Environmental Engel Curves with predicted consumption of high-income households, applied to Ecuador

Markus Nabernegg*[1], Stefan Nabernegg[2], and Thomas Kopp[1]

[1] *School of Economic Disciplines, University of Siegen, Germany*
[2] *Wegener Center, University of Graz, Austria*

June, 2023

Abstract

This paper develops a method to account for the emissions embodied in consumption for the upper tail of the income distribution, which is systematically underrepresented in household surveys. This method is first accounts for the income of the missing rich by integrating tax microdata to the structure of the income and expenditure survey. In a second step, consumption for these introduced rich households is predicted with income and demographic characteristics. Finally, emissions are mapped to the consumption of households with environmental input-output tables. The application to Ecuador, which is a good example for the underrepresentation of high income households in surveys. Results indicate that Environmentel Engel Curves that account for those households question the prevalence of a concave income-emission relationship. Further, the integration method increases income inequality measures sharply, but suggests a lower income-emission elasticity than the one stemming from uncorrected surveys.

Keywords: Environmental Engel Curves; emission inequality; missing rich; Ecuador.
JEL Keywords: E01, D63, Q54.

1 Introduction

During the last decades, scholars have produced an increasing number of studies on the relationship between economic inequality and environmental damage. From a theoretical perspective, transmission channels associated with the political economy of a society, such as improved social norms towards environmental awareness, less corruption, or the reduction of conspicuous consumption are expected to help reducing emissions when a society becomes more equal (Leach et al., 2018; Klasen et al., 2016; Veblen, 1899). Nevertheless, existing empirical studies also find that on the household level, poorer households consume a greater part of their income, which implies that redistribution policies counteract environmental sustainability efforts (Kopp and Nabernegg, 2022; Grunewald et al., 2017; Rojas-Vallejos and Lastuka, 2020).

If such a social dilemma between inequality reduction and environmental protection (Sager, 2019) indeed existed, it would require societies to decide upon a trade-off, making it imperative for science

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

to shed light on this relationship. For climate policies, assigning greenhouse gas emissions to income is still a contested field, and several methodological questions arise when emissions are associated with income, such as whether to use total greenhouse gases (GHG) or only carbon dioxide (CO_2) emissions, or the appropriateness of different sources for the emission intensity of consumption goods. One additional question that we are interrested in this paper is aligned to the work of scholars that correct distributions for the "missing rich" (Lustig, 2020, p.2). As recent research on 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 correct insights on distributional topics. 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. We have identified that existing studies lack completeness as they do not account for the upper part of the distribution (Hardadi et al., 2021; Levinson and O'Brien, 2019), which reduces the credibility of these findings. On the other hand, the work of income distribution corrections (Blanchet et al., 2020) has recently been extended for the assessment of carbon inequality (Chancel, 2022), but this method has the weakness of assuming a constant elasticity between income and emissions, without taking into account differences in consumption patterns for different income groups.

This paper corrects the income distribution of households, and in addition predicts the consumption of rich households from income and sociodemographic characteristics. This allows for associating emission intensities for consumption categories with the consumption of households and therefore to provide a more detailed picture of emissions from different income groups. We implement this methodology for the case of Ecuador, where we have privileged access to survey and register data.

In the remainder of this study, we first provide an overview of existing methods that link emissions to income in section 2. Section 3 explains the methodological approach and the construction of the comparable scenarios. Results are presented in section 4, followed by a discussion on the limitations (section 5) and the conclusions (section 6).

2 Literature review

Over the last decades, scholars have intensively investigated how income, inequality and environmental damage are related. Since the first seminar papers (Grossman and Krueger, 1995; Boyce, 1994), the discussion has diverged into different strands: The Environmental Kuznets Curve (EKC) literature (Grossman and Krueger, 1995; Dinda, 2004; Stern, 2017) raises the question of how average income of a country is associated with the environmental impacts of the country, following the classic Kuznets Curve, which represents the relationship between average income and inequality. It is reasonable to belief that economic growth, and therefore higher material welfare, leads to greater a demand of environmental resources, but scholars still dispute whether environmental damage eventually reaches a maximum, from which environmental pressure is reduced even when the economy keeps growing (Stern, 2017). In a more recent strand, Environmental Engel Curves (EEC, Levinson and O'Brien, 2019; Sager, 2019) - sometimes also referred to micro (or household) EKCs (Heerink et al., 2001) - describe how environmental damage varies with changes in income, in the style of classic Enge Curves, which associate expenditure changes to income changes. Finally, scholars have also discussed, how income inequality within an economy drives environmental damage (Grunewald et al., 2017; Kopp and Nabernegg, 2022) to evaluate potential trade-offs between the two political goals of inequality reduction and environmental sustainability.

Despite different nuances in their focus, these strands are intertwined for different reasons. First, when predicting environmental damage (such as carbon emissions) with the level of income by estimating the EKC, results can be biased if the distribution of income is not taken into account (Heerink et al., 2001; Boyce, 1994). Second, the transmission channels from income inequality to emissions can be split up in two: The political economy mechanisms (increased social norms, less conspicuous consumption), and the underlying income-emission relationship on the household level. While political economy suggests that emissions decrease when resources are distributed more equally, a concave relationship between income and emissions on the household level means that - by definition - any redistribution from richer households to poorer households increases emissions. Consequently, the credibility of results from the EKC and EEC literature - and therefore also of the existence of a social dilemma (Sager, 2019) or a social optimum (Rao and Min, 2018) between inequality and emission reduction - depends on reliable information about the distribution of income and the distribution of emissions to the households along this income distribution.

Considerable progress has been made in improving the information on income and wealth distributions. There are two outstanding initiatives that have worked on reliable and extensive inequality indicators that are consistent with national aggregates for a broad range of countries: First, the OECD-Eurostat Expert Group on Disparities in National Accounts EG-DNA (OECD and EUROSTAT, 2020; Zwijnenburg et al., 2021), which focuses on household income and uses household surveys and imputation methods to establish coherence with national accounts for income, consumption and savings on an aggregate level without the need of tax data. Second, there is the method of Distributional National Accounts DINA (Blanchet et al., 2020) by the team of the World Inequality Database WID (Alvaredo et al., 2017), which established a procedure that uses tax microdata as main source and combines this data with information from household surveys to obtain detailed information about income and wealth of the highest percentiles.

