

Fleurbaey, M. 1995. "Three Solutions for the Compensation Problem". *Journal of Economic Theory* 65 (2): 505–521.

Hufe, P., A. Peichl, J. E. Roemer, and M. Ungerer. 2017. "Inequality of Income Acquisition: The Role of Childhood Circumstances". *Social Choice and Welfare* 143 (3-4): 499–544.

Jusot, F., S. Tubeuf, and A. Trannoy. 2013. "Circumstances and Efforts: How Important is their Correlation for the Measurement of Inequality of Opportunity in Health?" *Health Economics* 22 (12): 1470–1495.

Kanbur, R., and A. Wagstaff. 2016. “How Useful Is Inequality of Opportunity as a Policy Construct?” Chap. 4 in *Inequality and Growth: Patterns and Policy*, ed. by K. Basu and J. E. Stiglitz, 1:131–150. London: Palgrave Macmillan UK.

Klasen, S., et al. 2018. “Economic Inequality and Social Progress”. In *Rethinking Society for the 21st Century: Report of the International Panel on Social Progress*, ed. by IPSP, 1:83–140. Cambridge University Press.

Niehues, J., and A. Peichl. 2014. “Upper bounds of inequality of opportunity: Theory and evidence for Germany and the US”. *Social Choice and Welfare* 43 (1): 73–99.

OECD. 2018. “Employment Rate by Age Group”. Visited on 05/12/2018. https://data.oecd.org/emp/employment-rate-by-age-group.htm.

Peichl, A., and M. Ungerer. 2016. “Accounting for the spouse when measuring inequality of opportunity”. *Social Choice and Welfare* 47 (3): 607–631.

Ramos, X., and D. Van de gaer. 2016. “Empirical Approaches to Inequality of Opportunity: Principles, Measures, and Evidence”. *Journal of Economic Surveys* 30 (5): 855–883.

Roemer, J. E. 1993. “A Pragmatic Theory of Responsibility for the Egalitarian Planner”. *Philosophy & Public Affairs* 22 (2): 146–166.

— . 1998. *Equality of Opportunity.* Cambridge: Harvard University Press.

Roemer, J. E., and A. Trannoy. 2015. “Equality of Opportunity”. Chap. 4 in *Handbook of Income Distribution*, ed. by A. B. Atkinson and F. Bourguignon, vol. 2A, 217–300. Amsterdam: Elsevier.

Son, H. H. 2013. “Inequality of Human Opportunities in Developing Asia”. *Asian Development Review* 30 (2): 110–130.

Tibshirani, R. 2011. “Regression shrinkage and selection via the lasso: a retrospective”. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73 (3): 273–282.

Van de gaer, D. 1993. “Equality of Opportunity and Investment in Human Capital”. PhD thesis, University of Leuven.

World Bank. 2018. “World Bank Country and Lending Groups”. Visited on 05/12/2018. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519.

A Data Overview

Table A.1: Overview

Country	Panel Period	Waves	Data	Source	Individual	Household		Sample Weights Available
					Income	Income	Expenditure	
Argentina	2003-2015	12	Encuesta Permanente de Hogares	Instituto Nacional de Estadística y Censos (INDEC)	Net	Net	-	Yes
Chile	2006-2009	4	Encuesta Panel CASEN	Ministerio de Desarrollo Social, Chile	Net	Net	-	Yes
China	1988-2014	10	China Health and Nutrition Survey	Carolina Population Center at the University of North Carolina at Chapel Hill and National Institute for Nutrition and Health (NINH) at the Chinese Center for Disease Control and Prevention (CCDC)	Net (Labor)	Net (Labor)	Yes	No
Ethiopia	1994-2009	6	Ethiopia Rural Household Survey	International Food Policy Research Institute (IFPRI), Washington DC	Net	Net	Yes	No
Indonesia	1992-2013	5	Indonesian Family Life Survey (IFLS)	RAND Social and Economic Well-Being	Net (Labor)	Net (Labor)	Yes	No
Malawi	1998-2010	4	Malawi Longitudinal Study of Family and Health	Population Studies Center at the University of Pennsylvania and College of Medicine at the University of Malawi and Invest in Knowledge (IKI) in Zomba, Malawi	-	-	Yes	No
Mexico	1999-2009	7	Encuesta Evaluation de los Hogares (ENCEL)	International Food Policy Research Institute (IFPRI), Washington DC	Net	Net	-	No
Peru	1998-2011	14	Encuesta Nacional de Hogares, Condiciones de Vida y Pobreza	Instituto Nacional de Estadística e Informatica	Gross	Gross	Yes	Yes
Russia	1994-2017	22	Russia Longitudinal Monitoring Survey (RLMS)	National Research University "Higher School of Economics", OOO “Demoscope”, Carolina Population Center, University of North Carolina at Chapel Hill and the Institute of Sociology of the Federal Center of Theoretical and Applied Sociology of the Russian Academy of Sciences	Net (Labor)	Net (Labor)	Yes	Yes
South Africa	2008-2017	5	National Income Dynamics Study (NIDS)	Southern Africa Labour and Development Research Unit (SALDRU), University of Cape Town	Net (Labor)	Net (Labor)	Yes	Yes
Tanzania	1991-2010	6	Kagera Health and Development Survey	Economic Development Initiatives	-	-	Yes	No
Thailand	1997-2017	21	Townsend Thai Data	The Townsend Thai Project	-	Gross	Yes	No

