Grantham Research Institute on Climate Change and the Environment

Centre for Climate Change Economics and Policy

Income inequality and carbon consumption: evidence from environmental Engel curves

Lutz Sager

November 2017

Centre for Climate Change Economics and Policy Working Paper No. 319
ISSN 2515-5709 (Online)

Grantham Research Institute on Climate Change and the Environment
Working Paper No. 285
ISSN 2515-5717 (Online)

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

The Grantham Foundation for the Protection of the Environment

University of Leeds

LSE The London School of Economics and Political Science

The Centre for Climate Change Economics and Policy (CCCEP) was established by the University of Leeds and the London School of Economics and Political Science in 2008 to advance public and private action on climate change through innovative, rigorous research. The Centre is funded by the UK Economic and Social Research Council. Its second phase started in 2013 and there are five integrated research themes:

1. Understanding green growth and climate-compatible development
2. Advancing climate finance and investment
3. Evaluating the performance of climate policies
4. Managing climate risks and uncertainties and strengthening climate services
5. Enabling rapid transitions in mitigation and adaptation

More information about the Centre for Climate Change Economics and Policy can be found at: www.cccep.ac.uk.

The Grantham Research Institute on Climate Change and the Environment was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute. It has six research themes:

1. Sustainable development
2. Finance, investment and insurance
3. Changing behaviours
4. Growth and innovation
5. Policy design and evaluation
6. Governance and legislation

More information about the Grantham Research Institute on Climate Change and the Environment can be found at: www.lse.ac.uk/grantham.

Income Inequality and Carbon Consumption: Evidence from Environmental Engel Curves

Lutz Sager

London School of Economics and Political Science

November 2017

Abstract: This paper analyses the relationship between the distribution of income and the carbon dioxide content of household consumption. Household carbon is estimated by linking expenditure data to productive sectors and their carbon intensity derived through input-output analysis. Environmental Engel curves (EECs) are estimated, which describe the relationship between household income and CO_2 in the United States between 1996 and 2009. A second-degree polynomial specification in income is found to approximate well the fit of more flexible nonparametric models. These parametric EECs are used to decompose the within-year household carbon inequality as well as the evolution of household carbon over time. In both cases, household income appears to be a main driver of carbon consumption. A potential "equity-pollution dilemma" is described and a method to quantify it is proposed. Assuming (conditional) homogeneity in preferences, EEC estimates predict that progressive income transfers would raise household carbon by 5.1% at the margin and by about 2.3% under complete income redistribution in 2009.

Keywords: Income, consumption, pollution, redistribution.

JEL codes: D12, D31, E21, H23, Q52.

Address: Grantham Research Institute on Climate Change and the Environment, Department of Geography and Environment, LSE, Houghton Street, London, WC2A 2AE, United Kingdom; telephone: +44-20-7107-5027; e-mail: l.sager@lse.ac.uk.

Acknowledgements: I gratefully acknowledge financial support by the Grantham Foundation for the Protection of the Environment and the UK's Economic and Social Research Council (ESRC). For valuable comments, I am thankful to Frank Cowell, Simon Dietz, Angela Druckman, Roger Fouquet, Ian Gough, Ben Groom, Antony Millner, as well as participants at the LSE Climate Change, Inequality & Social Policy Seminar and the LSE PhD Seminar on Environmental Economics. All remaining errors are my own.

1 Introduction

Income inequality has been rising in many developed countries since the 1970s and its consequences are today the focus of much research (for an overview see e.g. Atkinson et al., 2011). At the same time, manmade climate change is now recognised as a major threat to well-being and sustainable development in the long-run. This paper aims to improve the understanding of the *interplay between the distribution of income within a country and the environmental burden related to household consumption.*

In a first step, we estimate the carbon dioxide (CO_2) content of the consumption baskets of a sample of households in the United States, covering the period between 1996 and 2009. We then estimate Environmental Engel curves (EECs), which represent household carbon at different positions in the income distribution. Just like EECs for air pollutants (Levinson and O'Brien, 2015), we find EECs for CO_2 to be upward-sloping, concave, and shifting downwards over time. We then demonstrate the usefulness of both nonparametric and regression-based estimates of EECs for further analysis.

We first use nonparametric EECs to derive suggestive evidence of the contributions of technology, income growth and expenditure dynamics to trends in aggregate household carbon. We then exploit parametric estimates for EECs for a more systematic decomposition of the evolution of average household carbon over time and the distribution of household carbon with a given year. We find that income (and even more so total expenditure) is the main driver of household carbon both over time and between households within time. Meanwhile, other household characteristics appear to play only a minor role in shaping household carbon.

This regression-based decomposition based on quadratic EEC estimates is a useful addition to the existing literature on consumption-based household carbon footprints and their drivers, which has often relied on more descriptive analyses and single estimates of income elasticities (e.g. Weber and Matthews, 2008; Buechs and Schnepf, 2013). We demonstrate that a second-order polynomial specification for EECs approximates well the relationship between income and household carbon found by higher-order polynomial models and more flexible nonparametric estimation techniques.

We then consider the *consequences of income redistribution for consumption patterns and household carbon.* Much of the existing research assessing the relationship between the distribution of income and the environment has focused on how social groups are differentially affected by environmental pressures, adding a layer of environmental inequalities often related to economic ones. Growing evidence points to regressive effects of both local environmental externalities, such as air pollution (Currie and Neidell, 2005; Holland et al., 2016), as well as global ones, such as climate change (Mendelsohn et al., 2006; Hsiang et al., 2017). This paper is interested in the inverse of that relationship, asking if and how the distribution of income affects aggregate environmental outcomes.

Based on the observation of concave EECs, we formulate and quantify what we call the *"equity-pollution dilemma" – namely that positive income redistribution may raise aggregate household carbon.* To the best of our knowledge, this is the first attempt to quantify this dilemma using microdata on household consumption within a single country. It thus builds on the literature which resulted from the initial formulation of the dilemma by Scruggs (1998) and proposed empirical investigations using cross-country analyses following Heerink et al. (2001). We propose a simple method to quantify the "equity-pollution dilemma" which relies on the quadratic specification of EECs as well as the dispersion measure known as Gini's mean difference. Assuming (conditional) homogeneity in preferences, we predict that income transfers would raise household carbon by 5.1% at the margin and by about 2.3% under complete income redistribution in 2009. For hypothetical scenario under which the distribution of household incomes in the United States is distributed in a similar fashion to that in Sweden, we predict an increase in household carbon of about 1.5%. The estimated magnitude of the "equity-pollution dilemma" is larger for CO_2 than for two other greenhouse gases - methane (CH_4) and nitrous oxide (N_2O) - which we also analyse. We hope that the proposed metric for the "equity-pollution dilemma" will inspire future work assessing the relationship between the distribution of income and environmental burden using microdata across different countries, time periods and pollutants.

The rest of this paper is structured as follows. Section 2 reviews the previous literature. Section 3 discusses the methodology and data used. Section 4 presents evidence from nonparametric EECs, while Section 5 presents quantitative results from regression-based, parametric EECs. Section 6 quantifies the "equity-pollution dilemma". Section 7 concludes.

2 Previous literature

In this paper, we investigate the relationship between the distribution of income, the consumption decisions of individual households, and the carbon content of that consumption. In doing so, we contribute to two growing literatures. The first literature is the one asking how income inequality within a country affects aggregate greenhouse gas emissions (and environmental burden more broadly). The second literature is concerned with accounting for the carbon footprint of household consumption, assessing its distribution over households, and understanding its principal drivers.

Distributional causes of environmental pressure:

This paper adds to an emerging literature assessing the potential contribution of economic inequality to growing environmental pressures caused by economic activity. The existing literature has focused on two channels through which the shape of the income distribution in an economy may affect environmental outcomes – through *consumer choice* or *political economy* dynamics.

The first channel builds on the observation that the level and composition of aggregate consumption result from a combination of consumer preferences and budgets. This transmission channel was first proposed by Scruggs (1998) and then formalised by Heerink et al. (2001). Essentially, the observation that consumers at different income levels allocate varying budget shares to different product categories, leads to the proposition that *redistribution of income will change the composition of aggregate consumption and in consequence the environmental burden linked to it.*

The second transmission channel relies on a political economy perspective. It presupposes that environmental policy is the result of differential political power and tastes along the income distribution (Boyce, 1994). From that perspective, the distribution of income reflects differences in political influence between groups of varying concern for the environment.

However, existing empirical evidence does not support a systematic relationship between inequality and pollution (see survey by Berthie and Elie, 2015). Baek and Gweisah (2013) find a positive association between income inequality (measured as Gini index) and per capita CO_2 emissions in the United States for different years between 1967-2008. Meanwhile, Heerink et al. (2001) find a negative association between the Gini index and per capita CO_2 emissions across 180 countries in the period 1961-2001. For air pollution, Torras and Boyce (1998) find a positive association between inequality (Gini) and air pollution levels in a number of cities and countries between 1977-1991.

Results from these studies are rather mixed, and appear to vary with choice of pollution type (air, water, waste, etc.), regional scale of analysis, timing and empirical specification. It is worth mentioning the inherent limitations to drawing inference about the relationship between income inequality and aggregate pollution from such cross-country studies. Arguably, both the degree of income inequality and the pollution attributed to a country respond to a variety of structural, cultural, economic, and political factors.

This paper contributes to that literature by relating consumer choice to environmental outcomes *within one country*. It builds on the empirical literature concerned with estimating the pollution intensity of household consumption using microdata.

Consumption-based household carbon accounting:

Over the past decades, research into the greenhouse gas (GHG) emissions attributable to individual countries, regions, sectors, firms and households has been growing rapidly. Spurred on by international efforts to mitigate GHG emissions, most countries have by now implemented detailed accounting for GHG emissions *produced* within their territory. More recently, *consumption-based* GHG accounting has grown in popularity (Davis and Caldeira, 2010). As opposed to territorial or production-based GHG accounting, consumption-based GHG accounting attributes the emissions embedded in a good produced in country A but consumed in country B to the account of the latter. A key motivation for consumption-based emissions accounting is the quantification of so-called "carbon leakage", describing the carbon emissions embedded in trade between producing and consuming countries (see surveys by Wiedmann, 2009; Sato, 2014).

At the micro-scale, a growing literature is aiming to quantify the carbon content of individual products (e.g. Tukker and Jansen, 2006) or of the consumption basket of households *within a country* (e.g. Weber and Matthews, 2008). The latter is the approach most relevant to this paper, as we are aiming to relate the income and socio-economic characteristics of individual households to the carbon content of their consumption.

Similar to the literature on "carbon-leakage" at the economy level, the literature quantifying the greenhouse gas content of individual households' consumption baskets is motivated by a consumer responsibility perspective (Druckman et al., 2008; Lenzen, 2008). That literature has thus far focused on understanding the drivers of emissions as contained in household consumption (Weber and Matthews, 2008; Buechs and Schnepf, 2013) and quantifying the "rebound effect" (Thomas and Azevedo, 2013; Chitnis et al., 2014).

A key finding of that literature is that measures of consumption-based GHG emissions are increasing with income. For example, Weber and Matthews (2008) construct measures of household carbon footprint (HCF) based on expenditure data from the Consumer Expenditure Survey in the United States. They find that income and household expenditure are the strongest predictors of the HCF, with high income households generating more than 10 times the emissions of low income ones. Findings are similar for studies that focus only on certain portions of household consumption, such as fossil fuel use (Papathanasopoulou and Jackson, 2009) or the energy content of household consumption (Lenzen et al. 2006). Some further factors that have been found to predict household emission budgets are household size, age, employment status, educational attainment, urban vs. rural location, and the quality of housing stock (for a recent survey of the literature see Druckman and Jackson, 2016).