While the strength of the EG-DNA method lies in its extensive use of household information, including detailed information of household characteristics, its level of disaggregation is limited to a reduced number of groups (e.g. quintiles, where each group represents 20% of the population). The DINA method produces synthetic micro datasets and therefore allows for a very detailed disaggregation (e.g. the top 0.1% of a distribution), but as its starting point is tax data, household surveys are primarily used for correction purposes and not to obtain further characteristics that stem from the survey.[1] The WID initiative has put great effort into the application of DINA for countries on a global level, and the difficulty to obtain official tax data has fostered the development of standardized intermediate methods and programs for countries with data limitations (Blanchet et al., 2022b; Blanchet et al., 2022a).

The correction of income has the advantage of counting on an additional data source from tax records, but finding the emissions of the missing rich is a more puzzling objective. As we will describe

[1]For a detailed comparison of both methods, see Zwijnenburg (2019).

in the methodological section, emissions are usually associated to households through their consumption and the emissions embodied in these consumption items. This means, that the consumption of the household is the linkage between income and emissions. Despite efforts to capture consumption structures of different social groups - e.g., with scanner data or structuring online shopping behavior - there is no structured additional data source for high-income households, such as for income the tax records, that allow for an integration with the existing survey data. The consumption of the missing rich is not observable.

Scholars have worked around this problem with two solutions, neither of which we consider to be satisfying. The first and most straightforward approach is to use income and consumption directly from the survey, or even to exclude the highest income percentiles from the analysis to avoid distortions due to insufficient reliable information for the upper tail of the distribution (Levinson and O'Brien, 2019; Sager, 2019). By transparently using this approach, one can apply the detailed linkage from income to consumption (when captured within the same survey), and from consumption to emissions (via emission intensities of consumption items stemming from a consumption-emission bridging matrix), which is the most comprehensive method available and is called the *bottom-up approach* (Chakravarty et al., 2009).[2] The obvious disadvantage of this approach is that nothing can be said about the emissions of the richest households.

The other approach departs from the DINA method. Lately, scholars from the WID have made attempts to also estimate emission inequalities considering the corrected income distribution. Chancel and Piketty (2015) and Chancel (2022) use income inequality estimates corrected for the missing rich from a great number of countries in the world and apply a *top-down approach* to distribute emissions along the income distribution. The association of emissions to income is achieved by the use of income-emission elasticities provided by Bruckner et al. (2022). The advantage of this second method is that no consumption data is needed as an intermediary between income and emissions. Nevertheless, the emission elasticities for different countries in Bruckner et al. (2022) are obtained from original household surveys which do not account for the missing rich. Therefore, it can not be assumed that the elasticity holds for the upper part of the income distribution.

There are reasons to belief that consumption habits change drastically for the richest groups in society. Otto et al. (2019) calls for a shift from the focus from the poor to the rich to achieve carbon mitigation goals, due to their large individual carbon impact. Gössling (2019) finds excessive emission patterns of high-income individuals, and Barros and Wilk (2021) quantify the outsized carbon footprints of the super-rich. When a society's wealthiest individuals are not covered or do not provide correct information in surveys, inequality statistics fail to reflect both their income sources and consumption habits, and therefore also their carbon emissions associated to these habits (Eckerstorfer et al., 2016; Otto et al., 2019; Levinson and O'Brien, 2019). Scholars have proposed methods to assess the gap between consumption reported in surveys and the level of consumption stemming from national accounts, also because this gap is increasing over time (Browning et al., 2014; Aguiar and Hurst, 2005; Attanasio and Pistaferri, 2016). Nevertheless, as far as we know, a method

[2]This procedure allow for a very detailed interlinkage between consumption and its embodied emissions in the first place, but also needs to integrate income and consumption information on the household level when this information does not stem from one survey, to finally produce emission inequalities based on income.

for the systematic imputation of consumption for the missing high-income households has not been developed yet. The effort of this paper is to develop such a method, which is outlined in the next section.

3 Methodology and Data

We have shown that both the bottom-up and the top-down approach face limitations. The methodology proposed here is aimed at increasing reliability of the income-emission nexus on the household level. This comes with the cost of an additional step to predict consumption for the missing rich, and is itself subject to limitations, which will be discussed in section 5.

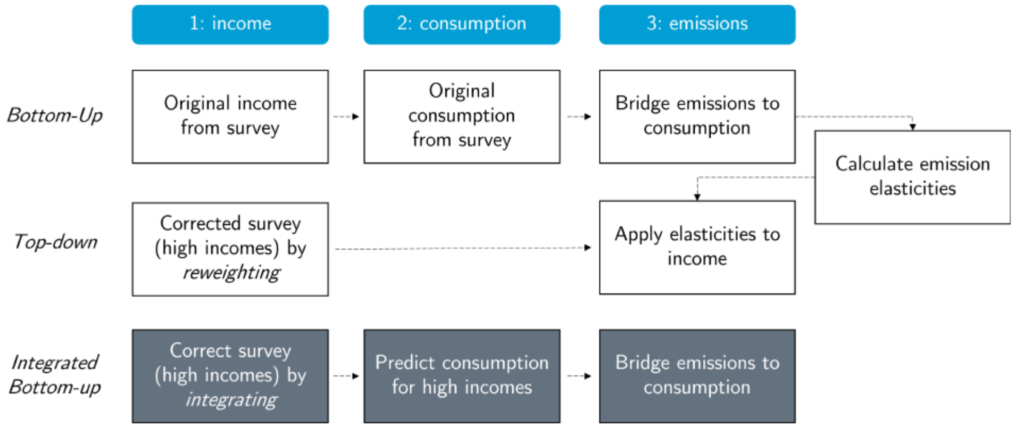
The general structure of the existing approaches and the newly proposed approach, which we call *integrated bottom-up approach* are presented in figure 1. The bottom-up and the integrated bottom-up approach follow three steps from income to emissions, while the top-down approach goes from income to emissions without the consumption steps and depends on the emission elasticities from the bottom-up approach. In the bottom-up approach (Levinson and O'Brien, 2019; Sager, 2019; Irfany and Klasen, 2017; Hardadi et al., 2021), income and consumption are taken directly from the household survey without any correction for the missing rich, or even excluding the upper tail of the distribution due to potential small sample bias in this region. For the step from consumption to emissions, this approach uses bridging tables from environmental multi-regional input-output tables (EMRIO), which contain the emission intensity of consumption goods and services. As a result, emission elasticities can be calculated, either for consumption or for income (Bruckner et al., 2022).

