Household consumption and emissions from corrected microdata: A DINA application for Ecuador

Markus Nabernegg*[a], Thomas Kopp[a], and Stefan Nabernegg[b]

[a]School of Economic Disciplines, University of Siegen, Germany

[b]Wegener Center, Karl-Franzens University Graz, Austria

Extended Abstract

Motivation

One key question within the existing literature on the interrelation between income, inequality and carbon emissions, is whether the combination of the normative goals of lowering inequality and reducing carbon emissions constitutes a virtuous social optimum (Rao and Min, 2018), or a social dilemma in which reducing inequality increases carbon emissions (Grunewald et al., 2017; Rojas-Vallejos and Lastuka, 2020).

If such a social dilemma indeed existed, it would require societies to decide upon a trade-off, making it imperative for science to shed light on this relationship. As recent research about different inequality dimensions shows (Alvaredo et al., 2017; Burdin et al., 2022; De Rosa et al., 2020; Blanchet et al., 2019), data quality and completeness on the one hand, and precision of the methods on the other hand are essential first elements to provide these necessary insights. The same condition holds for the estimation of inequalities in carbon emissions, which are to a great extend embodied in personal consumption and therefore related to income. Nevertheless, we have identified that existing studies either lack completeness as they do not account for the upper part of the distribution (Hardadi et al., 2021; Levinson and O'Brien, 2019), or they lack precision due to a very simple elasticity assumption between income and emissions (Chancel et al., 2021).

This study aims to combine the completeness of the household income distribution with the precision of linking emissions to income not through a constant elasticity, but through the intermediary of household consumption. The pivot challenge is therefore to predict consumption patterns of high-income households, in order to allocate their emissions correctly. We implement our methodology for the case of Ecuador.

*markus.nabernegg@uni-siegen.de, Kohlbettstraße 17, 57072 Siegen, +49 271 74012899.

Methods

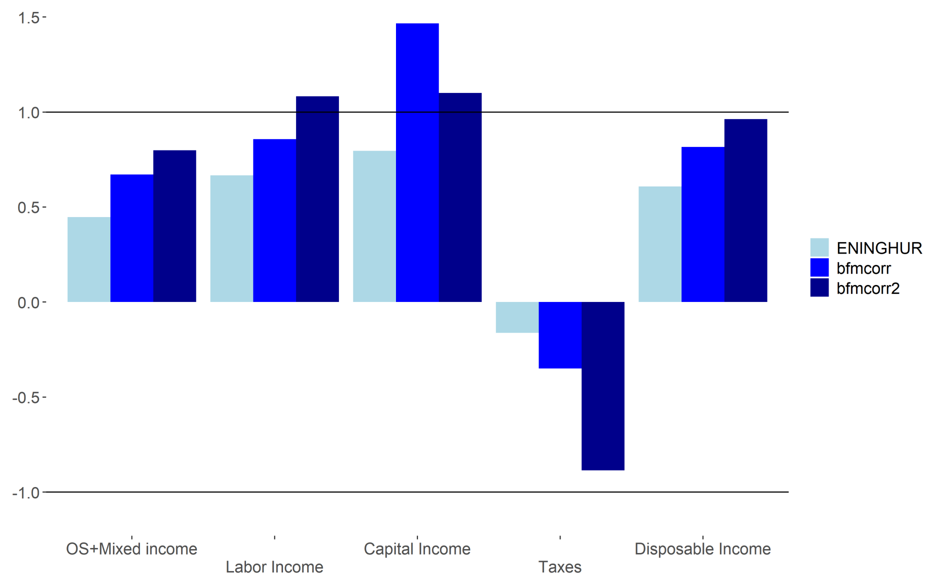
To correctly account carbon emissions to households, one relevant limitation is systematic in its nature, namely the lack of information on the "missing rich" (Lustig, 2020, p.2). These observations lack because consumption data is generated via household surveys which systematically fail to reach out to the most affluent individuals due to under-reporting and under-coverage - i.e. selected households not responding.

Our method assess this limitation and relies on three steps: First, we build a synthetic micro-dataset that is corrected for high-income households, stemming from exhaustive household survey and tax register micro-data. In the integration process, high-income individuals from the tax data are integrated to households from the survey, which allows to maintain some characteristics from these integrated individuals beside disposable income (e.g. wages, capital income, residence). In addition, the dataset contains detailed information about income and consumption categories for the households in the survey. For the implementation, we amplify the existing data integration method `bfmcorr` from Blanchet et al. (2022), which was designed to integrate survey and tax information for countries and years where tax microdata is not available, to the case where the tax authority provides detailed tax declarations. One further advantage of this amplified method is that we do not lose valuable information from income that is computed on the household level (e.g. imputed rents or social benefits), as we integrate the richest individuals from tax data (with its income and demographic characteristics) into the richest households from the survey (with their income and sociodemographic characteristics). Therefore, we already cover a very high percentage of household income from National Accounts and depend less on additional assumptions for upscaling to economy wide totals.