We contribute to this line of work by analysing in detail the distribution of household carbon in the United States. Our main contribution to the literature exploring the drivers of household carbon is that *we decompose the variation in household carbon into the respective contributions of socio-economic characteristics*. This regression-based decomposition analysis is possible because we rely on the concept of Environmental Engel curves (proposed by Levinson and O'Brien, 2015), which is introduced below. This provides a prototype in moving beyond descriptive statistics and income elasticity estimates used in the literature so far.

A related literature has used estimates of consumption-based household carbon footprints and especially its association with household income to derive estimates of the global distribution of greenhouse gas emissions. Policy implications derived include the allocation of global carbon reduction targets to nations according to the principle of "common but differentiated responsibilities" (Chakravarty et al., 2009) and highlight the disproportionate responsibility on the part of the rich independent of nationality (Chancel and Piketty, 2015).

Another insight emerging from this literature is that household carbon is not a linear function of income, but that households tend to increase budget shares of less carbon intensive goods as they become richer (e.g. Buechs and Schnepf, 2013; Chitnis et al., 2014). This finding has important implications for the likely welfare effects of environmental policy such as pollution taxes. It is often argued that carbon taxes will be regressive by disproportionally affecting poorer households who will be harder hit from price increases to carbon-intensive necessities such as heating fuel (e.g. Pearce, 1991; Grainger and Kolstad, 2010). Similarly, knowing the carbon content of certain types of consumption baskets can help inform the feasibility of emissions targets given current technologies. For example, Druckman and Jackson (2010) estimate minimal GHG emissions requirements based on "minimum income standard" budgets needed to provide a "decent life".

The exact shape of the relationship between income and household carbon is still debated in the literature. Early contributions hypothesised an inverted U-shaped relationship between household income and the pollution intensity of consumption (Kahn, 1998; Heerink et al., 2001). More recent empirical evidence shows that the pollution burden per unit of expenditure is indeed decreasing in income, suggesting concavity if not an inverted U-shape (e.g. Liu et al., 2013; Buechs and Schnepf, 2013). In the literature on consumption-based CO_2 emissions, this observation is usually summarised by an expenditure elasticity of CO_2 below 1, with most estimates between 0.8-1.0 (Chakravarty et al., 2009). In this paper, *we will go beyond a single estimate of income elasticity and demonstrate the usefulness of estimating Environmental Engel curves* – which describe more fully the carbon content of demand schedules as they are related to income.

This approach explicitly allows for income elasticities of demand to differ at various income levels in line with recent evidence on energy services which constitute an important portion of household carbon budget. Fouquet (2014) estimates long-run income elasticities for energy services (domestic heating, lighting, passenger transport) and finds income elasticities which are rising at lower levels of incomes up to a certain point and subsequently tend towards zero. Similar trends can be observed in our data when assessing expenditure shares of energy services at different points of the income distribution (Figure A.4 in the Appendix). It is apparent that energy services in aggregate represent a slightly growing budget share in total expenditures at low levels of household income and only exhibit diminishing budget shares at household incomes above about USD 40k. The composition of expenditures on energy services reveals further interesting patterns (Figure A.5). While electricity can clearly be described as a necessity (shrinking expenditure shares all along the income distribution), gasoline appears to be a luxury good at incomes below USD 50k and only exhibits clearly diminishing budget shares at incomes above USD 100k. Our point estimates of concave Environmental Engel curves are consistent with such "saturation effects".

Environmental Engel curves:

We use parametric estimates of Environmental Engel curves (EECs) for decomposition analyses and to construct a measure for the degree to which income redistribution may affect aggregate emissions embedded in consumption.

I doing so, we follow Levinson and O'Brien (2015), who construct EECs describing the relationship between income and air pollutants embodied in the consumption of households in the United States. They focus on PM_{10}, but find similar results for VOC, NO_x, SO_2 and CO. EECs are useful visualisations of the income-pollution relationship. Levinson and O'Brien (2015) find EECs for air pollutants to be upward sloping and concave.

A key contribution of this paper is that *we estimate parametric EECs for CO_2 emissions embedded in the consumption of households in the United States between 1996 and 2009*. Similar to Levinson and O'Brien (2015), we also find the carbon EECs to be upward sloping and concave.

Parametric estimation of EECs as proposed by Levinson and O'Brien (2015) opens up a range of avenues for more theoretical considerations based on empirical estimates from consumption microdata. In this paper, we use estimates of EECs to generate insights into the relationship between the distribution of income and aggregation consumption-based carbon emissions. We demonstrate that simple parametric EECs that include a quadratic term for income (i.e. second-degree polynomial) match well the relationship estimated using more flexible nonparametric methods. One advantage of the parametric (quadratic) specification is that it makes possible the decomposition of household carbon inequality by contributing factors, decomposing the evolution of average household carbon over time, and quantifying the potential trade-off between income redistribution and emissions reduction. Our results yield systematic evidence of income being a main driver of household carbon, both over time and between households within time.

The "equity-pollution dilemma":

Much empirical work remains to discover shapes of EECs for different types of pollutants, in different economic contexts and across time. As we demonstrate below, EECs change over time with the composition of consumption and production technologies. Analytically, the concavity of EECs may have important consequences for redistributive considerations. As discussed above, it has been a long-standing argument that mitigation policies may be regressive by disproportionally raising prices for carbon-intensive necessities with income elasticities below 1 (Pearce, 1991; Grainger and Kolstad, 2010; Gough, 2013). We focus on the flip-side of this, which we call the "equity-pollution dilemma":

Given the higher pollution intensity of consumption per unit of expenditure by poorer households, progressive redistribution may result in higher aggregate pollution from consumption.

Based on the constructed EECs for household carbon, we assess whether or not the "equity-pollution dilemma" is likely to hold and what might be its magnitude. We use the derived EECs to illustrate under which assumptions an "equity-pollution dilemma" may arise. We propose a method to quantify the "equity-pollution" dilemma based on parametric EECs using consumption microdata for households within one country. We hope that this adds to the literature concerned with the inequality-pollution relationship, which often relies on single income elasticity estimates and cross-country data (Scruggs, 1998; Heerink et al., 2001).

It is noteworthy here that concavity of EECs which pass through the origin implies an income elasticity below 1. However, we believe that the analysis throughout this paper demonstrates the usefulness of estimating the shape of EECs in more detail, rather than focusing on a single estimate of income elasticity.

We thus see three major contributions of this paper. In a first instance, we generate estimates of consumption-based household carbon for the United States between 1996 and 2009 in the form of Environmental Engel curves. These estimates are useful tools for descriptive analyses, such as separating the contributions to changes in emission over time from changes in technologies, savings rates, and the composition of consumption. Secondly, we demonstrate how parametric estimates of EECs can be used for regression-based decomposition of household carbon over time and within time. Thirdly, we rely on estimates of quadratic EECs to generate a simple formula for the quantification of the "equity-pollution dilemma" under (conditionally) homogenous preferences.

3 Data and methodology

We construct Environmental Engel curves (EECs) for carbon dioxide (CO_2) contained in the consumption of households in the United States. We focus on households in the United States, because it has some of the highest consumption-based CO_2 per household (e.g. Chancel and Piketty, 2015). At the same time, detailed data are available on the income and consumption patterns of households. We estimate the CO_2 attributable to the consumption of all energy, fuels, goods and services by households at different positions in the income distribution. The focus of this exercise is thus on *total emissions* contained in household consumption. This includes *direct* emissions from the consumption of fossil fuel based energy (e.g. heating, electricity, transportation fuels) as well as *indirect* or "embedded" emissions from the production of goods and services consumed.

We then combine information on yearly expenditures of households on different consumption items (in dollars) with estimates of the carbon intensity of these different goods and services (kg of CO_2 per dollar) to construct EECs following the methodology proposed by Levinson and O'Brien (2015). Our emissions accounting methodology is based on Environmentally-Extended Input-Output Analysis as is standard in the literature on consumption-based greenhouse gas emission accounting (Wiedmann, 2009).

Data:

Information on household income, consumption expenditures and socio-demographic characteristics comes from the United States Consumer Expenditure Survey (CEX). The Bureau of Labor Statistics provides anonymised public use micro-data from 1996. We make use of the interview portion of the CEX, containing information on survey responses by "consumer units" (CU). In what follows, we will refer to "consumer units" as households. Our main source of information is the collection of "monthly expenditures" files (MTBI), which contain information on a household's expenditures (and incomes) split into over 800 categories assigned universal classification codes (UCC). We combine these with income and socio-demographic characteristics contained in the "consumer unit characteristics and income" files (FMLI). To allocate emissions intensities to consumption categories, information from the World Input-Output Tables (WIOD) is used. WIOD contains information on 35 production sectors in 40 countries. Notably, WIOD publishes "Environmental Accounts", which include information on emissions and gross output per sector. We use these to allocate to each sector a *direct emissions intensity* (CO_2 per \$ output).

Estimation of carbon content of consumption:

We use the input-output portion of WIOD to attribute to each sector a *total emissions intensity*, taking into account the full chain of intermediate inputs from other sectors ad infinitum. This is done, following the procedure proposed by Leontief (1970) assuming a linear relationship between the sector outputs and the required inputs (i.e. linear production function and constant returns to scale). The *total emissions intensity* (kg of CO_2 per USD output) of a sector is what we refer to below as production *technology*.

The CEX consumption expenditures are then each allocated to one WIOD production sector. It is in this step, where a number of judgements by the researcher are necessary. We follow where possible the matching procedure used by Levinson and O'Brien (2015) to link UCC to IO codes used in the input-output tables of the Bureau of Economic Analysis[1]. Appendix A.1 contains a detailed description of the procedure including the assumptions necessary to arrive at a complete matching of expenditure categories to production sectors[2]. Table A.1 lists the 34 WIOD sectors and estimated emissions intensities for the years 1996 and 2009.

Multiplying the consumption expenditures of a household with the matched *total emissions intensity* yields a rough estimate of the CO_2 embedded in the yearly consumption of that household, which we shall call *estimated household carbon / CO_2*.

Direct emission factors for high-carbon goods:

To improve the precision of our estimates, we allocate emissions intensities to certain high-carbon consumption categories directly. We do so for expenditures on home electricity, heating oil, natural gas, gasoline for vehicles (incl. Diesel and motor oil), and air travel. Data on end consumer prices for electricity, heating oil, natural gas, and gasoline are provided by the U.S. Energy Information Administration (2017). Emissions factors for gasoline, heating oil, natural gas, and kerosene are those used by the U.S. Environmental Protection Agency in guidelines for the Greenhouse Gas Inventory (EPA, 2009). The emissions intensity of residential electricity is taken from the EPA's Emissions & Generation Resource Integrated Database (EPA, 2017). An overview of the resulting emission factors used is given in Appendix Table A.2.

[1] We are grateful to Arik Levinson and James O'Brien for kindly sharing their matching from UCC categories to IO codes used in their forthcoming paper and for answering our questions regarding their methodology. As there are many more UCC categories than IO sectors, the matching procedure applied by Levinson and O'Brien (2015) relies on a number of subjective judgements, which they outline in an online appendix to their forthcoming paper.

[2] Matching of CEX UCC codes to WIOD sectors to be provided as online appendix for eventual publication.

We believe that this methodology significantly improves the precision of our estimates of household carbon embedded in consumption. The implementation of direct emission factors for these consumption categories increases aggregate household carbon by about 25% (e.g. from 25.0t on average with only WIOD factors to 31.0t with added direct emission factors in 2009).

