

2022

IARIW-TNBS Conference on “Measuring Income, Wealth and Well-being in Africa”

Paper Prepared for the IARIW-TNBS Conference, Arusha, Tanzania, April 14-15, 2022

“Recovering Within-Country Inequality from Trade Data”

Dorothee Hillrichs (Université Laval, Canada), Gonzague Vannoorenberghe (Université Catholique de Louvain, Belgium)

Collecting data on within-country inequality in a reliable and consistent manner across countries is challenging. For the period 2016-2018, the United Nations' World Income Inequality Database reports income inequality data for only 55 countries.

Inequality data are typically based on surveys and fiscal data to supplement the survey information. Yet, surveys are costly to conduct, and questionnaires are designed for country-specific needs impeding a meaningful cross-country comparison. Measurement error due to under-reporting or attrition pose further challenges to obtain a reliable assessment of inequality. Harmonization of the inequality reports often takes statisticians at international agencies several years. Fiscal data is available only where income is taxed though a large informal sector biases inequality measures based on tax data especially in less developed economies. Missing values are pervasive particularly among developing countries, among them many African nations, where budgetary constraints to conduct surveys are most stringent. Common drivers of within-country inequality, therefore, remain under-studied.

We present a novel method of measuring within-country inequality. We infer within-country inequality from import patterns. Since imports are globally observed with high quality, we can fill in the missing data using a consistent methodology. Our method builds on the link between income and spending by product variety. As households become richer, they shift their expenditure from normal to luxury varieties. In the aggregate, the relative imports of luxuries to normal varieties can, therefore, be traced to the dispersion of income in a country accounting for the average income of the economy. We extend the data coverage especially among less developed economies and provide a consistent measure of inequality for 160 countries over 24 years.

To guide our empirical analysis, we develop an Armington type international trade framework embedding non-homothetic preferences. In the Armington world, a product variety is identified by an exporter country. For example, Swiss and German watches are two varieties of watches. We maintain the constant elasticity of substitution (CES) preferences structure, standard in international trade models, while allowing for non-linear income effects. Aggregating households' expenditure on a variety - an exporter-product pair - we derive a non-homothetic gravity equation: alongside standard gravity determinants such as bilateral trade costs and multilateral resistance terms, imports depend on a country's income distribution. We approximate the income distribution by the average income in the economy and a function of the country's Gini coefficient.

We exploit the link between a country's income distribution and its imports following a two-step approach. First, we use a subsample of destinations for which high quality inequality data is observed to estimate the parameters of the aggregate non-homothetic gravity equation. The variety-specific slope and curvature of the expenditure function are identified from the variation of per capita income and of the Gini coefficient across destinations, respectively. Crucially, we find that the parameters of the expenditure function vary across varieties, i.e., an income redistribution affects imports of Swiss watches differently to imports of German watches. We estimate income and inequality elasticities for each of the more than 6500 exporter-product combinations in our sample.

We impose that the income and inequality estimates are constant within groups of destinations that share similar expenditure patterns. Constant preference parameters across countries allow us to use a subsample of countries in the first step, and to infer inequality for any country in the second step - irrespective of whether the country forms part of the first sample. This identifying assumption implies that households in different countries rank the various (watch) varieties in the same way in their spending decision. In the example, Tanzanians and Kenyans are both assumed to rank Swiss watches above German watches.

We obtain the trade-based inequality index in the second step. Here we exploit the variation of the varieties' income elasticities within goods. Inverting the demand system from the first step, we find a country's Gini index from a regression of imports on the parameter governing the curvature of the expenditure function. We ask, given country A's average income, its distance from Switzerland and other trade determinants, how many Swiss watches do we expect country A to import relative to its average imports of watches? We attribute deviations from the predicted imports at the average income to income inequality.

The indirect method of generating inequality data proposed in this paper takes advantage of the reliability and availability of trade data. Trade flows are registered by customs administration both of the exporting and the importing country. Product level trade flows are

recorded according to the Harmonized System, a product classification system shared by all countries worldwide. The double information allows the reconciliation of reports from both trading countries. This makes trade data a reliable source of information on a country's expenditure pattern, from which we infer inequality. In addition, unlike survey or fiscal data, availability of trade data does not necessarily depend on a country's statistical capacity. Nor is the frequency with which it is provided subject to governmental decisions. Besides, our method is relatively cost and time-efficient and therefore allows much quicker data generation than conducting surveys.

We find that our within-country inequality estimates closely track inequality measured by high-quality survey data. The median difference between the predicted Gini and UN survey data is three Gini points in a 10-fold cross-validation. We also contrast the trade-based inequality estimates to two survey databases that have not been used in the first step of our estimation (LIS and WDI). We show, our estimate is close most of the time to at least one survey data. The overall high correlation underscores the credibility of our method for filling in missing inequality data.

Our inequality estimates pave the way to gaining further insight in common determinants of inequality across countries. The consistent methodology applied to all countries and time periods make comparisons across countries meaningful. To illustrate such use, we examine the development of within-country inequality by world region, income group and natural resource dependency.