The top-down approach corrects the survey for the missing high-income households. There is a number of different methods to correct the distribution of income, ranging from exclusive imputation methods for surveys without any secondary data (Zwijnenburg et al., 2021; Ederer et al., 2020), to tax register based approaches that correct the distribution with survey data (Alvaredo et al., 2017), and survey based approaches that do not integrate microdata from tax records, but correct the survey with assumptions based on tax tabulates (Blanchet et al., 2022a)[3]. For the case of Ecuador, where we want to apply the integrated bottom-up method, only the most simple correction approach has been implemented yet, which is the correction of the weights of the households from the survey, such that they resemble the distribution including income stemming from tax registers (Chancel, 2022). While practical for general insights about the income distribution of a country, this method produces highly unstable results at the top of the distribution, as they may be sensitive to a handful of observations in the survey that will be assigned extremely high weights. The top-down approach does not rely on any consumption information (which makes it especially interesting for countries without a unified income and consumption survey), but rather relies on emission elasticities from other studies which are then applied to the corrected income distribution in order to obtain the emission distribution. This is, again, a very practical solution for many countries, but the assumption that emission elasticities stemming from uncorrected surveys correctly represent the income-emission relationship can be challenged.

We introduce two innovations, one for the income correction and a second for the consumption

[3] This is the case for countries with limited access to micro tax records.

Figure 1: Methods overview



Own design.

correction, which together enable to use the consumption-emission bridging. The details are presented in the following subsections.

3.1 Income correction

In the first step, the method creates a synthetic micro-dataset that is corrected for high-income households, stemming from exhaustive household survey and tax register micro-data. This follows the guidelines from the Distributional National Accounts methodology, widely applied for countries all over the world (Blanchet et al., 2020). In the integration process, high-income individuals from the tax data are introduced 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. (2022a), 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 higher percentage of household income from national accounts and depend less on additional assumptions for upscaling to economy wide totals.

3.2 Consumption prediction and emissions embodied in consumption

The second step meets 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 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). This new approach is aimed at allowing for nuances that go beyond total income as a predictor for emissions and therefore allows for different emission patterns

Table 1: Method summary

Scenario	Examples	Income correction for rich households	Consumption prediction for rich households	Emissions calculation
Bottom-up	Levinson and O'Brien, 2019; Sager, 2019; Bruckner et al., 2022	-	-	EMRIO
Top-down (Rw)	Chancel, 2022	✓ (reweighting)	-	$\varepsilon_{y,e}$
Integrated (Rp) Bottom-up		✓ (integrating)	✓	EMRIO

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 want to test whether this elasticity is not constant along the entire income distribution.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 simple statistical learning method that regresses each consumption category on income and household characteristics is applied. The model is first trained within the survey, as this source contains households with information on income and consumption. The income variables have to be the same as in the tax register (e.g., wages, capital income, social benefits). After training the model, the consumption for observations from the tax register is predicted for different categories (following the international COICOP standard classification) with the available income variables.

The third step maps the calculated household consumption onto the corresponding carbon emissions through an environmental multi-regional input-output (EMRIO) model, accounting for embodied emissions of household consumption. This allows us to add up the emissions from households to the corresponding total of the economy. One of the restrictions of this method at this stage is the lack of bridging tables for direct emissions (e.g., from transportation, heating or cooking), and which will potentially underestimate total emissions for richer households. Table 1 provides a resume for the three methodological approaches.

3.3 Data

The development of the outlined method requires the combination of a number of data sources. We have gained privileged access to data from Ecuador, where we count on all necessary micro datasets. First, income and expenditure were mutually collected by the national statistic office of Ecuador (INEC, 2013) in the National Income and Expenditure Household Survey (ENINGHUR) in 2012. This survey contains around 40,000 households with their corresponding declarations on different income sources, consumption of more than 40,000 detailed items (following the international COICOP classification) and sociodemographic characteristics. Using this combined income and consumption survey has the advantage that there are no additional steps to be taken to match independent income surveys with consumption surveys. Second, the Internal Revenue Service of Ecuador (SRI) provided access to process tax microdata from the year 2012. The information includes income tax declarations from more than 3,000,000 individuals, working as employees or independently. After integrating the survey and the tax register, and after predicting consumption, results for disposable income, final consumption and savings are scaled up to national accounts total, which stem from the Integrated Economic Tables published by the Ecuadorian Central Bank (BCE). The association of emissions to consumption is achieved with GTAP MRIO Database and the bridging matrix from (Cai and Cherkassky, 2009).

4 Results for the case of Ecuador

This section starts with the results from the income correction methods. We will see that the choice of the correction method already influences the outcome of the income-emission relationship. The second part of this section presents how the consumption of the upper tail changes with predicted consumption. The last part presents EECs, emission elasticities and inequality indicators for the three scenarios.