Limitations and refinements:

Input-output based accounting for consumption-based CO_2 emissions is by now a common methodology. The major advantage of the approach is that, in theory, it allows for a comprehensive account of emissions related to all types of expenditures by a household. A key weakness of this method is that it cannot account for systematic differences in price/quality of goods consumed. In our methodology, $5 spent on a premium organic loaf of bread will be estimated to have five times the CO_2 content than $1 spent on a more mass-market industrial loaf. Assuming that the consumption of goods from the same category but with higher price-per-CO_2 ratio is generally increasing with income, we may thus underestimate the concavity of EECs.

Furthermore, some input-output based emissions accounting assumes a closed economy and ignores international trade, assuming instead in the calculation of emissions intensities that the value chain of all goods is entirely based within the United States. This might introduce a bias in final estimates of consumption-based CO_2 emissions, especially if the content of traded inputs into a sector is large. This is likely true for certain sectors, as the literature on embedded carbon in trade has highlighted (surveyed in Sato, 2014). For example, Weber and Matthews (2008) estimate that approximately 30% of CO_2 emissions from US household consumption occurred outside the US. This can matter for our analysis especially if we suspect that households at different income levels might consume goods with different import shares. To overcome this limitation, we rely on the multi-region input-output (MRIO) tables included in WIOD to explicitly account for both a global supply chain and trade in final goods.

Global supply chains: Estimates of carbon intensities of consumption by US consumers are derived accounting for the global nature of production supply chains. This can be problematic if certain goods rely on a higher portion of intermediate goods from countries where those sectors are relatively more (or less) emissions intensive than in the United States. We resolve this issue by expanding the input-output analysis described above is applied to the 34 WIOD sectors in 41 countries (including the United States and "rest of the world").

This results in estimates of the emissions intensities of the 34 WIOD sectors in the United States, but taking into account intermediate inputs from 1394 (41x34) WIOD sectors around the world. The procedure is described in more detail in Appendix A.1

Trade in final goods: In addition to the global nature of supply chains (i.e. trade in intermediate goods), misguided estimates may arise when a certain share of final goods consumed by US households is directly imported from other countries. We exploit information contained in WIOD on "final consumption expenditure by private households" to take into account the share of final demand by US consumers per WIOD sector that is demanded from countries outside of the United States. The inclusion of global supply chains raises average estimates of household emissions by about 7.4% in 2009 as compared to the closed economy assumption (from 31.0t to 33.3t), while the consideration of trade in final goods adds another 1.8% (from 33.3t to 33.9t). However, these changes have slightly different effects at different points of the income distribution, with a higher proportional effect for higher income households as shown in Figure A.1 in the Appendix.

Other greenhouse gases: Finally, other greenhouse gases, such as methane (CH_4) and nitrous oxide (N_20), are usually emitted alongside CO_2, but in much smaller quantities. Excluding these gases may introduce a bias in the analysis if their relationship with income and consumption systematically differs from CO_2 for certain types of consumption (e.g. food). We thus complement the methodology to include emissions of CH_4 and N_2O, information on which is also contained in WIOD Environmental Accounts. Finally, we generate an overall measure of greenhouse gas emissions by converting CH_4 and N_2O into CO_2 equivalents (CO_2e) based on their respective global warming potential. On average, this raises estimates of household greenhouse footprint by about 42% in 2009, this time with a slightly higher increase for low-income households. Further detail on the methodology can be found in Appendix A.1.

Notwithstanding remaining limitations of the methodology, we believe that this approach yields a useful first estimate of household carbon.

Final sample:

We supplement data on expenditures and estimated CO_2 with further information on household income, composition and socio-economic characteristics taken from the FMLI interview files of CEX. Households are surveyed in five consecutive quarter-yearly interview rounds. There are thus different waves of households starting the survey procedure in every quarter of every year. To generate yearly cross-sections, we assign households to the year in which their 2nd interview took place, independent of the specific date of the interview. To obtain the most representative mapping from household income to expenditures to emissions, we limit our sample to those observations for which a complete record from five interviews is available. We further limit our sample to those households classified by CEX as "complete income reporters"[3]. Our final sample then consists of 51,265 households, surveyed between 1996 and 2009, that completed 5 quarterly interviews, and for which both expenditures and reported incomes are available for 12 months preceding the fifth interview. Only households with a positive reported annual after-tax income are included to avoid distortion from those households declaring financial losses. Due to lacking information at the upper tail of the income distribution, we limit the sample to households with after-tax income below USD 400k (real 2009). Table 1 provides summary statistics of select key variables in the final sample.

Table 1: Summary Statistics

	(1) N	(2) mean	(3) sd	(4) min	(5) max	(6) Gini (2009)
Income before tax (k$)	51,265	54.88	50.73	0.00100	510.1	0.45
Income after tax (k$)	51,265	51.86	47.09	0.00100	389.0	0.44
Expenditure (k$)	51,265	42.14	35.98	2.439	1,411	0.33
HH CO_2 (kg, closed)	51,265	34,371	18,992	515.8	435,572	0.28
HH CO_2 (kg, open)	51,265	36,915	20,919	627.7	479,490	0.28
HH CO_2 (kg, open+trade)	51,265	37,574	21,545	656.3	517,434	0.29
HH CH_4 (kg, open+trade)	51,265	320.5	182.8	5.284	6,206	0.29
HH N_2O (kg, open+trade)	51,265	11.59	6.252	0.0890	105.9	0.28
HH GHG (kg CO_2e, open+trade)	51,265	51,927	29,039	915.5	759,985	0.28
Age (HH head)	51,265	51.63	16.85	15	94	
Family size	51,265	2.586	1.496	1	14	
Population weight	51,265	15,882	5,940	460.4	81,398	
Year	51,265	2,003	4.109	1,996	2,009	

Notes: Estimates for household emissions contained in consumption expenditure according to methodology described. All other variables from the US Consumer Expenditure Survey.

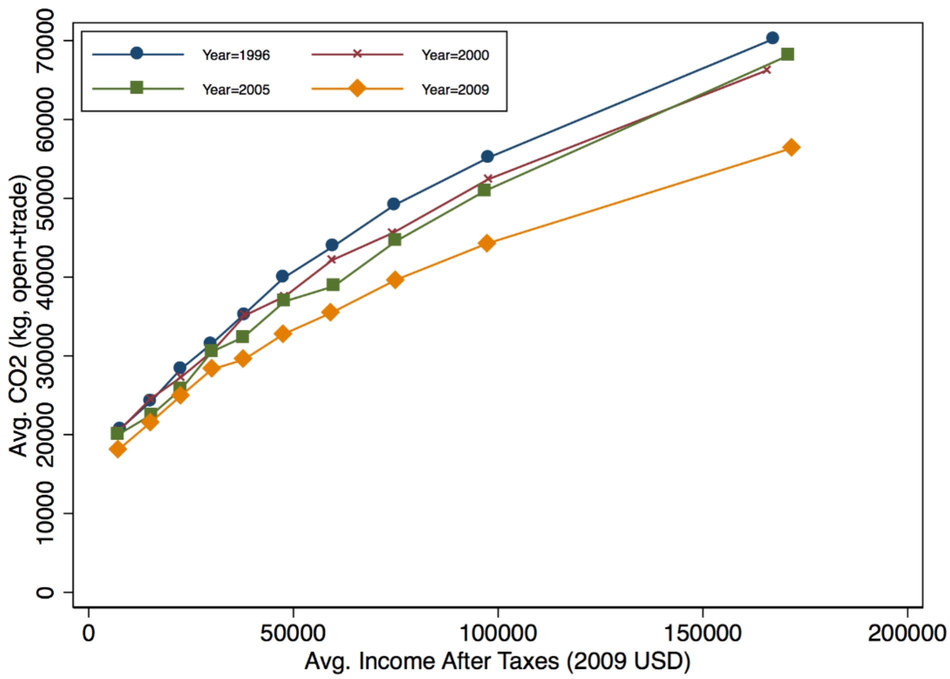
[3] The CEX data contains imputed values for incomes of those not considered to be "complete reporters" from 2004 onwards. To ensure comparability, we limit our sample to "complete reporters" throughout.

4 Descriptive Environmental Engel curves

Following the methodology proposed by Levinson and O'Brien (2015), we construct both parametric and nonparametric estimates for the Environmental Engel curves (EECs) for consumption-based carbon dioxide emissions. The advantage of the nonparametric approach is that it does not impose any functional structure on EECs, and thus is a natural starting point for descriptive analysis.

Figure 1 presents nonparametric estimates of the EECs. It represents the estimated CO_2 contained in the yearly consumption expenditures of households at different positions in the income distribution. Households are divided into income deciles, for which average after-tax incomes and CO_2 are calculated. The CO_2 content of consumption reported here is that based on calculations considering a global supply chain and including direct imports of final goods ("open+trade"). A breakdown of household carbon in 2009 by major consumption categories and information on other greenhouse gases can be found in Appendix Figure A.1. To avoid confusion with the more involved nonparametric smoothing techniques applied below, we shall call these "descriptive" EECs.

Figure 1: Descriptive Environmental Engel curve – Household CO_2



Notes: Decile averages of household income after tax (2009 USD) and estimated CO_2-content of consumption (current technology). Household weights as provided by CEX sample. Households with negative reported after-tax income are excluded.

We display descriptive EECs for the years 1996, 2000, 2005, and 2009. Figure 1 visually suggests the following characteristics of consumption-based carbon:

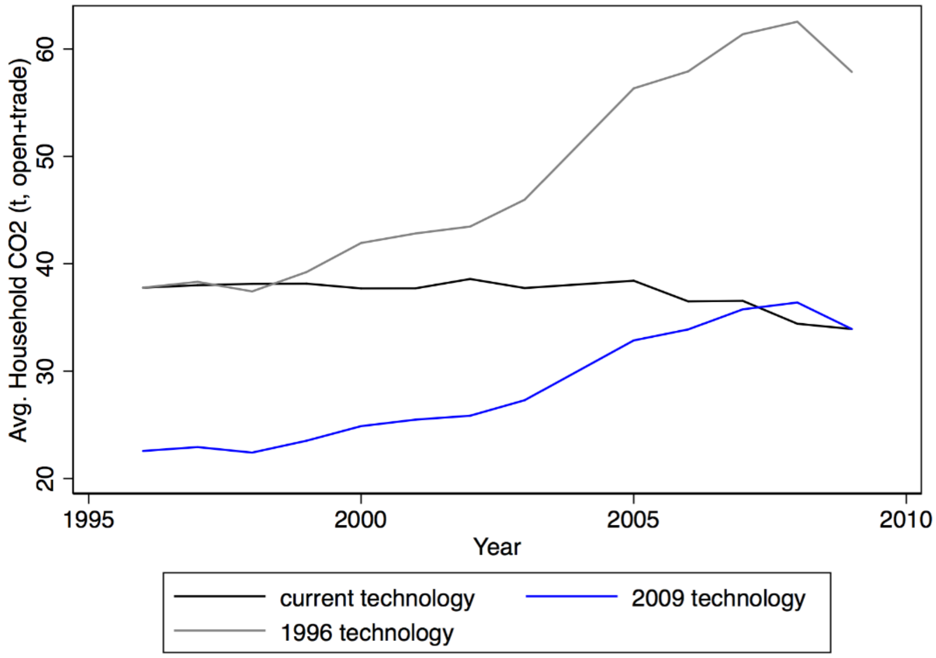
1) *EECs are increasing*: The average households higher up in the income distribution are responsible for significantly more CO_2 contained in their consumption. For example, in 1996 we estimate a carbon content of 21t for the yearly consumption of the average household in the bottom decile, while the number for the top decile is over 70t.
2) *EECs are concave:* Households with higher income have on average a less carbon-intensive consumption mix, i.e. the carbon intensity of the average dollar spent is decreasing with income.
3) *EECs shift down over time:* The average carbon-content of consumption decreases with time across the income distribution. For example, the average CO_2 embedded in the consumption of the top income decile was reduced from 70t in 1996 to ca. 56t in 2009. Two effects might contribute to this shift:
    a. Composition effect: Consumers are shifting to a less carbon-intensive mix
    b. Technology effect: Carbon intensity (kg/USD) is decreasing in most industries

These observations are in line with those made by Levinson and O'Brien (2015) about EECs for air pollutants. Moreover, our estimates for consumption-based household carbon are broadly in line with previous estimates. For example, Weber and Matthews (2008) estimate an average pollution intensity of aggregate consumption of 0.7 kg CO_2/\$ in the US in 2004. Our aggregate average in that year is 0.82 kg CO_2/\$ (0.68 kg CO_2/\$ when using only WIOD-based emission factors).