Second, we want to meet the challenge of the omission of high-income households from household consumption surveys. This step is at the center of our contribution, as existing literature (i) either omits high-income households because they are not part of the survey (Hardadi et al., 2021) or excluded on purpose (Levinson and O'Brien, 2019), or (ii) omits the step from income to consumption and then to emissions by applying an elasticity between income and emissions directly (Kartha et al., 2020; Chancel et al., 2021). While this second strand of literature gives valuable first insights, our new approach allows for nuances that go beyond total income as a predictor for emissions and therefore allows for different emission patterns of high-income households, depending on their characteristics. Our reasoning is in line with Pottier (2022), as we want to distinguish between income-emission elasticity and consumption-emission elasticity, and we belief that this elasticity is not constant for an entire economy, let alone for the whole world. To achieve this, the procedure has to predict consumption for high-income households with the different income characteristics that were integrated through the procedure laid out in the first step (including the "missing rich"). A model is developed and trained[1] for households where we have information on income and consumption to predict consumption categories (following the international COICOP standard classification) with characteristics that we also find in the tax register (e.g. wages, capital income, residence). The model is then

[1] At the current stage of development, we use simple statistical learning methods, such as regressions for each consumption categories. Simultaneously, we are also working on machine learning techniques, using the available consumption patterns in the existing survey data to complete the distribution of consumption via multi-output and multi-target regressions (Borchani et al., 2015; Masmoudi et al., 2020).

Figure 1: Coverage of income categories



Own elaboration based on data from ENINGHUR-INEC, SRI and BCE. The figure shows the percentage of each income category in comparison to National Accounts. ENINGHUR indicates results from original household survey, bfmcorr from the reweighting method (Blanchet et al., 2022) and bfmcorr2 the direct integration method described in this paper.

applied to high-income households.

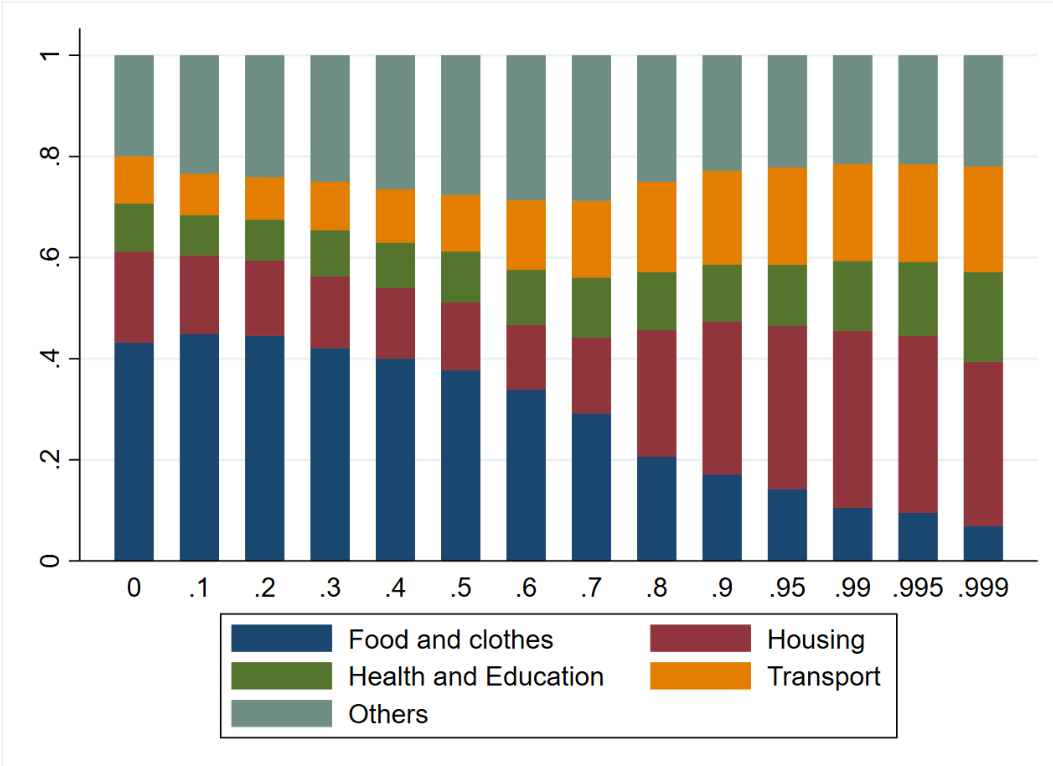
Finally, we map the calculated household consumption onto the corresponding carbon emissions through a Multi-Regional Input-Output model, accounting for direct and embodied emissions of household consumption. This allows us to add up the emissions from households to the corresponding total of the economy.