4.1 Income distribution results

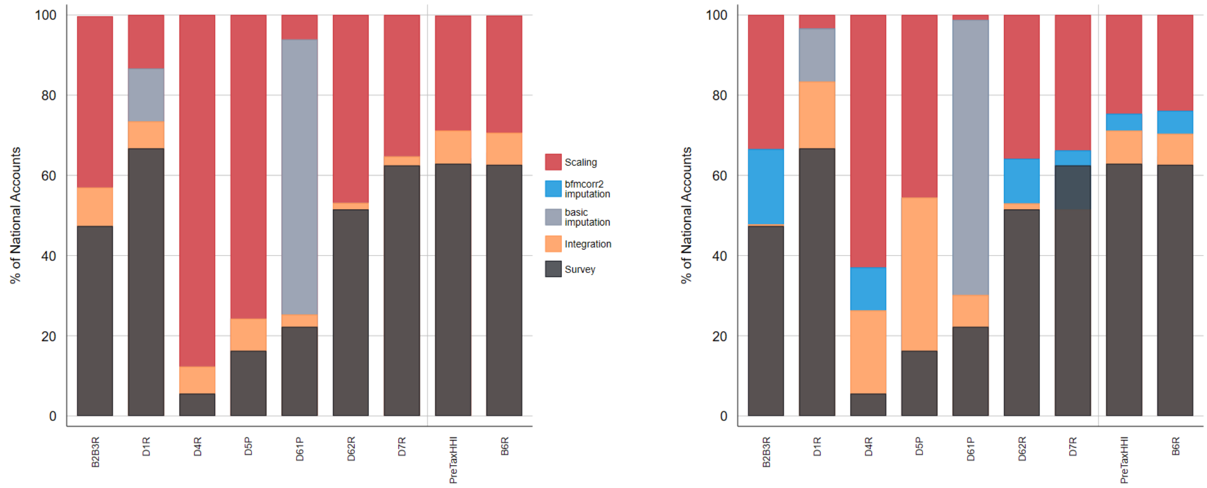
In section 3, we referred to two different methods to correct for the missing rich: reweighting survey observations and integrating tax register observations into the survey. Figure 2 indicates, how well both approaches increase the coverage of different income components, compared to the benchmark of national accounts totals. The dark brown part of the bars is the initial coverage of the household survey (and is exactly the same for figure 2a and 2b). The first seven bars represent different income components, from operating surplus and mixed income (B2B3) to other transfers (D7), while the last two bars represent pretax household income (PreTaxHHI) and disposable income (B6R). For disposable income, the survey covers slightly more than 60% of what is captured in national accounts, but there are some income components (D5P: taxes, D61P: social contributions), which hardly reach 20% or are even almost not covered by the survey at all (D4: capital income). Coverage increases with the integration of tax information (in figure 2a through reweighting existing survey observations to reproduce the tax register information, in figure 2b via the real insertion of individuals from the tax register), represented by the orange color. Together with the imputations that come directly from the survey (blue bar), the newly developed approach reaches significantly higher coverage than the reweighting approach. Finally, basic imputations reduce the gap to national accounts totals (such as imputing social contributions associated to declared income following official tax rates). What is left (red bar) has to be scaled up. Our method increases coverage considerably, compared to the reweighting method (e.g., more than 20 percentage points for capital income D4R or more than 30 percentage points for taxes), and therefore relies less on imputations and simple upscaling.[4]

For the association of income with consumption (and later with emissions), this the method uses not only total income, but information of different income components. Figure 3 displays how the income composition varies for different income groups. The left figure represents the original survey, the middle figure the reweighting approach, and the right figure the integrated approach. In all scenarios, mixed income increases for the hightest percentiles, while wage income is first increases in

[4]Taxes and social contributions are subtracted from other income components. This is why the blue bar for PreTaxHHI and B6R increases only by slightly less than percentage points.

Figure 2: Coverage with reweighting and integration correction methods

(a) Reweighting. (b) Integration.

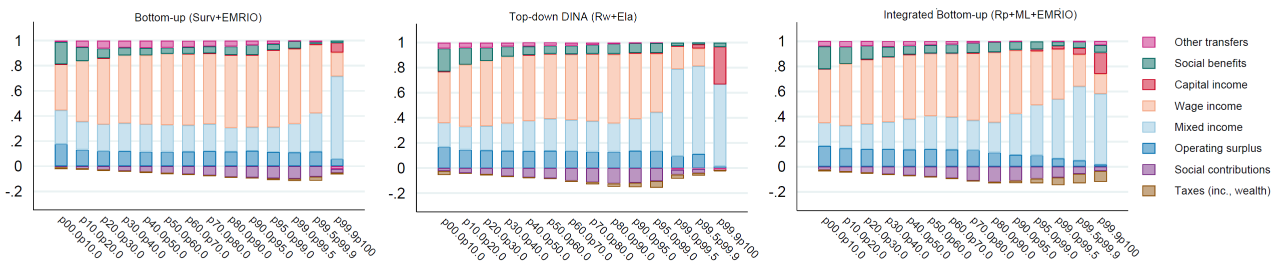


Source: Own production.
B2B3R: operating surplus and mixed income, D1R: compensation of employees, D4R: capital income, D5P: taxes, D61P: social contributions, D62R: social benefits (other than social transfers in kind), D7P: other transfers, PreTaxHHI: pretax household income, B6R: disposable income.

importance, but loses relevance in the upper tail of the distribution.[5] Capital income is absent for the lower 90% of the household distribution. Beside these similarities, figure 3 also indicates that the reweighting approach potentially introduces a small sample bias for the richest income groups. There is no wage income at all in the last group, capital income suddenly represents 30% of income, and taxes are zero. Also two groups earlier, mixed income importance raises by 40%. The integrated approach in the right figure shows smoother transitions from one income group to the next, and produces more credible results, e.g. for the last income group paying at least 10% of income taxes.

[5]Not that the last five bars in each figure represent ever smaller groups of households. The last bar represents the richest 0.1% of all households.

Figure 3: Income composition by income group



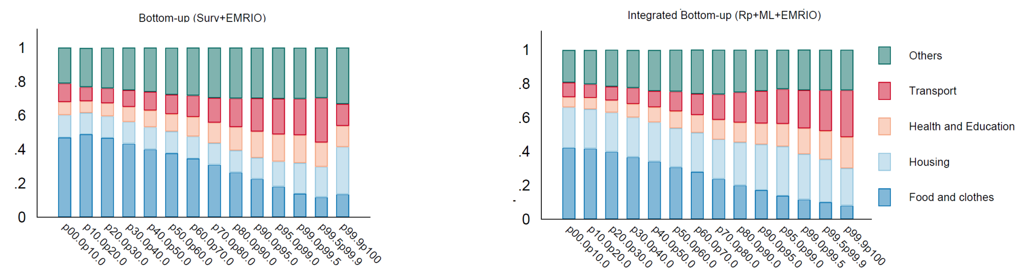
Own elaboration based on data from ENINGHUR-INEC, SRI and BCE. The bars present the percentage from different income categories as a share of disposable income in each income-percentile group. The horizontal axis shows percentiles.

4.2 Consumption prediction results

The integrated bottom-up approach predicts consumption for each household with information on different income components. The impact of predicting consumption can be observed by how the composition of consumption components changes for rich households, compared to the results from the original survey (figure 4). Both subfigures indicate the expected declining importance of food and clothes with rising income, and an increasing share of health and education, as well as transport consumption. The difference, again, is in the detailed upper part of the distribution, where sudden changes can be observed for the original survey, which disappear in the integrated method.