Notes: All year specifications refer to the year in which the outcome (income/expenditure) was realized.
Source: Own calculations based on the panel survey data described in Table A.1.

B Baseline Results

Table B.2: Absolute and Relative Inequality of Opportunity, Baseline Specification

						Absolute (MLD)				Relative (%)			
Country	Outcome	Year	N (Single Year)	N (FE)	Inequality (MLD)	LB1	LB2	LB3	UB	LB1	LB2	LB3	UB
Argentina	Individual Income	2015	3020	6040	0.285	0.025	0.027	0.025	0.159	8.798	9.389	8.706	55.700
Chile	Individual Income	2009	2808	8424	0.395	0.021	0.069	0.051	0.190	5.369	17.420	13.030	48.160
China	Individual Income	2014	261	771	0.542	0.052	0.061	0.022	0.113	9.512	11.240	4.086	20.830
Ethiopia	Individual Income	2009	697	2533	0.893	0.003	0.238	0.138	0.255	0.295	26.720	15.410	28.600
Indonesia	Individual Income	2013	786	2036	0.533	0.013	0.091	0.056	0.238	2.521	17.030	10.530	44.660
Mexico	Individual Income	2009	3050	16552	0.186	0.024	0.025	0.025	0.032	12.983	13.430	13.427	17.180
Peru	Individual Income	2011	2193	5878	0.695	0.067	0.148	0.126	0.279	9.629	21.355	18.092	40.096
Russia	Individual Income	2017	1181	10816	0.234	0.024	0.047	0.026	0.137	10.180	19.950	11.180	58.270
South Africa	Individual Income	2017	670	2331	0.418	0.013	0.128	0.091	0.303	3.198	30.730	21.740	72.460
Argentina	Household Income	2015	3020	6040	0.236	0.002	0.004	0.003	0.147	0.826	1.563	1.455	62.420
Chile	Household Income	2009	2808	8424	0.261	0.001	0.031	0.023	0.133	0.268	11.730	8.728	50.880
China	Household Income	2014	261	771	0.501	0.004	0.017	0.000	0.075	0.858	3.349	0.000	14.960
Ethiopia	Household Income	2009	697	2533	0.973	0.003	0.277	0.158	0.314	0.306	28.500	16.280	32.290
Indonesia	Household Income	2013	786	2036	0.482	0.044	0.112	0.084	0.226	9.089	23.250	17.390	47.000
Mexico	Household Income	2009	3050	16552	0.198	0.005	0.006	0.006	0.017	2.343	3.096	3.095	8.563
Peru	Household Income	2011	2193	5878	0.561	0.015	0.092	0.080	0.238	2.676	16.404	14.279	42.454
Russia	Household Income	2017	1181	10816	0.194	0.001	0.018	0.006	0.105	0.504	9.515	2.907	54.100
South Africa	Household Income	2017	670	2331	0.475	0.015	0.171	0.121	0.351	3.236	35.960	25.500	73.940
Thailand	Household Income	2017	467	6371	0.288	0.008	0.022	0.016	0.153	2.944	7.751	5.439	53.200
China	Household Expenditure	2014	261	771	1.431	0.033	0.068	0.001	0.261	2.305	4.745	0.049	18.270
Ethiopia	Household Expenditure	2009	697	2533	0.618	0.004	0.089	0.038	0.105	0.670	14.390	6.198	16.970
Indonesia	Household Expenditure	2013	786	2036	0.398	0.018	0.067	0.042	0.143	4.531	16.830	10.550	35.810
Malawi	Household Expenditure	2010	503	1278	1.262	0.013	0.128	0.088	0.290	1.014	10.130	6.944	22.970
Peru	Household Expenditure	2011	2193	5878	0.187	0.001	0.030	0.027	0.106	0.714	16.19	14.423	56.881
Russia	Household Expenditure	2017	1181	10816	0.472	0.000	0.037	0.009	0.151	0.045	7.918	1.944	32.000
South Africa	Household Expenditure	2017	670	2331	0.503	0.001	0.203	0.153	0.340	0.192	40.360	30.380	67.570
Tanzania	Household Expenditure	2010	203	753	0.614	0.024	0.065	0.007	0.081	3.828	10.540	1.140	13.250
Thailand	Household Expenditure	2017	467	6371	0.762	0.017	0.052	0.033	0.450	2.258	6.837	4.310	59.020