To show potential results, this paper applies the developed method for the case of Ecuador in the year 2012, which is the year of the most recent Income and Expenditure Survey (ENINGHUR), and which contains the information on income and consumption that is needed to map household emissions later. First, we combine this household survey with information from extensive tax micro records provided by the Internal Revenue Service (SRI) and National Accounts from the Ecuadorian Central Bank (BCE) to derive the entire income distribution of the household sector of the economy (Alvaredo et al., 2020; OECD and EUROSTAT, 2020). Second, we predict consumption for the corrected high-income households and, third, we use GTAP Regional Input-Output tables to map the emissions embodied in consumption to the households as described above.

Results

This paper is the first to provide insights about the distribution of income, consumption and carbon emissions on the household level stemming directly from corrected household micro data. The newly developed method is applied for Ecuador and indicates that the integration procedure allows for a much higher coverage of the bench-

Figure 2: Consumption categories by income group



Own elaboration based on data from ENINGHUR-INEC, SRI and BCE. The bars present the percentage that is spent on different consumption categories as a share of total expenditure in each income-percentile group. The horizontal axis shows percentiles.

mark household income provided by national accounts. Figure 1 shows that almost 96% of the aggregate household income is replicated by the new method, which alleviates the problem of depending on additional assumptions to scale income up to 100%.

Our model predicts that consumption patterns change also within the highest percentiles. This is shown in figure 2, where for example more than 40% of total expenditure of the poorest households is spent on food and clothes, whereas the richest 0.1% spends less than 7% on these items. Some categories also show considerable variation within the highest decile, such as spending on housing (between 0.30% and 0.35%). Health and education spending also increases substantially not only from the lowest income groups to the highest, but also within the richest 10%.

The integration method applied in step 1, in combination with consumption prediction and carbon emission mapping increases inequality indicators, as shown in table 1. The Gini index for income jumps from 0.44 to 0.59, whereas consumption consumption and emissions are concentrated less unequal on the income scale (both are concentration indexes ranked by income).

Further, mapping income of the ”missing rich” to their consumption, we provide improved estimates for classic Engel-curves and Environmental Engel-curves (Levinson and O'Brien, 2019) as shown in figure 3. Our findings suggest that also for Ecuador carbon emissions marginally decrease as income raises, which is in line with findings

Table 1: Concentration and Gini coefficients

income concept	ENINGHUR	bfmcorr2
Income	0.44	0.59
Consumption	0.38	0.52
Emissions	0.42	0.48

Own elaboration based on data from ENINGHUR-INEC, SRI and BCE. The table shows Gini Index for income, and Concentration indexes for consumption and emissions. The ranking variable is income. ENINGHUR indicates carbon emission mapping to original household survey, bfmcorr2 indicates results for the integration method described in this paper.

on the macro level where a trade-off between income redistribution and a reduction in carbon emission has been suggested by several authors (Grunewald et al., 2017; Rojas-Vallejos and Lastuka, 2020; Hailemariam et al., 2020).

Final results will be enriched by an improved prediction method and a comparison between the incidence of going from income-emission elasticity or consumption-emission elasticity to a emission mapping on the household level.

Bibliography

Alvaredo, Facundo, Anthony Atkinson, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2020). "Distributional National Accounts (DINA) guidelines: Concepts and methods used in the World Inequality Database". In: *updated and extended version of the DINA Guidelines first published as WID Working Paper 2016/2*.

Alvaredo, Facundo, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2017). "Global inequality dynamics: New findings from WID. world". In: *American Economic Review* 107.5, pp. 404–09.