To additionally check for the plausibility of our consumption results, we want to assure that our method does not predict extreme values. Appendix 9a compares the expenditure of the highest household quintile from European countries in absolute terms and for components, with the expenditure from Ecuador with our method. Only the highest percentile in Ecuador reaches values above the richest 20% of Europe, which indicates that values are not disproportionate in our model.

Figure 4: Consumption categories by income group



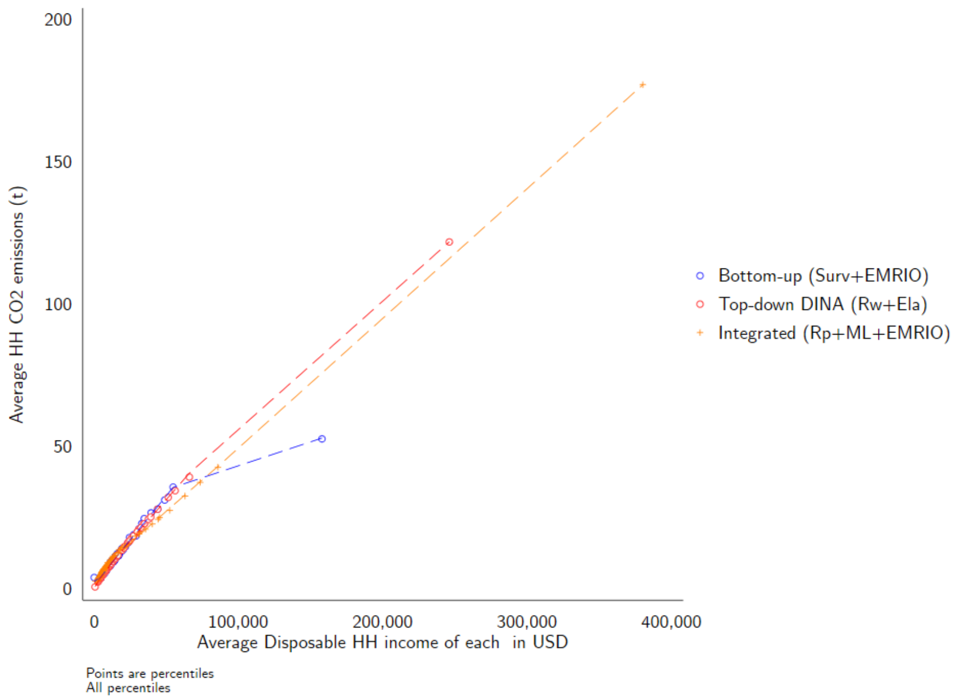
Own elaboration based on data from ENINGHUR-INEC, SRI and BCE. The bars present the percentage that is spent on different consumption categories as a share of total expenditure in each income-percentile group. The horizontal axis shows percentiles.

4.3 Environmental Engel Curves

After consumption prediction, the final step from figure 1 for each scenario is the calculation of emissions. For the bottom-up approaches, the bridging with EMRIO tables was implemented, while for the top-down approach the income-emission elasticity from the bottom-up approach was extracted and imputed to the income distribution. With all the information at hand, Environmental Engel Curves can provide insights to the relationship between income and emissions on the household level. We are especially interested, whether this relationship is concave, and for this purpose provide two EECs, with different dissaggregations: Figure 5 plots one point for each percentile and connects the points on the graph, where the x-coordinate is average disposable income and the y-coordinate the average CO2 emissions in tons. The bottom-up approach (blue color) shows linear patterns up to the 95th percentile and a marginal decrease in emission growth with increasing income only in the highest percentile. For the top-down (red color) and the integrated approach (orange color), this linearity even applies for upper tail of the income distribution.

For an even more detailed analysis of the very upper tail of the distribution, figure 6 presents more points for the the highest percentiles. Therefore, the income from the richest groups goes beyond the ones observable in figure 5. Take, for example, the integrated approach: While the richest 1% of

Figure 5: Environmental Engel curves for percentiles



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

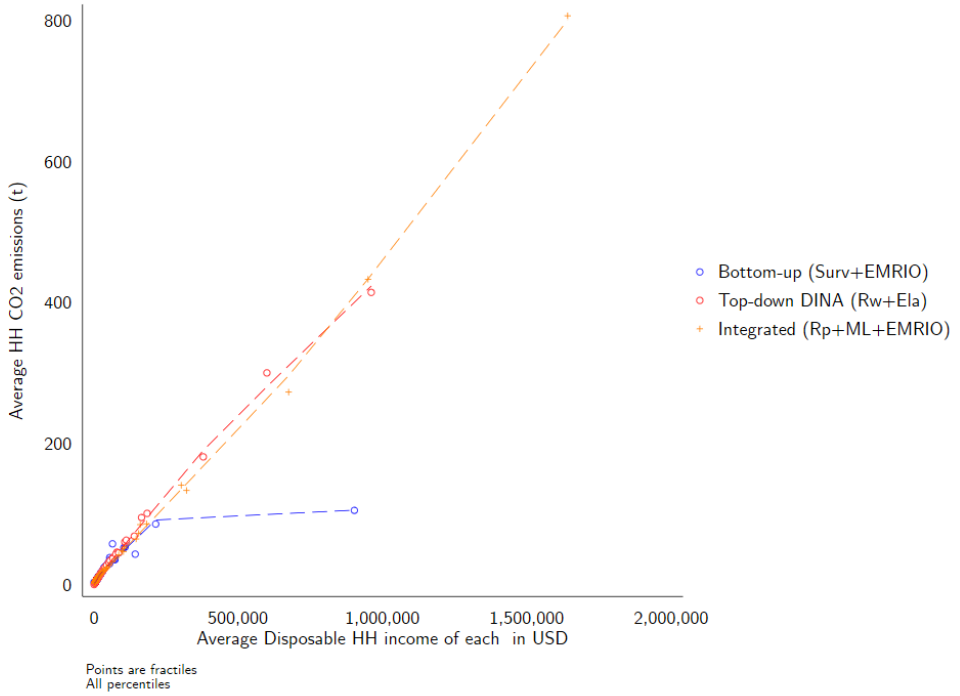
households reached average incomes of about 400,000 USD per year and CO_2 equivalent emissions of around 200t per year (rightest point in figure 5), the richest 0.01% of the population (rightest point in figure 6) had incomes of almost 2,000,000 USD and emissions of around 800t per year. This group covers only 300 households, while the richest percentile represents 30,000 households. The concavity of the EECs for the bottom-up approach (blue color) is further reduced to the most affluent group, with much volatility in emissions for the groups around 250,000 USD. It can also be observed, that in this detailed analysis, the income and emissions for the richest groups between the top-down and the integrated approach diverge, but the general linear tendency in both approaches still holds.