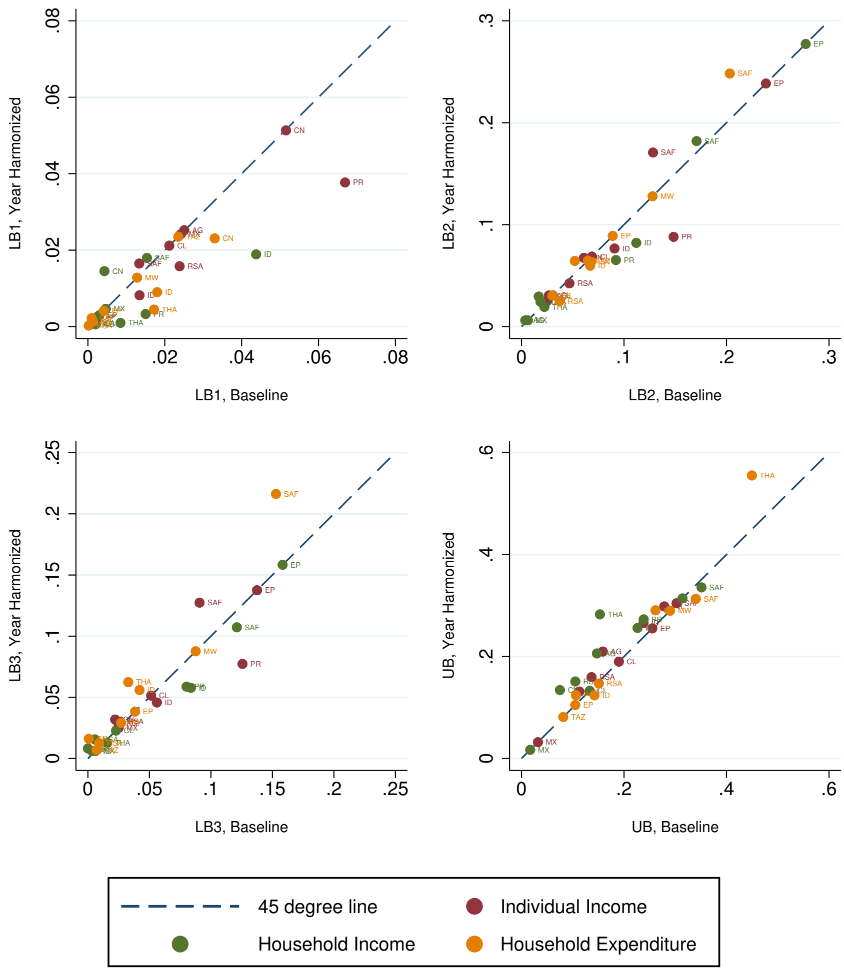
Notes: All year specifications refer to the year in which the outcome (income/expenditure) was realized.
Source: Own calculations based on the panel survey data described in Table A.1.