The role of income growth, technology, and consumption composition:

Descriptive EECs make possible a range of insights. Following Levinson and O'Brien (2015), we will here decompose the aggregate CO_2 embedded in US household consumption into three effects: income growth and changes in income distribution (shifts along the EECs), changes in expenditure levels per unit of income, and composition/technology effects (shifts of the EECs). We note that, had technologies not improved, the consumption of the average household would be responsible for significantly more CO_2 than at current technologies. Figure 2 shows exactly that. It compares the actual CO_2 content of the consumption of the average household (at current technologies) to hypothetical estimates assuming constant 1996/2009 technologies (i.e. carbon-intensities, in kg per \$ of final demand). For example, the average household carbon of 2009 consumption levels would have been linked to 57.9t of CO_2 if technology had not improved since 1996 (instead of 33.9t at current technology). However, improvements in technology have outweighed these dynamics, and average household carbon at current technology has decreased from 37.8t in 1996 to 33.9t in 2009.

Figure 2: Technology improvement

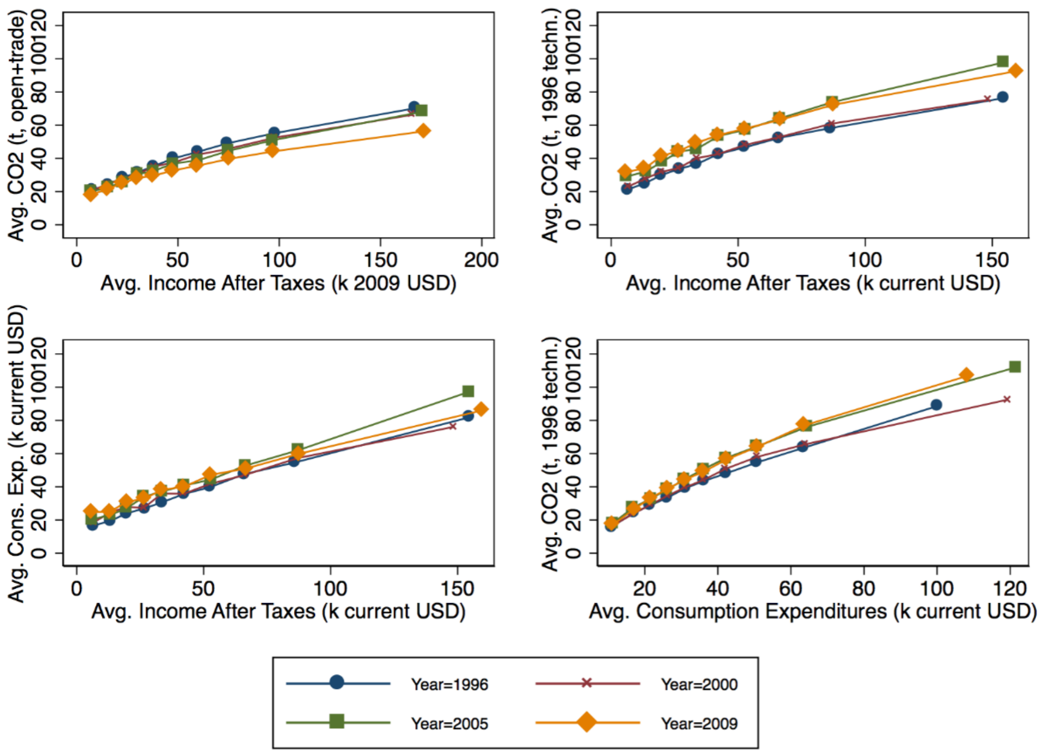


Notes: Averages of estimated CO_2-content of consumption (current technology, constant 1996, and constant 2009 technology). Household weights as provided in CEX sample. Households with negative reported after-tax income and income above USD 400k excluded.

This perspective only highlights the changes in the technology dimension and cannot account for income growth and changes in the composition of expenditures. EECs are a useful tool to disentangle these dynamics. Figure 3 compares different representations of the EECs for the years 1996, 2000, 2005, and 2009. The top left panel plots the EECs based on current technologies and real household income (2009 dollars). It is equivalent to Figure 1 discussed above. In the top right panel, we repeat the decile-based estimation of EECs relative to household income. However, here we hold the technology constant at 1996 levels. This comparison makes apparent that had technology not changed (in the sense of significant reductions in carbon-content per dollar output in most WIOD sectors), EECs would have shifted upwards. Clearly, without significant reductions in the emissions intensity of production, current consumption of US households would be responsible for significantly higher levels of CO_2 across the income distribution.

Figure 2 also illustrates that, when in this paper we refer to a change in the emissions intensity of goods as "technology", this includes price variations for example in the price of oil. For example, the observation that emissions would have been higher in 2008 at 2009 emission factors (blue line above dark grey line in 2008), is driven by the strong decline in oil prices between 2008 and 2009, which resulted in an increase of emission factors for gasoline, heating fuel and natural gas (this is not observed when using WIOD factors only). More broadly, in the case of fossil fuel combustion, changes in *technology*, i.e. variation in direct emissions intensities (kg of CO_2 per USD of output), are largely driven by changes to retail prices rather than gains in combustive efficiency.

Figure 3: Descriptive Engel curve variations – Technology and savings



Notes: Decile averages of household income after tax (current USD and constant 2009 USD), household consumption expenditure (2009 USD), and estimated CO_2-content of consumption (current technology and constant 1996 technology). Household weights as provided in CEX sample. Households with reported after-tax income below USD 10k excluded.

This increase in the CO_2 content of consumption can have two explanations: (a) households with the same nominal income spend more on carbon-intensive goods (according to 1996 technology), and (b) households at a given income level spend more on aggregate. Indeed, when comparing aggregate dollar consumption expenditures to aggregate dollar after-tax incomes (bottom left panel), it becomes apparent that nominal spending[4] is higher for

[4] It is important here to mention that throughout this paper we refer to as expenditures/spending only those expenditures that we have linked to WIOD sectors and thus to a carbon intensity. Significant portions of consumer spending that may be left out are for example the acquisition of housing via mortgages or debt-financed purchases of vehicles.

households with the same nominal income in 2009 than it was in 1996. However, even when accounting for this difference in aggregate spending (or savings rates), there appears to be a compositional effect. In the bottom right panel, we plot EECs relative to nominal aggregate consumption expenditures. It is apparent that, even for the same level of aggregate expenditures (and assuming the same emissions intensities), households consumed more carbon-intensive mix of goods in 2009 than in 1996.

The above analysis has shown that different representations of EECs can provide useful evidence on structural changes over time in consumption and its carbon content. A key insight is that there has been a significant downward trend the emissions intensity of consumption - what we refer to as *technology*. Keeping technology constant, income (and expenditure) growth appears to be a main driver of household carbon over time. Furthermore, we observe a compositional shift in the emissions intensity of expenditure (holding technology constant). While visual inspection of difference versions of descriptive EECs is clearly a useful first step of analysis, it is limited in its potential to disentangle the relative importance of these trends.

Below, we will investigate these suggested insights further using systematic decomposition analysis relying regression-based estimates of EECs.

5 Parametric Environmental Engel curves

Above, we showed that descriptive (or nonparametric) EECs are useful tools for comparison of consumer behaviour and its environmental burden between income groups and over time. Of course, the consumption pattern of a given household will not only depend on the income available (as an approximation of the budget set), but also on the needs, attitudes and habits of the household members (preferences). It is likely that households at different positions of the income distribution will also differ with respect to other characteristics related to consumer preferences. Obvious examples of household characteristics that vary with income and may influence consumption plans are household size, education, location (e.g. local weather, infrastructure, and culture), and many more (e.g. Buechs and Schnepf, 2013). To account for some of this heterogeneity, we turn to parametric estimation of EECs based on a linear regression model:

$$y_{it} = \beta_{1t} m_{it} + \beta_{2t} m_{it}^2 + \boldsymbol{x_{it}}' \boldsymbol{\delta_t} + \varepsilon_{it} \quad (1)$$

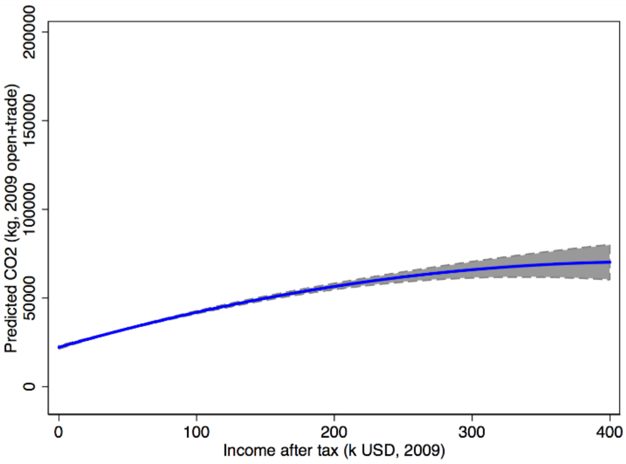
For each yearly cross-section of CEX data, we run a linear regression using estimates of the consumption-based CO_2 emissions y_{it} of household i living in year t as the dependent variable. Independent variables include after-tax household income m_{it} (real 2009 USD), its square, and a vector of household characteristics $\boldsymbol{x_{it}}$. We should note that this approach does not presuppose a model of causal relationships, but is simply a tool to elucidate partial linear associations between the variables of interest. The advantage of using a linear regression model will become apparent in subsequent analyses presented below, which will make possible the decomposition of changes in household carbon into contributing factors such as income and expenditure growth, the decomposition of inequality of household carbon, and the quantification of the "equity-pollution dilemma" based on a quadratic term for income.

Quadratic vs. nonparametric fit:

The inclusion of a term for squared income in a linear regression model is a standard ad hoc procedure when nonlinear relationships with income are suspected. However, to account for the possibility of a more complex relationship between income and household carbon, we compare the fit of our quadratic specification with a semiparametric one. We control for the same set of covariates in a linear fashion and then fit a nonparametric Gaussian kernel weighted local polynomial to describe the relationship between after tax income and household carbon[5]. Results of these two approaches are presented in Figure 4. The left panel presents the fitted values of the quadratic specification (Figure 4a) and 95% confidence intervals (relying on Huber-White heteroscedasticity-robust standard errors). The right panel (Figure 4b) compares the quadratic model with the nonparametric fit.