In a similar way, the results from the EEC can be interpreted through income-emission elasticities along the income distribution (see figure 7). The blue line for the bottom-up approach indicates relative stable elasticity up to an income between 10,000 and 100,000 USD[6], and - by definition - the top-down approach uses the average elasticity from this bottom-up approach, which lies constantly at 0.887. On the contrary, with the integrated bottom-up approach, the income-emission elasticity lies below the value from the survey (0.779), but increases for the highest income groups.

The integration method applied in step 1, in combination with consumption prediction and carbon emission mapping increases inequality indicators in general, compared to the bottom-up scenario (see figure 8). The Gini index for income jumps from 0.34 to 0.48, which is significantly higher than with

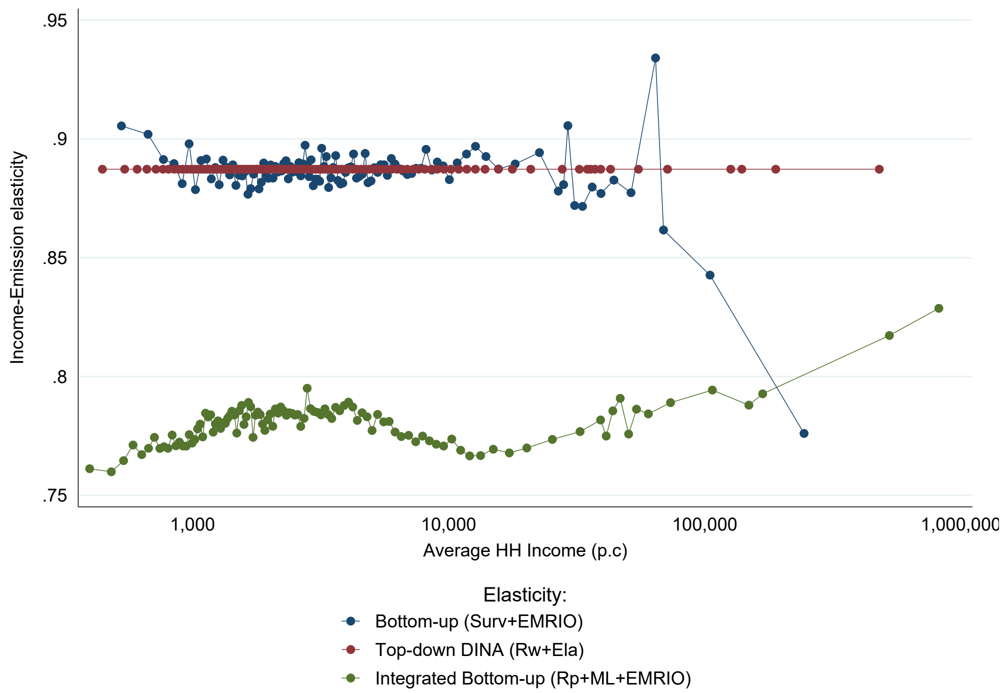
[6]Note that the x-coordinate is in log-scale.

Figure 6: Environmental Engel curves for higher dissagregation



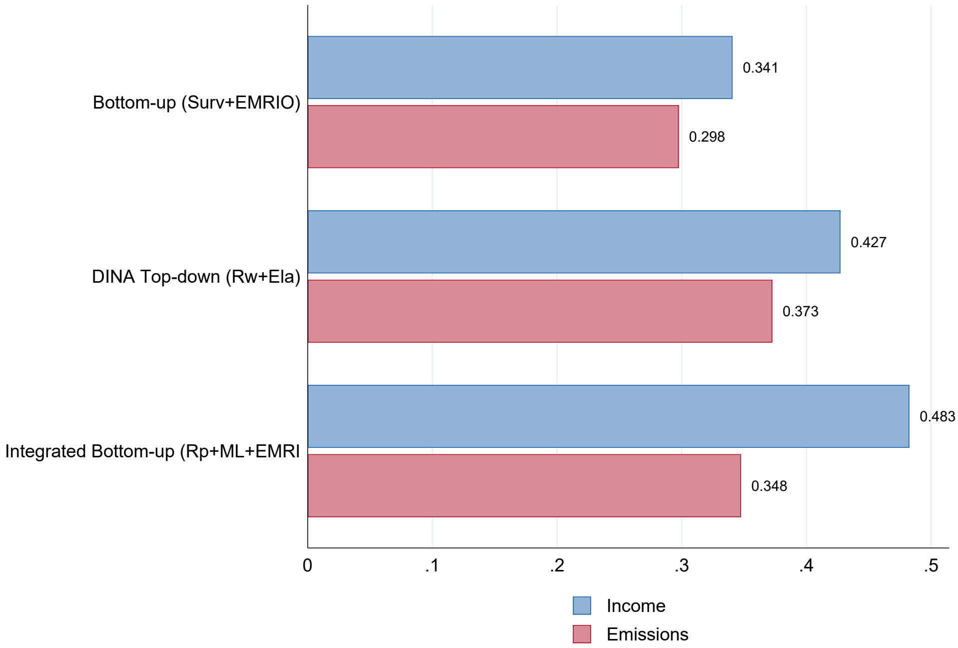
Own elaboration based on data from ENINGHUR-INEC, SRI and BCE. The figure shows disposable income on the horizontal axis, emissions on the vertical axis.

Figure 7: Income-Emission elasticities



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

Figure 8: Gini and concentration indexes for income and emissions



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

the reweighting method in the top-down approach (0.43). Nevertheless, emission concentration[7] only rises from 0.3 to 0.35 in the integrated bottom-up approach, whereas the top-down approach reaches a concentration index of 0.37.

The takeaway from the results is, that the combination of detailed income distribution correction (which increases income inequality indicators) with consumption prediction and emission bridging reduces the income-emission elasticity and therefore also the impact of higher income inequality on emission inequality.