C Sensitivity Analysis

Table C.3: Year of Interest, Baseline and Harmonized

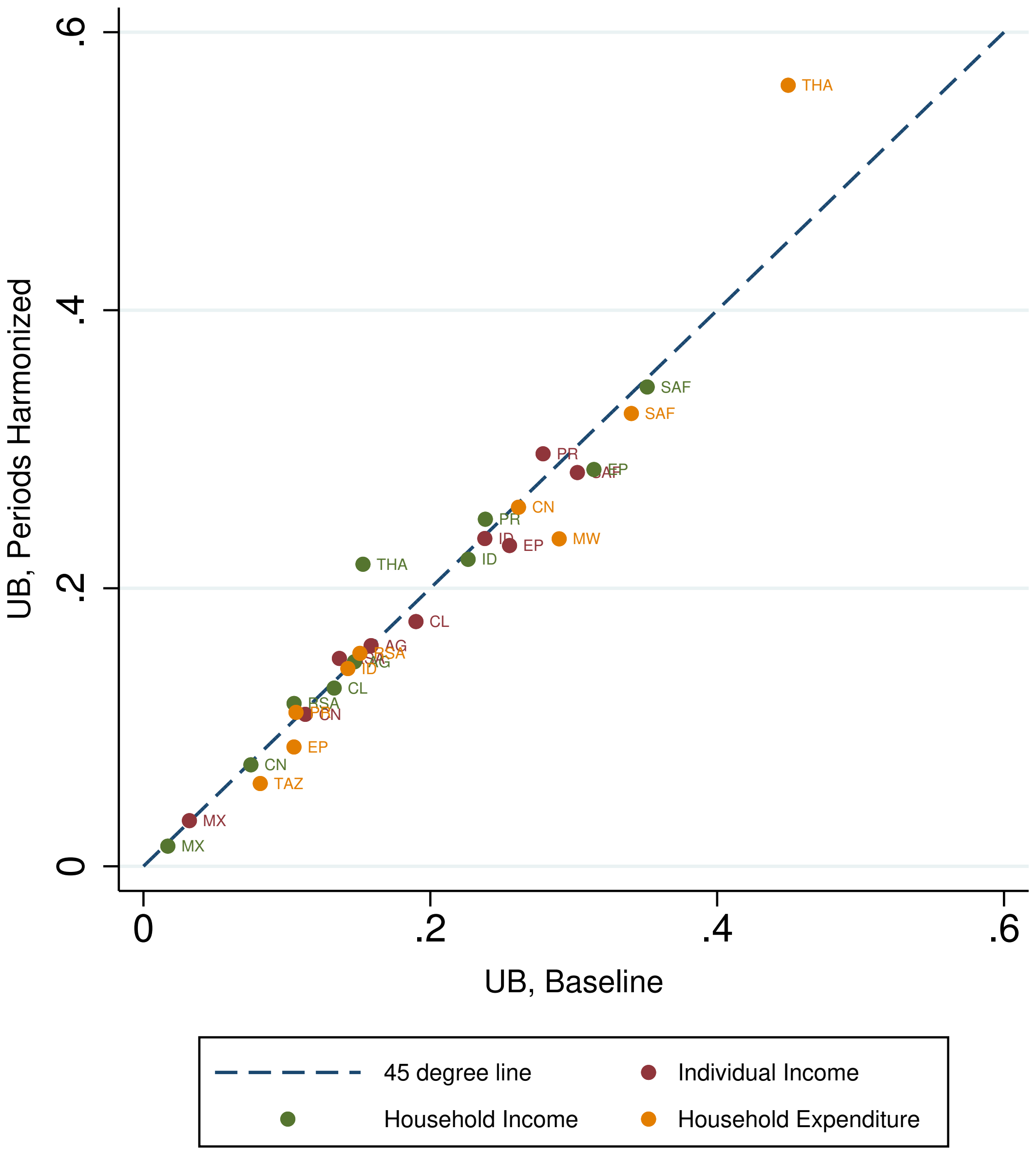
	Year Baseline	Year Harmonized	Difference in Years
Argentina	2015	2013	2
Chile	2009	2009	0
China	2014	2010	4
Ethiopia	2009	2009	0
Indonesia	2013	2006	7
Malawi	2010	2010	0
Mexico	2009	2009	0
Peru	2010	2009	1
Russia	2017	2009	8
South Africa	2017	2008	9
Tanzania	2010	2010	0
Thailand	2017	2009	8

Figure C.1: Comparison of Baseline and Year Harmonized Estimates



Notes: The four panels plot baseline estimates of absolute inequality of opportunity against absolute inequality of opportunity estimates for a harmonized year according to Table C.3. LB1 (first panel) and LB2 (second panel) show estimates based on the circumstance sets indicated in Table 1. LB3 (third panel) uses the same set of circumstances as LB2 but employs a lasso estimation to account for sampling variance. UB (fourth panel) indicates the upper bound estimate based on the fixed effect specification.
Source: Own calculations based on the panel survey data described in Table A.1.

Figure C.2: Comparison of Baseline and Period Harmonized Upper Bound Estimates



Notes: The figure plots baseline estimates of absolute inequality of opportunity for the upper bound against upper bound estimates of a period harmonized version. In this period harmonized specification, we use for each unit of observation the three most recent observations only. Upper bound estimates (UB) are based on the fixed effect specification.
Source: Own calculations based on the panel survey data described in Table A.1.