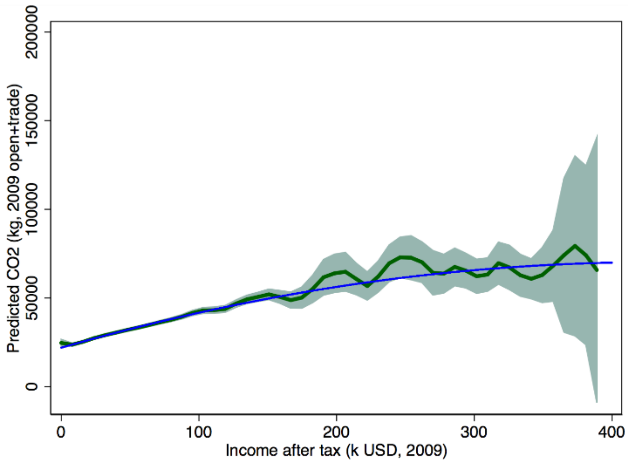
Figure 4: Environmental Engel curves – CO_2 – 2009

Figure 4a: Quadratic fit



Note: Blue = fitted values of quadratic model (holding other covariates constant at mean);Grey = 95% confidence intervals

Figure 4b: Nonparametric fit



Note: Green = fitted values of semiparametric model & 95% confidence intervals; Blue = fitted values of quadratic model

To test the appropriateness of a quadratic specification in income (polynomial of degree 2), we implement a test for equivalence between a parametric (polynomial) and nonparametric models as proposed by Hardle and Mammen (1993). Table 2 represents the results for the 2009 sample and different degrees of polynomial fit. The null hypothesis each time is that the polynomial adjustment of degree *n* is appropriate. We are thus looking for the lowest degree of polynomial for which we clearly fail to reject the null hypothesis. As can be seen from Table 2, this is the case for the quadratic model.

Table 2: Goodness of fit – Nonparametric vs. polynomial

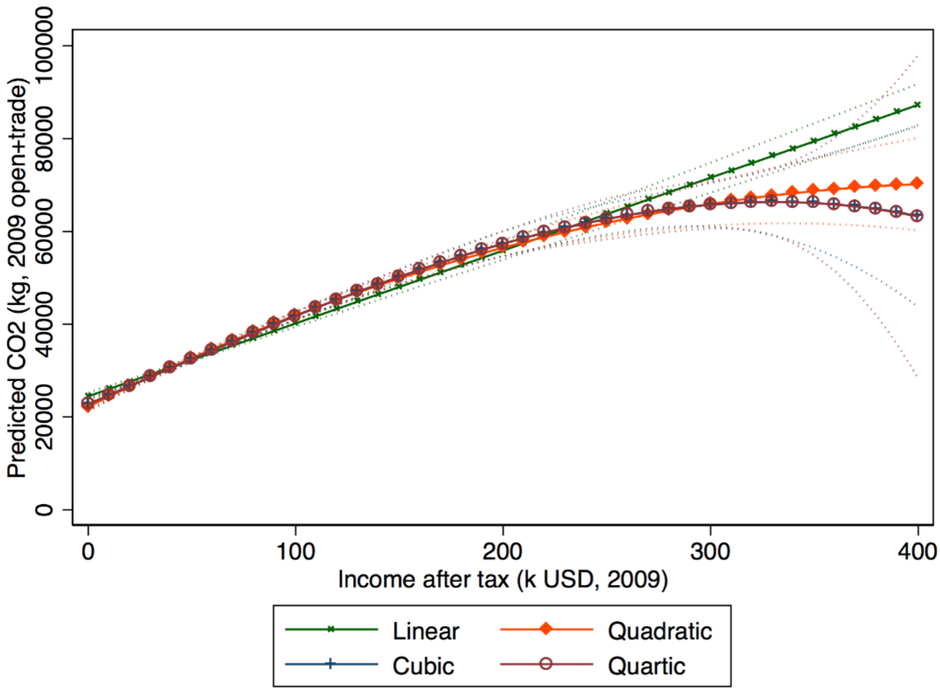
Polynomial degree tested	(0) None	(1) Linear	(2) Quadratic	(3) Cubic	(4) Quartic
T test (standardised)	26.395***	1.911*	0.792	0.770	0.596
[p value]	[0.00]	[0.09]	[0.73]	[0.84]	[0.97]

Notes: Hardle and Mammen (1993) test for goodness of fit of polynomial adjustment; different polynomial degrees by column; 2009 data.
*** p<0.01, ** p<0.05, * p<0.1.

[5] The semiparametric specification includes the following linear covariates: family size, family size (squared), age of HH head, age (squared), marital status, education, race, region. Estimates are derived using the Stata module SEMIPAR, which estimates Robinson's (1988) double residual estimator.

This is confirmed visually by Figure 5, which compares the change in model fit when moving to higher-order polynomials. It is visible how the quadratic model (red) diverges significantly from the predictions of the linear model. However, higher-order polynomials, which include a cubic and quartic term, do not seem to deviate significantly from the fit of the quadratic specification.

Figure 5: Engel curves – Quadratic vs. higher-order polynomial (2009)



Notes: Fitted values of multiple linear regression models including polynomial terms (of orders 1 through 4) for income after tax. Covariates are family size, family size (squared), age of HH head, age (squared), marital status, education, race, region. Dotted lines mark 95% confidence intervals using heteroscedasticity robust standard errors.

We interpret results presented in Table 2 and Figures 4 and 5 as evidence that the quadratic specification used throughout this section is an adequate approximation capturing a large portion of the relationship between after tax income and household carbon after controlling for covariates. We now turn to estimation of this quadratic model and applications making use of parametric EECs.

Parametric (quadratic) Environmental Engel curves:

Table 3 presents parameter estimates from the model specified in (1) for survey years 1996 and 2009. In line with the nonparametric representation of Engel curves, the results support EECs for consumption-based CO_2 which are upward sloping ($\hat{\beta}_{1t} > 0$) and concave ($\hat{\beta}_{2t} < 0$).

While household characteristics other than income appear to be associated with household carbon, the signs and magnitudes of the income coefficient estimates remain similar when controlling for these characteristics (Columns 2 and 4 respectively). This is important, because it indicates that differences in the composition and carbon intensity of consumption between households with different incomes are not primarily due to structural differences between these households (e.g. education levels). With regards to a potential "equity-pollution dilemma", this would indicate that an income transfer from a richer to a poorer household might add to aggregate CO_2 emissions even when holding constant the households' other characteristics.

We will show below that estimates of the coefficient for quadratic income $\hat{\beta}_{2t}$ are useful to characterise the magnitude of the "equity-pollution dilemma". Inclusion of socio-demographic controls thus significantly reduces the estimated magnitude of the dilemma. Of course, when assessing the impact of policies targeting inequality in the long-run, such as through education policy, the nonparametric EECs might provide a more appropriate vision, as in the long-run household incomes and other characteristics (education, size, environmental awareness, etc.) are largely co-determined.

Table 3: Parametric estimates of quadratic EECs (1996 / 2009)

	1996		2009	
	(1) OLS (income)	(2) OLS (full)	(3) OLS (income)	(4) OLS (full)
Income (k USD, after tax)	597.537***	397.392***	333.674***	223.187***
	(30.6475)	(33.8508)	(12.5338)	(13.3885)
Income squared (k USD, after tax)	-1.264***	-0.566**	-0.538***	-0.258***
	(0.2389)	(0.2478)	(0.0571)	(0.0571)
Family size		7,224.712***		6,045.746***
		(721.2440)		(640.5012)
Family size squared		-531.372***		-390.455***
		(96.7207)		(89.3228)
Age of household head		882.973***		602.852***
		(83.5928)		(68.3003)
Age squared		-7.216***		-4.566***
		(0.7774)		(0.6224)
Married (binary)		3,017.720***		3,498.022***
		(727.3970)		(516.9155)
Race (Black)		-4,538.612***		-2,222.663***
		(833.7596)		(625.6325)
Race (Native American)		-4,061.459***		-3,850.197
		(1,517.4194)		(2,381.0824)
Race (Asian / Pacific)		-6,459.371***		-3,523.863***
		(1,242.5257)		(1,202.1452)
Race (Pacific Islander)				-5,189.483**
				(2,595.3759)
Race (Multi-race)				3,073.647
				(2,920.3731)
Education (below high school)		1,543.111**		1,527.981**
		(758.2393)		(595.5659)
Education (high school)		3,874.106***		3,552.079***
		(804.1852)		(612.2637)
Education (some college/vocational)		4,583.578***		3,130.905***
		(979.1615)		(743.7333)
Education (college degree or higher)		3,360.628**		3,048.080***
		(1,425.7927)		(1,113.4889)
Region (Midwest)		-147.868		-2,074.284***
		(792.3300)		(631.2095)
Region (South)		1,582.209**		-499.459
		(800.7617)		(604.0257)
Region (West)		-1,986.629**		-2,938.677***
		(846.8349)		(682.1159)
Constant	18,110.522***	-17,674.053***	17,360.021***	-10,358.121***
	(686.5682)	(2,350.5535)	(446.1919)	(2,047.4550)
Observations	3,069	3,069	4,407	4,378
R-squared	0.450	0.552	0.402	0.506

Notes: Estimates from linear regression. Household weights as provided in CEX sample. Households with reported after-tax income below USD 10k excluded. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

The role of income growth, expenditure, and consumption composition:

Above, we have discussed the evidence (Figure 3) based on nonparametric EECs which suggests that increased carbon-content of household consumption between 1996 and 2009 was due to increases in income, but also due to changes in expenditure per unit of income and the composition of consumption. We will now quantify these effects using Oaxaca-Blinder decomposition, which was initially suggested to decompose wage differentials between population groups (Oaxaca, 1973; Blinder, 1973).

Table 4: Movement along parametric EECs - CO_2 (1996 vs. 2009)

	Change due to movement along EECs	
	(1)	(2)
Income after tax	4.9*	
Income squared	-1.0*	
Expenditure		7.7*
Expenditure squared		-0.8*
Family size	-0.1	0.0
Family size squared	0.1	0.0
Age	1.0*	0.8*
Age squared	-0.7*	-0.6*
Married	0.0	0.0
Race dummies	0.0	0.0
Education dummies	0.1*	0.0
Regional dummies	-0.1*	-0.1*
Total change due to income (movement along EECs)	**3.9**	
Total change due to expenditure (movement along EECs)		**6.9**
Total change due to other demographics	**0.4**	**0.2**
Unexplained difference (shift in EECs)	**7.0**	**4.4**

Notes: Estimates based on Oaxaca-Blinder decomposition. Movement along EECs in column 1 is calculated as coefficient estimates from regression model (Table 1, column 2) multiplied by difference by corresponding changes in variable levels. Column 2 is constructed in parallel fashion but replacing after-tax income with aggregate consumption expenditure in the regression and decomposition. CO_2 content is estimates based on method described in Section 3, using CEX and WIOD data. Weights as provided by CEX survey. * regression coefficient significant at $p<0.05$.

The consumption-based CO_2 budget of the average household at constant 2009 technology increased by 11.3t between 1996 and 2009 (from 22.6t to 33.9t). Table 4 displays results of an Oaxaca-Blinder decomposition, which relies on coefficient estimates from the regression-based estimation of EECs. Essentially, the changes in levels of the outcome variable (here household CO_2) are divided into (i) changes in levels of explanatory variables when assuming constant regression coefficients, (ii) changes in regression coefficients holding variable levels constant, and (iii) an interaction thereof. For more details about Oaxaca-Blinder decomposition, the reader is referred to Appendix A.2 and the summary in Fortin et al. (2011).