5 Limitations

The purpose of the integrated bottom-up approach is to overcome certain limitations of existing methods in the association of emissions to income. At the same time, the approach faces some limitations itself: The first limitation concerns the extensive data requirements, as access to tax microdata from tax authorities is needed. The second limitation is the performance of the prediction method, which is a parsimony supervised statistical learning method via regressions. The use of non-linear models or unsupervised machine learning methods could potentially increase the predictive power. The inclusion of more information about the consumption of rich households, for instance from other countries or regions, could additionally increase the precision of the prediction method.

Beside the method, the study at hand is also limited in some aspects: The extensive data re-

[7]Emissions inequality is measured here through a concentration index ranked by income.

quirement restricts the application to a country with access to tax register, which is why we have only obtained results for Ecuador at this stage. Further, the most current income and expenditure survey was executed by INEC in 2012, and the survey before this stems from 2004, a year without reliable tax records. This reduces our findings to one observation for one country. Future studies in this field should amplify the geographic and temporal scope. For the case of Ecuador, we also lack information to bridge direct emissions to consumption. In further studies, this significant part of consumption should be included into the method, as missing the direct emission components probably underestimates the emissions from rich households.

6 Conclusions

This paper has proposed a new method to associate emissions to income on the household level. Limitations of the two existing methods motivated the development of this new method: One group of studies (Levinson and O'Brien, 2019; Sager, 2019) ignore the missing rich in household surveys and therefore derive results on income, consumption, and emission distributions with limited data (bottom-up approach). The other group of studies (Chancel, 2022) corrects for income but ignores potential heterogeneous patterns of consumption of the rich households by applying a constant income-emission elasticity to all households (top-down approach).

The integrated bottom-up approach presented in this paper first corrects the income distribution by integrating tax register observations into the household survey, maintaining valuable variables for income categories otherwise lost. One of the advantages of this integration method is a higher coverage of income in relation to totals from national accounts, compared to the existing methods. In a second step, consumption is predicted with a statistical learning procedure. The objective of this step is to obtain consumption patterns for the newly introduced rich households. Finally, emission intensities for consumption categories are mapped onto the consumption of each household.

With the privileged data access to household survey and tax microdata for Ecuador, we applied the method to the case of Ecuador for the year 2012. The resulting Environmental Engel Curves (EEC) and income-emission elasticities show different patterns for the three approaches: The bottom-up approach suggests a concave income-emission relationship, at least for the upper tail of the income distribution. The top-down and the integrated bottom-up approach indicate a linear relationship also for the richest income groups. This challenges the idea of a social dilemma between inequality and carbon emission reduction. On the other hand, the income-emission elasticity is significantly lower (although increasing in the last percentiles) with the integrated bottom-up approach, compared to the simple bottom-up approach that does not correct for the missing rich. Consequently, although income inequality measures increase sharply from the survey-based bottom-up approach to the integrated approach (+14 points), the concentration index for emissions only rises by 5 points.

The need for a correct understanding of the income-emission relationship is relevant for climate policy distributional policy design at the same time. Future research in this field can build upon the methodological suggestions of this study, and amplify coverage in terms of space an time.

Bibliography

AGUIAR, Mark and Erik HURST (2005). "Consumption versus expenditure". *Journal of Political Economy* 113.5, pp. 919–948.

ALVAREDO, Facundo, Lucas CHANCEL, Thomas PIKETTY, Emmanuel SAEZ, and Gabriel ZUCMAN (2017). "Global inequality dynamics: New findings from WID. world". *American Economic Review* 107.5, pp. 404–09.

ATTANASIO, Orazio P and Luigi PISTAFERRI (2016). "Consumption inequality". *Journal of Economic Perspectives* 30.2, pp. 3–28.

BARROS, Beatriz and Richard WILK (2021). "The outsized carbon footprints of the super-rich". *Sustainability: Science, Practice and Policy* 17.1, pp. 316–322.

BLANCHET, Thomas, Lucas CHANCEL, Ignacio FLORES, and Marc MORGAN (2020). "Distributional National Accounts (DINA) guidelines: Concepts and methods used in the World Inequality Database". *updated and extended version of the DINA Guidelines first published as WID Working Paper 2016/2*.

BLANCHET, Thomas, Lucas CHANCEL, Amory GETHIN, et al. (2019). "How unequal is Europe? Evidence from distributional national accounts, 1980-2017". *WID. world working paper* 6.

BLANCHET, Thomas, Ignacio FLORES, and Marc MORGAN (2022a). "The weight of the rich: improving surveys using tax data". *The Journal of Economic Inequality* 20.1, pp. 119–150.

BLANCHET, Thomas, Juliette FOURNIER, and Thomas PIKETTY (2022b). "Generalized pareto curves: Theory and applications". *Review of Income and Wealth* 68.1, pp. 263–288.

BOYCE, James K (1994). "Inequality as a cause of environmental degradation". *Ecological economics* 11.3, pp. 169–178.

BROWNING, Martin, Thomas F CROSSLEY, Joachim WINTER, et al. (2014). "The Measurement of Household Consumption Expenditures". *Annual Review of Economics* 6.1, pp. 475–501.

BRUCKNER, Benedikt, Klaus HUBACEK, Yuli SHAN, Honglin ZHONG, and Kuishuang FENG (2022). "Impacts of poverty alleviation on national and global carbon emissions". *Nature Sustainability* 5.4, pp. 311–320.

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". *World Development* 152, p. 105783.

CAI, Feng and Vladimir CHERKASSKY (2009). "SVM+ regression and multi-task learning". *2009 International Joint Conference on Neural Networks*. IEEE, pp. 418–424.

CHAKRAVARTY, Shoibal, Ananth CHIKKATUR, Heleen DE CONINCK, Stephen PACALA, Robert SOCOLOW, and Massimo TAVONI (2009). "Sharing global CO2 emission reductions among one billion high emitters". *Proceedings of the National Academy of Sciences* 106.29, pp. 11884–11888.

CHANCEL, Lucas (2022). "Global carbon inequality over 1990–2019". *Nature Sustainability*, pp. 1–8.

Chancel, Lucas and Thomas Piketty (2015). “Carbon and inequality: From Kyoto to Paris Trends in the global inequality of carbon emissions (1998-2013) & prospects for an equitable adaptation fund World Inequality Lab”.

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

Dinda, Soumyananda (2004). “Environmental Kuznets curve hypothesis: a survey”. *Ecological economics* 49.4, pp. 431–455.