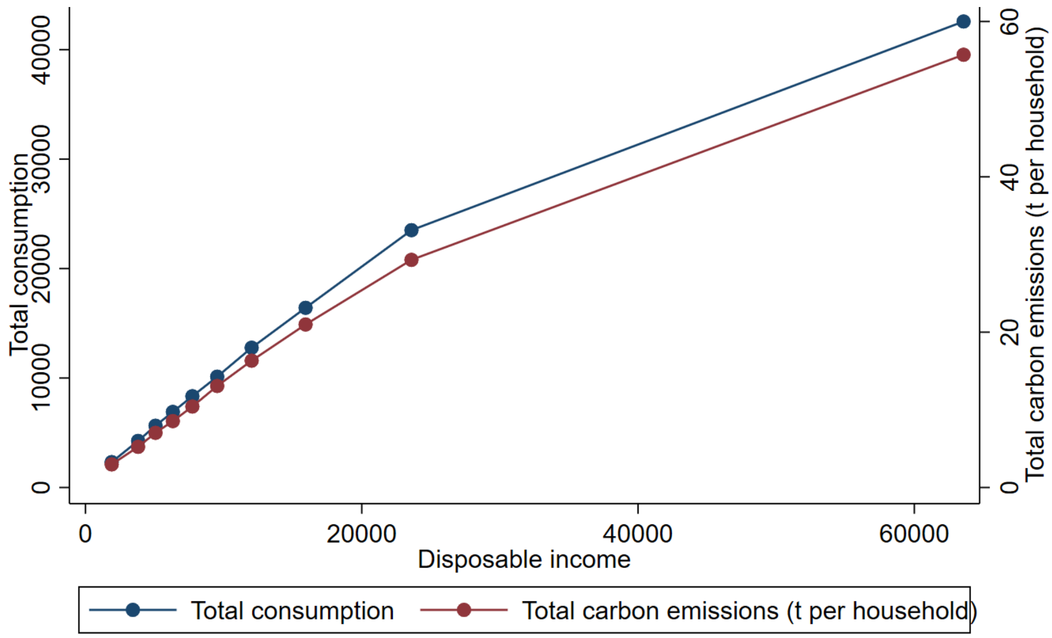
Blanchet, Thomas, Lucas Chancel, Amory Gethin, et al. (2019). "How unequal is Europe? Evidence from distributional national accounts, 1980-2017". In: *WID. world working paper* 6.

Blanchet, Thomas, Ignacio Flores, and Marc Morgan (2022). "The weight of the rich: improving surveys using tax data". In: *The Journal of Economic Inequality* 20.1, pp. 119–150.

Borchani, Hanen, Gherardo Varando, Concha Bielza, and Pedro Larranaga (2015). "A survey on multi-output regression". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 5.5, pp. 216–233.

Burdin, Gabriel, Mauricio De Rosa, Andrea Vigorito, and Joan Vila (2022). "Falling inequality and the growing capital income share: Reconciling divergent trends in survey and tax data". In: *World Development* 152, p. 105783.

Figure 3: Consumption and Environmental Engel curve



Own elaboration based on data from ENINGHUR-INEC, SRI and BCE. The figure shows disposable income on the horizontal axis, consumption and emissions on the vertical axis.

Chancel, Lucas, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman (2021). "World inequality report 2022". In: *World Inequality Lab*.

De Rosa, Mauricio, Ignacio Flores, and Marc Morgan (2020). "Inequality in Latin America revisited: insights from distributional national accounts". In: *World Inequality Lab – Technical Note* (02).

Grunewald, Nicole, Stephan Klasen, Inmaculada Martínez-Zarzoso, and Chris Muris (2017). "The Trade-off Between Income Inequality and Carbon Dioxide Emissions". In: *Ecological Economics* 142, pp. 249–256.

Hailemariam, Abebe, Ratbek Dzhumashev, and Muhammad Shahbaz (2020). "Carbon emissions, income inequality and economic development". In: *Empirical Economics* 59.3, pp. 1139–1159.

Hardadi, Gilang, Alexander Buchholz, and Stefan Pauliuk (2021). "Implications of the distribution of German household environmental footprints across income groups for integrating environmental and social policy design". In: *Journal of Industrial Ecology* 25.1, pp. 95–113.

Kartha, Sivan, Eric Kemp-Benedict, Emily Ghosh, Anisha Nazareth, and Tim Gore (2020). "The Carbon Inequality Era: An assessment of the global distribution of consumption emissions among individuals from 1990 to 2015 and beyond". In.

Levinson, Arik and James O'Brien (2019). "Environmental Engel Curves: Indirect Emissions of Common Air Pollutants". In: *The Review of Economics and Statistics* 101 (1). issn: 0034-6535.

Lustig, Nora (2020). "The "Missing Rich" in Household Surveys: Causes and Correction Approaches". In: *ECINEQ Working Paper Series 2020 520* 520, pp. 1–37.

Masmoudi, Sahar, Haytham Elghazel, Dalila Taieb, Orhan Yazar, and Amjad Kallel (2020). "A machine-learning framework for predicting multiple air pollutants' concentrations via multi-target regression and feature selection". In: *Science of the Total Environment* 715, p. 136991.

OECD and EUROSTAT (2020). "Guidelines of the OECD-Eurostat joint Expert Group on Disparities in a National Accounts framework - Version December 2020". In.

Pottier, Antonin (2022). "Expenditure elasticity and income elasticity of GHG emissions: A survey of literature on household carbon footprint". In: *Ecological Economics* 192, p. 107251.

Rao, Narasimha D. and Jihoon Min (2018). "Less global inequality can improve climate outcomes". In: *WIREs Climate Change* 9 (2). issn: 1757-7780.

Rojas-Vallejos, Jorge and Amy Lastuka (2020). "The income inequality and carbon emissions trade-off revisited". In: *Energy Policy* 139, p. 111302.