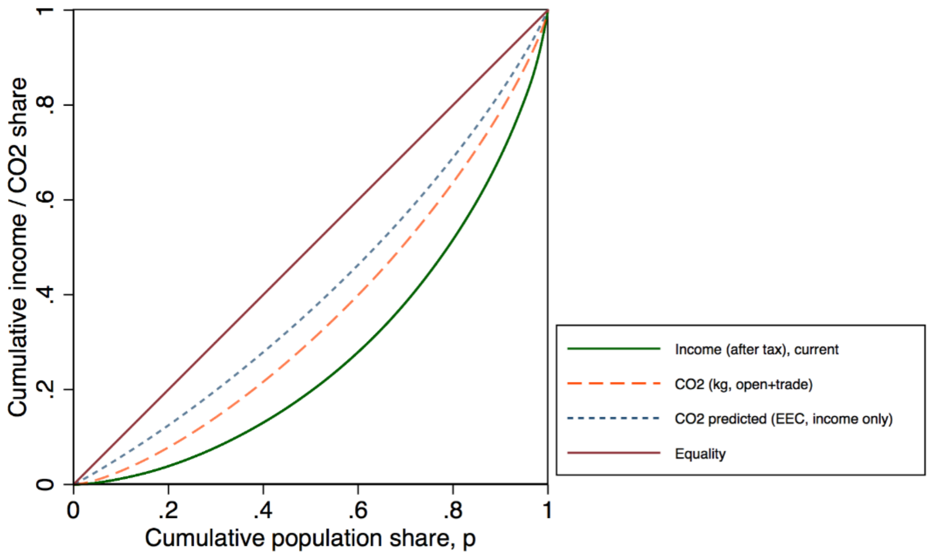
Table 4 Column 1 shows that changes in (i) income after tax, essentially movement along EECs, can account for about 3.9t (4.9 - 1.0) of the 11.3t overall change in household carbon between 1996 and 2009 (at constant 2009 technology). Changes in demographic characteristics contribute very little (0.4t of combined effects) to aggregate change. Meanwhile, effects (ii) and (iii), essentially shifts in the EECs, account for 7.0t of the difference. Column 2 makes clear that a significant portion of the unexplained shift in EECs is due to changes in expenditure levels at a given income. When replacing after-tax income with aggregate consumption expenditures (in the linear regression model and the decomposition), movement along the EECs accounts for 6.9t of the overall 11.3t change in household carbon.

In sum, changes in aggregate expenditure levels, which represent the combination of income growth and higher expenditure at given income, account for roughly 55% (6.9t out of 11.3t) of the total increase of average household CO_2 holding technology constant at 2009 levels. Meanwhile, shifts of EECs, which represent a change in the composition of consumption at a given expenditure level, account for about 35% (3.9t out of 11.3t) of the change.

As we have shown in Figure 2, improvements in technology have outweighed these dynamics, and average household carbon at current technology has decreased from 37.8t in 1996 to 33.9t in 2009.

Decomposing carbon inequality:

Figure 6: Lorenz curves – Income and household carbon (2009)



Notes: Cumulative population share and cumulative values of after-tax income (current USD), estimated household carbon contained in consumption (kg) and predicted values based on linear regression model with income and its square as independent variables. Household weights as provided by CEX sample. Households with reported after-tax income below USD 10k excluded.

Estimates of the carbon content of household consumption allow us to characterise the distribution of CO_2 in the population. A useful visual representation of distributions is given by the (generalised) Lorenz curve, plotting cumulative population shares against cumulative values of the variable of interest. In Figure 6, we present such Lorenz curves for after-tax incomes in 2009 and the estimated CO_2 content of household consumption. A few interesting insights are immediately suggested by visual inspection of Figure 6. Firstly, incomes were more unevenly distributed than consumption-based CO_2 in 2009 (Gini of 0.44 and 0.29 respectively). Secondly, it suggests that income inequality is an important driver of CO_2 inequality. This can be seen when comparing the CO_2 levels predicted (blue line) based on a linear regression of CO_2 on income and its square (Table 3, Column 3) with the estimated CO_2 levels based on expenditures (orange line). Figure 6 suggests that the distribution of income alone can reproduce a large portion of the inequality in household carbon (Gini of 0.22 and 0.29 respectively).

However, it is important to note that such visual inspection is merely suggestive, ignoring individual heterogeneity and associations with other relevant variables. In particular, the ordering of households in the income and CO_2 distributions may not be identical.

A more systematic method of quantifying the contribution of different variables to the dispersion of household CO_2 is again based on the coefficient estimates from Table 3. We follow the regression-based approach suggested by Fields (2003) and building on factor decomposition initiated by Shorrocks (1982). A brief description of this method can be found in Appendix A.3.

Results of the inequality decomposition are presented in Table 5. They confirm that income appears to be the key determinant in the distribution of household carbon as suggested by upward-sloping EECs. Depending on model specification, after-tax income accounts for about 31-40% of the dispersion of CO_2 in 2009. Interestingly, the weight of income in explaining household carbon dispersion appears to be decreasing over time (from 34-45% in 1996 to 31-40% in 2009). Family size is the second most important factor out of those included, accounting for about 13% and 12% in 1996 and 2009 respectively. Table 4 also suggests that there is a significant portion of the dispersion in CO_2, which is not accounted for by income or other variables. Residual dispersion is 45% and 49% in 1996 and 2009 respectively. This suggests that a significant role for household heterogeneity in preferences or for additional demographic characteristics not included here.

Table 5: Inequality decomposition – Household CO_2 (1996 / 2009)

	(1) 1996 (income)	(2) 1996 (full)	(3) 2009 (income)	(4) 2009 (full)
Income after tax	0.642	0.427	0.606	0.407
Income (squared)	-0.192	-0.0861	-0.204	-0.0984
Famiy size		0.215		0.207
Family size (squared)		-0.0889		-0.0773
Age		-0.0902		-0.0597
Age (squared)		0.112		0.0686
Married		0.0327		0.0407
Race (sum)		0.012		0.004
Education (sum)		0.018		0.012
Region (sum)		0.001		0.002
Residual	0.550	0.448	0.598	0.494
Observations	3,069	3,069	4,407	4,378
Total contribution of income	**45%**	**34%**	**40%**	**31%**
Total contribution of other demographics	**NA**	**21%**	**NA**	**20%**
Unexplained (residual)	**55%**	**45%**	**60%**	**49%**

Notes: Inequality decomposition based on coefficient estimates from linear regression models (Table 2). Calculations made using Stata module INEQRBD by Fiorio and Jenkins (2007). Household weights as provided in CEX sample. Households with reported after-tax income below USD 10k excluded.

6 The "equity-pollution dilemma"

We have demonstrated that Environmental Engel curves (EECs) are a useful tool in the analysis of household carbon, its' drivers and its' distribution over households. EECs for greenhouse gases from household consumption are clearly upward-sloping and concave. Assuming conditional heterogeneity of preferences, this concavity implies what we call the *"equity-pollution dilemma" – progressive redistribution of income may increase the emissions content of aggregate consumption.* While this dilemma has been acknowledged (Scruggs, 1998; Heerink et al., 2001), it has yet to be quantified using microdata. We propose a method to do so below.

Quantifying the "equity-pollution dilemma" with parametric (quadratic) Engel curves:

We have demonstrated above that a linear specification of EECs that includes a quadratic term (second-degree polynomial) approximates well the relationship between (after tax) income and household carbon while allowing for additive covariates. This quadratic specification yields a simple formula for the "equity-pollution dilemma". We continue to assume that households have homogenous preferences, i.e. that households move in parallel to the EECs when their incomes change (at least conditional on other linear associations included in the model). The marginal change in consumption-based CO_2 of household i when her income changes is then:

$$\frac{\partial y_i}{\partial m_i} = \beta_1 + 2\beta_2 m_i \tag{2}$$

A marginal transfer from household j to household i has the following effect on total CO_2:

$$\frac{\partial y_i}{\partial m_i} - \frac{\partial y_j}{\partial m_j} = -2\beta_2(m_j - m_i) \tag{3}$$

This leaves us with a useful result to quantify the "equity-pollution dilemma":

The expected change in aggregate CO_2, when choosing at random two households from the population, and re-distributing a small amount of income from the richer to the poorer, can be expressed as a function of the coefficient estimate $\hat{\beta}_2$ and Gini's mean difference[6] *Ψ (GMD), giving*

$$E_{ij}\left(\frac{\partial y_i}{\partial m_i} - \frac{\partial y_j}{\partial m_j}\middle| m_j > m_i\right) = -2\hat{\beta}_2 E_{ij}\left(m_j - m_i \middle| m_j > m_i\right) = -2\hat{\beta}_2 \Psi\big(F(m)\big)$$
$$\text{where } \Psi\big(F(m)\big) = \int\int |y - z|\, dF(y)\, dF(z) \tag{4}$$

[6] The GMD is equivalent to the "average self-distance" proposed by Koszegi and Rabin (2007) in their analysis of reference-dependent risk preferences.

In this simple quadratic approximation of EECs, and under the assumption of homogenous preferences (conditional on included covariates, household carbon moves in parallel to estimated EEC), the expected effect of a small progressive redistribution of income is thus negatively proportional to $\hat{\beta}_2$ as well as the dispersion measure Ψ.[7] The more dispersed the distribution of incomes and the more negative is $\hat{\beta}_2$, the larger the "equity-pollution dilemma".

For example, in our sample of US households in the year 2009, $\Psi = 55.3$ (in k USD) and $\hat{\beta}_2 = -0.26$ give an estimated increase of about 28.5 kg of household CO_2 for a *marginal redistribution* of 1000 USD from a higher income to a lower income household (both drawn at random). That constitutes about 5% of the carbon related to 1000 USD of income on average (514 kg).

Table 6: The "equity-pollution dilemma" – Comparison of pollutants (2009)

	(1) CO_2	(2) CO_2e	(3) CH_4	(4) N_2O
Income (k USD, after tax)	223.187***	304.581***	1.996***	0.045***
	(13.3885)	(18.3258)	(0.1285)	(0.0040)
Income squared (k USD, after tax)	-0.258***	-0.336***	-0.002***	-0.000***
	(0.0571)	(0.0785)	(0.0006)	(0.0000)
Observations	4,378	4,378	4,378	4,378
R-squared	0.506	0.525	0.506	0.476
HH characteristics	YES	YES	YES	YES
Implied "equity-pollution dilemma"				
Avg. emissions per income (kg per k USD)	563.3	789.9	5.186	0.169
$-2\hat{\beta}_2\Psi$	28.55	37.23	0.214	0.0047
Marginal effect of redistribution	**+5.1%**	**+4.8%**	**+4.2%**	**+2.8%**
Effect of full redistribution	+2.3%	+2.1%	+1.8%	+1.3%

Notes: Estimates from linear regression. Household weights as provided in CEX sample. Households with negative reported after-tax income and income above USD 400k excluded. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

[7] The discrete version of GMD can be defined as $\Psi = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{j=1}^{N} |m_i - m_j|$ for $i \neq j$.

Table 6 lists regression coefficient estimates and the implied magnitudes of the "equity-pollution dilemma" when comparing the embedded emissions of different greenhouse gases in 2009. Column 1 reproduces the estimates of Table 3 as well as the calculation described above. Columns 2-4 list estimates for totals greenhouse gases (CO_2e), methane (CH_4), and nitrous oxide (N_2O) respectively. For each of these pollutants, we do estimate concave EECs and thus a positive "equity-pollution dilemma". However, this dilemma seems to be the largest for CO_2, with estimates of the rise in pollution from a marginal redistribution at 4.2% and 2.8% for CH_4 and N_2O respectively.