Eckerstorfer, Paul, Johannes Halak, Jakob Kapeller, Bernhard Schütz, Florian Springholz, and Rafael Wildauer (2016). “Correcting for the Missing Rich: An Application to Wealth Survey Data”. *Review of Income and Wealth* 62 (4). issn: 14754991.

Ederer, Stefan, Predrag Ćetković, Stefan Humer, Stefan Jestl, and Emanuel List (2020). “Distributional national accounts (dina) with household survey data: Methodology and results for european countries”. *Review of Income and Wealth.*

Gössling, Stefan (2019). “Celebrities, air travel, and social norms”. *Annals of Tourism Research* 79, p. 102775.

Grossman, Gene M and Alan B Krueger (1995). “Economic growth and the environment”. *The quarterly journal of economics* 110.2, pp. 353–377.

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

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”. *Journal of Industrial Ecology* 25.1, pp. 95–113.

Heerink, Nico, Abay Mulatu, and Erwin Bulte (2001). “Income inequality and the environment: aggregation bias in environmental Kuznets curves”. *Ecological Economics* 38.3, pp. 359–367. issn: 0921-8009.

INEC (2013). *Encuesta Nacional de Ingresos y Gastos de los hogares urbanos y rurales 2011- 2012, Resumen Metodológico y Principales Resultados.*

Irfany, Mohammad Iqbal and Stephan Klasen (2017). “Affluence and emission tradeoffs: Evidence from Indonesian households’ carbon footprint”. *Environment and Development Economics* 22, pp. 1–25.

Klasen, Stephan, Katrin M Meyer, Claudia Dislich, Michael Euler, Heiko Faust, Marcel Gatto, Elisabeth Hettig, Dian N Melati, I Nengah Surati Jaya, Fenna Otten, César Pérez-Cruzado, Stefanie Steinebach, Suria Tarigan, and Kerstin Wiegand (2016). “Economic and ecological trade-offs of agricultural specialization at different spatial scales”. *Ecological Economics* 122, pp. 111 –120. issn: 0921-8009.

Kopp, Thomas and Markus Nabernegg (2022). “Inequality and Environmental Impact–Can the Two Be Reduced Jointly?” *Ecological Economics* 201, p. 107589.

Leach, Melissa, Belinda Reyers, Xuemei Bai, Eduardo S Brondizio, Christina Cook, Sandra Díaz, Giovana Espindola, Michelle Scobie, Mark Stafford-Smith, and Suneetha M Subramanian (2018). “Equity and sustainability in the Anthropocene: A social–ecological systems perspective on their intertwined futures”. *Global Sustainability* 1, e13.

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

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

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

Otto, Ilona M, Kyoung Mi Kim, Nika Dubrovsky, and Wolfgang Lucht (2019). “Shift the focus from the super-poor to the super-rich”. *Nature Climate Change* 9.2, pp. 82–84.

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

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

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

Sager, Lutz (2019). “Income inequality and carbon consumption: Evidence from Environmental Engel curves”. *Energy Economics* 84. issn: 01409883.

Stern, David I (2017). “The environmental Kuznets curve after 25 years”. *Journal of Bioeconomics* 19, pp. 7–28.

Veblen, Thorstein (1899). *The theory of the leisure class.* New York: Macmillan.

Zwijnenburg, Jorrit (2019). “Unequal distributions: EG DNA versus DINA approach”. 109, pp. 296–301.

Zwijnenburg, Jorrit, Sophie Bournot, David Grahn, and Emmanuelle Guidetti (2021). “Distribution of household income, consumption and saving in line with national accounts: Methodology and results from the 2020 collection round”. *OECD Statistics Working Papers 2021/01 No. 108.*

A Characteristics of data sources

Characteristics of data sources are displayed in table (2).

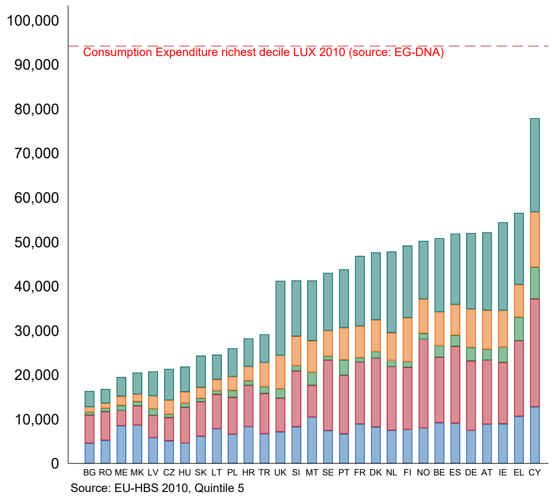
Table 2: Characteristics of data types

Data type	Example	Completeness/ representativeness	Level of disaggregation	Context information
Macro aggregate	National Accounts, Input-Output tables	high	low	low
Household survey	Labor market surveys, Expenditure surveys	middle (high for group A, low for group B)	high	high
Administrative record	tax record, social security records, social assistance records	middle (high for group B, low for group A)	high	middle

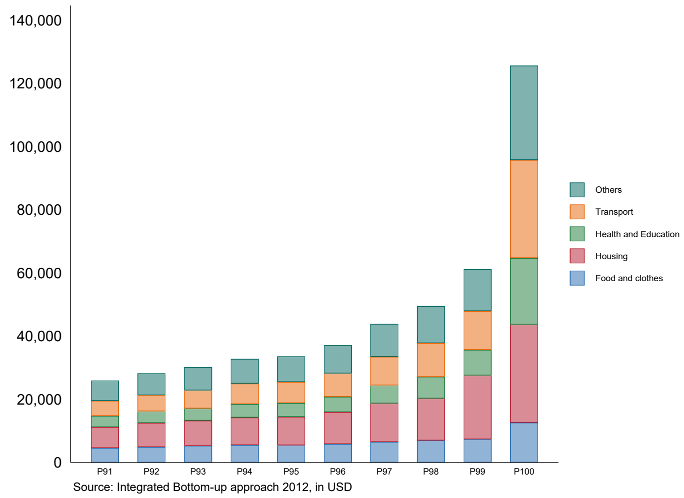
B Robustness of consumption prediction results

Figure 9: Comparison to European countries

(a) Highest quintile European countries.



(b) Percentiles Ecuador.



Expenditures of the highest household quintiles from European countries and all deciles from Ecuador with our method.

C Robustness of emission results

Figure 10: Comparison to super rich