Full redistribution:

Regression-based EECs also allow for the calculation of the change in predicted household carbon if all households had the same income equal to the mean:

The difference between the expected mean of household carbon under "full equality" and the current mean level at a given income distribution is given by:

$$\hat{\beta}_2 \left[\bar{m}^2 - \frac{1}{N} \sum_{i=1}^{N} (m_i)^2 \right]$$

In the case of our sample, average household carbon in 2009 is predicted to increase by 0.8t from 33.9t estimated currently to about 34.7t under full income equality, a rise of 2.3%. The respective increases in emissions when moving to full equality are 1.8% for CH_4 and 1.3% for N_2O.

The above quantification makes clear that estimates of the magnitude of the "equity-pollution dilemma" are sensitive to estimates of $\hat{\beta}_2$. For example, without including socio-demographic covariates (Table 3 Column 3), we estimated a much larger absolute $\hat{\beta}_2$ (0.54 instead of 0.26) and hence would have significantly overestimated the dilemma.

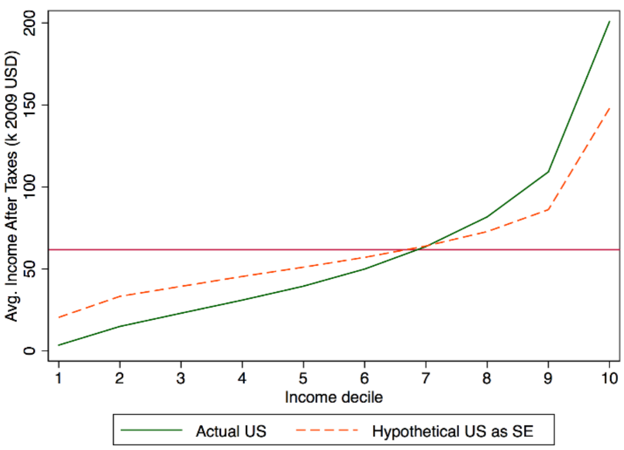
Hypothetical income distribution – Sweden:

Finally, we estimate the predicted change in average household carbon when moving from the 2009 distribution of household incomes in the United States to the income distribution of Sweden in the same year. To do so, we obtain decile average household incomes in 2009 (disposable income including capital income, equalised) as provided by Statistics Sweden (SCB, 2017). We then scale decile average incomes in the United States so that they match the decile shares in total income of the Swedish distribution. We rescale incomes to keep constant the aggregate mean income in the United States to avoid scale effects. Figure 7a compares these hypothetical average decile incomes (red) with the actual average household incomes by decile as observed in our sample for 2009 (green).

We then estimate the change in predicted household carbon when moving from the actual average income per decile to the hypothetical value emulating the Swedish income distribution. The effect of this change is predicted based on the coefficient estimates from our preferred specification (Table 3 Column 4). We predict that average household carbon would have been 0.5t higher under the Swedish income distribution, corresponding to an increase of about 1.5% relative to average household carbon of 33.9t in 2009. Figure 7b illustrates how that predicted increase in average household carbon is distributed over income deciles.

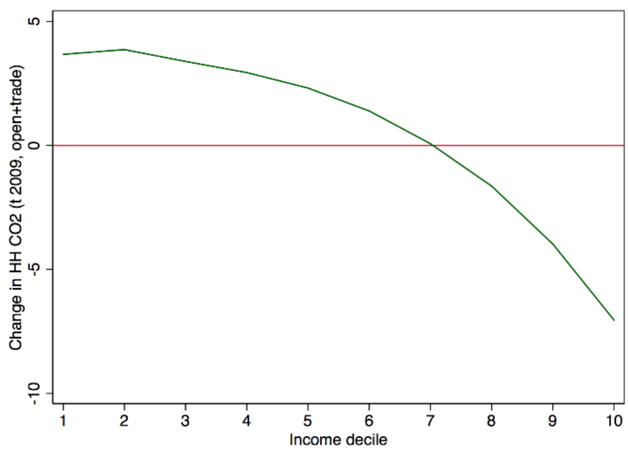
Figure 7: Hypothetical income distribution – Sweden – 2009

Figure 7a: Comparison – Household incomes



Note: Green = Average household income after as observed in analysis sample; Green = Average household income after scaling of US distribution to mirror decile shares of Swedish distribution of disposable household income. Both by income deciles, 2009 data.

Figure 7b: Predicted change in HH carbon



Note: Predicted difference between average household CO_2 by income decile between hypothetical distribution emulating Sweden and actual distribution in the United States. Calculations based on estimates reported in Table 3, Column 4. 2009 data.

Assumptions and limitations:

The methodology proposed above to quantify the "equity-pollution dilemma" is based on three critical assumptions. Firstly, we assume throughout that we have arrived at unbiased estimates of the carbon content of household consumption baskets along the income distribution. Limitations to input-output based carbon accounting have been discussed above. One important remaining concern is the assumption of constant emissions intensity per dollar expenditure along the income distribution. As discussed above, price/quality heterogeneity of products thus likely results in estimates of EECs that are more convex than true EECs, resulting in *underestimation* of $\hat{\beta}_2$ and consequently the "equity-pollution dilemma".

Secondly, we assume throughout that the linear model specified in equation (1) is adequate. We have shown above that a second-degree polynomial specification approximates well the relationship between income and household carbon as shown by more flexible nonparametric models. Relatedly, we assume homogeneity of household preferences conditional on income and the set of household characteristics included in (1) as covariates.

We thus assume that households will respond to a change in their income by moving in parallel to the estimated EECs[8]. This implies that there is no variable omitted from our specification of EECs that influences both incomes and consumption preferences at the same time. While this assumption is necessary for our analysis, there is some evidence to the contrary. For example, Lewbel and Pendakur (2017) find evidence of significant preference heterogeneity in the demand for energy. Such unobserved heterogeneity in preferences would pose a problem for our quantification of the "equity-pollution dilemma" if it means that the observed relationship between household income and the income elasticity of demand were driven by some unobserved factor. This might lead to households responding to income changes by not moving in parallel to the EECs, which is our fundamental assumption in quantifying the dilemma. Arguably, income and consumption preferences are shaped by a range of experiences, choices, and external factors over a household's life cycle. Alan et al. (forthcoming) find evidence of such co-dependence between income and preferences. This opens the possibility of bias in our hypothetical analysis underlying the "equity-pollution dilemma". However, we are not aware of convincing evidence that would predict the sign of such a bias nor of possible ways to overcome this limitation.

[8] Consider a hypothetical change in income for a household with actual income x to hypothetical income y. We thus assume throughout that this household would consume the same bundle of goods as a household with actual incomey (holding constant all other household characteristics to be included in the analysis).

We further assume that consumer preferences are not only homogenous (conditional on observed household characteristics), but also independent of the distribution of income. However, a growing literature finds evidence of relative preferences, such as conspicuous consumption based on a desire for status (Veblen, 1899; Bagwell and Bernheim, 1996; Charles et al., 2009). Allowing for preferences to be endogenous in such a fashion would mean that the shape of EECs themselves would change in response to changes in the distribution of income, negating our counterfactual analysis.

Finally, we assume throughout that external circumstance of consumption remain fixed when income is redistributed. In particular, our analysis is a partial equilibrium one and we assume that redistribution does not affect the emissions intensity of goods, implying no effect of income redistribution on production technologies and retail prices. However, it is conceivable that demand shifts towards less emissions intensive goods might induce changes in relative prices or stimulate innovation in production. Similarly, production technologies and market conditions may change if income redistribution would indeed influence the political landscape by shifting political influence between different demographics – the political economy channel proposed by Boyce (1994), which was not the focus on this paper.

The assumptions listed above are generally less restrictive when considering marginal or small-scale redistribution of income. Meanwhile, large-scale income redistribution might have wider-ranging implications which themselves feed back into production technologies and prices.

Welfare economic implications:

We believe that the above finding of a potential trade-off between income redistribution and carbon emissions – what we term the "equity-pollution dilemma" – is an important dynamic to consider when designing redistributive policies. However, the "equity-pollution dilemma" does not necessarily render income redistribution undesirable. The optimal degree of redistributive policy requires extensive welfare economic analysis and will rely on a variety of assumptions regarding market structure, household welfare and socially desirable outcomes. For example, the estimated increase of about 28.5 kg of household CO_2 for a *marginal redistribution* of 1000 USD in 2009 might represent a social externality cost of roughly 90 cents (applying a conservative estimate for the social cost of carbon of 31 USD following Nordhaus, 2017). An inequality-averse social planner might well believe that the benefits of redistributing 1000 USD may compensate for a social cost of 90 cents.

7 Conclusion

This paper contributes to the understanding of the interplay between the distribution of household income, expenditure, and the carbon content of consumption. Based on detailed expenditure data from the US Consumer Expenditure Survey (CEX) for the period 1996-2009, estimates of household carbon are derived based on input-output data from WIOD as well as energy emissions factors. Estimates of household carbon are used to derive Environmental Engel curves (EECs) for CO_2. This paper estimates parametric EECs for greenhouse gases, following Levinson and O'Brien (2015) who do so for air pollutants. EECs are found to be upward-sloping, concave, and shifting downwards over time. We find that a second-degree polynomial specification for EECs fits well the observed relationship between income and household carbon, after controlling for household characteristics. The paper proceeds with a range of simple descriptive/predictive analyses, which highlight the usefulness of such parametric estimates of EECs.

The paper finds that average household carbon has declined from 37.8t in 1996 to 33.9t in 2009. However, it would have risen significantly had technology remained constant. Based on coefficient estimates from regression-based EECs, an Oaxaca-Blinder decomposition suggests that changes in incomes can account for about 35% of this increase in household carbon at constant technology. Factoring in changes in savings behaviour, changes in expenditure levels even account for about 55% of the increase. We further find that there is significant inequality in household carbon, though it is lower than inequality of income and expenditure. Using regression-based inequality decomposition, we find that income is the strongest driver of carbon inequality out of the variables considered. Household income is found to account for about 31-40% of carbon inequality in 2009.

A key contribution of this paper is the quantification of the "equity-pollution dilemma": *Given the higher pollution intensity of consumption per expenditure by poorer households, progressive redistribution may result in higher aggregate pollution from consumption.* Assuming that households have (conditionally) homogenous preferences, we find that a marginal transfer of 1000 USD from a richer to a poorer household in 2009 may increase the CO_2 content of that income by about 28.5kg or 5%. Similarly, we predict that aggregate household carbon would have been about 1.5% higher under a hypothetical scenario of income distributed as in Sweden and 2.3% higher under full equality. We hope that the formal analysis relying on parametric estimates of EECs, and in particular the proposed quantification of the "equity-pollution dilemma" will inspire further systematic work on the relationship between household income and consumption-based pollution.

References:

Alan, S., M. Browning, and M. Ejrnaes (forthcoming). Income and consumption: A micro semi-structural analysis with pervasive heterogeneity. *Journal of Political Economy* forthcoming.

Atkinson, A., T. Piketty, and E. Saez (2011). Top incomes in the long run of history. *Journal of Economic Literature* 49(1), 3–71.

Attanasio, O. P. and G. Weber (1995). Is consumption growth consistent with intertemporal optimization? Evidence from the Consumer Expenditure Survey. *Journal of Political Economy* 103(6), 1121-1157.

Baek, J. and G. Gweisah (2013). Does income inequality harm the environment? Empirical evidence from the United States. *Energy Policy* 62, 1434–1437.

Bagwell, S. and B. D. Bernheim (1996). Veblen effects in a theory of conspicuous consumption. *American Economic Review* 86(3), 349–373.

Berthe, A. and L. Elie (2015). Mechanisms explaining the impact of economic inequality on environmental deterioration. *Ecological Economics* 116, 191–200.

Blinder, A. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8(4), 436-455.

Boyce, J. (1994). Inequality as a cause of environmental degradation. *Ecological Economics* 11, 169–178.

Buechs, M. and S. Schnepf (2013). Who emits most? Associations between socio-economic factors and UK households' home energy, transport, indirect and total CO2 emissions. *Ecological Economics* 90, 114–123.

Chakravarty, S., Chikkatur, A., de Coninck, H., Pacala, S., Socolow, R., and M. Tavoni (2009). Sharing global CO2 emission reductions among one billion high emitters. *Proceedings of the National Academy of Sciences* 106(29), 11884-11888.

Chancel, L., and T. Piketty (2015). *Carbon and inequality: From Kyoto to Paris.* Paris, France: Paris School of Economics.

Charles, K. K., E. Hurst, and N. Roussanov (2009). Conspicuous consumption and race. *Quarterly Journal of Economics* 124(1), 425–467.

Chitnis, M., S. Sorrell, A. Druckman, S. K. Firth, and T. Jackson (2014). Who rebounds most? Estimating direct and indirect rebound effects for different UK socioeconomic groups. *Ecological Economics* 106, 12-32.

Currie, J. and M. Neidell (2005). Air pollution and infant health: What can we learn from California's recent experience? *Quarterly Journal of Economics* 120(3), 1003–1030.

Davis, S.J., and K. Caldeira (2010). Consumption-based accounting of CO2 emissions. *Proceedings of the National Academy of Sciences* 107(12), 5687-5692.

Druckman, A., P. Bradley, E. Papathanasopooulou, and T. Jackson (2008). Measuring progress towards carbon reduction in the UK. *Ecological Economics* 66, 594-604.

Druckman, A., and T. Jackson (2010). The bare necessities: How much household carbon do we really need? *Ecological Economics* 69, 1794-1804.

Druckman, A., and T. Jackson (2015). Understanding households as drivers of carbon emissions. In: Clifts, R. and A. Druckman (eds.), *Taking stock of industrial ecology*. Cham, Switzerland: Springer.

EPA (2009). Mandatory Reporting of Greenhouse Gases. *Federal Register* 74(209), 56373-56519.

EPA (2017). *eGRID2014v2 – The emissions & generation resource integrated database*. Retrieved from https://www.epa.gov/energy/emissions-generation-resource-integrated-database-egrid on 30 March 2017.

Fields, G. (2003). Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States. *Research in labor economics* 22(3), 1-38.

Fiorio, C., and S. Jenkins (2007). *Regression-based inequality decomposition, following Fields (2003)*. Presentation at the UK Stata User Group meeting, 10 September 2007.

Fortin, N., Lemieux, T., and S. Firpo (2011). Decomposition methods in economics. *Handbook of labor economics* 4, 1-102.

Fields, G. (2003). Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States. *Research in labor economics* 22(3), 1-38.

Fouquet, R. (2014). Long-run demand for energy services: Income and price elasticities over two hundred years. *Review of Environmental Economics and Policy* 8(2), 186–207.

Grainger, C. A. and C. D. Kolstad (2010). Who pays a price on carbon? *Environmental and Resource Economics* 46(3), 359–376.

Hardle, W. and E. Mammen (1993). Comparing nonparametric versus parametric regression fits. *Annals of Statistics* 21(4), 1926–1947.

Harris, E., and J. Sabelhaus (2000). Consumer Expenditure Survey, family-level extracts, 1980:1-1998:2. *NBER Working Paper*, 2000. Retrieved from http://www.nber.org/data/ces_cbo.html on 12/12/2016.

Heerink, N., A. Mulatur, and E. Bulte (2001). Income inequality and the environment: Aggregation bias in environmental Kuznets curves. *Ecological Economics* 38, 359–367.

Heffetz, O. (2011). A test of conspicuous consumption: Visibility and income elasticities. *Review of Economics and Statistics* 93(4), 1101–1117.

Holland, S. P., E. T. Mansur, N. Z. Muller, and A. J. Yates (2016). *Distributional effects of air pollution from electric vehicle adoption.* Working Paper 22862, National Bureau of Economic Research.

Hsiang, S., R. Kopp, A. Jina, J. Rising, M. Delgado, S. Mohan, D. J. Rasmussen, R. Muir-Wood, P. Wilson, M. Oppenheimer, K. Larsen, and T. Houser (2017). Estimating economic damage from climate change in the United States. *Science* 356(6345), 1362–1369s.

Kahn, M. (1998). A household level environmental Kuznets curve. *Economics Letters* 59(2), 269–273.

Koszegi, B., and M. Rabin (2007). Reference-dependent risk attitudes. *American Economic Review* 97(4), 1047-1073.

Lenzen, M., Wier, M., Cohen, C., Hayami, H., Pachauri, S., and R.Schaeffer (2006). A comparative multivariate analysis of household energy requirements in Australia, Brazil, Denmark, India and Japan. *Energy* 31, 181-207.

Lenzen, M. (2008). Consumer and producer environmental responsibility: A reply. *Ecological Economics* 66, 547-550.

Leontief, W. (1970). Environmental repercussions and the economic structure: An input-output approach. *Review of Economics and Statistics* 52(3), 262-271.

Levinson, A. and J. O'Brien (2015). *Environmental Engel curves.* Working Paper 20914, National Bureau of Economic Research.

Lewbel, A. and K. Pendakur (2017). Unobserved preference heterogeneity in demand using generalized random coefficients. *Journal of Political Economy* 125(4), 1100-1148.

Liu, W., G. Spaargaren, N. Heerink, A. Mol, and C. Wang (2013). Energy consumption practices of rural households in north China: Basic characteristics and potential for low carbon development. *Energy Policy* 55, 128–138.

Mendelsohn, R., A. Dinar, and L. Williams (2006). The distributional impact of climate change on rich and poor countries. *Environment and Development Economics* 11, 159–178.

Myhre, G., D. Shindell, F.-M. Breon, W. Collins, J. Fuglestvedt, J. Huang, D. Koch, J.-F. Lamarque, D. Lee, B. Mendoza, T. Nakajima, A. Robock, G. Stephens, T. Takemura, and H. Zhang (2013). Anthropogenic and Natural Radiative Forcing. In Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge, United Kingdom and New York, NY: Cambridge University Press.

Nordhaus, W. D. (2017). Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences of the United States* 114(7), 1518-1523.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review* 14(3), 693-709.

Papathanasopoulou, E., and T. Jackson (2009). Measuring fossil resource inequality – A case study for the UK between 1968 and 2000. *Ecological Economics* 68, 1213-1225.

Pearce, D. (1991). The role of carbon taxes in adjusting to global warming. *The Economic Journal* 101(407), 938–948.

Robinson, P. M. (1988). Root-N consistent semiparametric regression. *Econometrica* 56(4), 931–954.

Sato, M. (2014). Embodied carbon in trade: A survey of the empirical literature. *Journal of Economic Surveys* 28(5), 831-861.

SCB (2017). Household finances statistics. Statistics Sweden. Retrieved from http://www.scb.se/he0103-en on 29 September 2017.

Scruggs, L. (1998). Political and economic inequality and the environment. *Ecological Economics* 26, 259–275.

Shorrocks, A. (1982). Inequality decomposition by factor components. *Econometrica* 50(1), 193-211.

Thomas, B. A. and I. L. Azevedo (2013). Estimating direct and indirect rebound effects for U.S. households with input-output analysis Part 1: Theoretical framework. *Ecological Economics* 86, 199-210.

Torras, M. and J. Boyce (1998). Income, inequality, and pollution: A reassessment of the environmental Kuznets curve. *Ecological Economics* 25, 147–160.

Tukker, A., and B. Jansen (2006). Environment impacts of products – a detailed review of studies. *Journal of Industrial Ecology* 10, 159-182.

U.S. Energy Information Administration (2017). Statistics on prices of electricity, natural gas, petroleum & other liquids. Retrieved from http://www.eia.gov on 30 March 2017.

Veblen, T. (1899). *The theory of the leisure class [reprint 1965]*. New York, NY: MacMillan.

Weber, C., and H. Matthews (2008). Quantifying the global and distributional aspects of American household carbon footprint. *Ecological Economics* 66, 379-391.

Wiedmann, T. (2009). A review of recent multi-region input-output models used for consumption-based emission and resource accounting. *Ecological Economics* 69, 211-222.

Appendix A.1: Estimation of emission content of consumption

We aim to construct Environmental Engel curves (EECs) for household carbon in the United States. We focus on households in the United States, because it has some of the highest consumption-based CO_2 per household (e.g. Chancel and Piketty, 2015). At the same time, detailed data are available on the income and consumption patterns of households. We estimate the CO_2 attributable to the consumption of all energy, fuels, goods and services by households at different positions in the income distribution. The focus of this exercise is thus on *total emissions* contained in household consumption. This includes *direct* emissions from the consumption of fossil fuel based energy (e.g. heating, electricity, transportation fuels) as well as *indirect* or "embedded" emissions from the production of goods and services consumed. We base our accounting methodology on Environmentally-Extended Input-Output Analysis as is standard in the literature on consumption-based emission accounting (Wiedmann, 2009).

We then construct EECs by following the methodology by Levinson and O'Brien (2015) in combining information on yearly expenditures of households on different consumption items (in dollars) with estimates of the carbon intensity of these different goods and services (kg of CO_2 per dollar).

Consumption data:

Information on household income, consumption expenditures, and socio-demographic characteristics, comes from the United States Consumer Expenditure Survey (CEX). The Bureau of Labor Statistics provides anonymised public use micro-data from 1996. We make use of the interview portion of the CEX, containing information on survey responses by "consumer units" (CU). In what follows, we will refer to "consumer units" as households. Our main source of information are the "monthly expenditures" files (MTBI), which contain information on a household's expenditures (and incomes) split into over 800 categories assigned universal classification codes (UCC). We combine these with income and socio-demographic characteristics contained in the "consumer unit characteristics and income" files (FMLI).

The CEX consumption expenditures are then each allocated to one WIOD production sector. It is in this step, where a number of judgements by the researcher are necessary. We follow where possible the matching procedure used by Levinson and O'Brien (2015) to link UCC to

IO codes used in the input-output tables of the Bureau of Economic Analysis[9]. We then match the IO codes to the smaller number of WIOD production sectors (34 sectors, excluding the "Private Households" sector). Due to significant overlap in definitions and coding conventions, the matching of BEA IO to WIOD codes is mostly unambiguous. Nevertheless, there are certain categories, where we used assumptions to arrive at a full and exclusive matching of expenditure categories to production sectors[10].

Multiplying the consumption expenditures of a household with the matched *total emissions intensity* yields a rough estimate of the CO_2 embedded in the yearly consumption of that household, which we shall call *estimated household carbon / CO_2*.

Emission content of consumption:

Total emissions z can be represented as two identities, depending on either total output **x** or final demand **y**:

$$z = \boldsymbol{x}'\boldsymbol{d} = \boldsymbol{y}'\boldsymbol{e}$$

As we have obtained estimates of final demand per household *k* (i.e. the vector $\mathbf{y}_k$) from the CEX data, we aim to multiply household final demand with total emission intensities **e** to arrive at estimates of the total emissions content of the consumption by household *k*:

$$z_k = \boldsymbol{y}_k{}'\boldsymbol{e}$$

We thus require estimates of the emissions intensity **e** per unit of final demand **y** per sector.

Input-output based emission factors:

In order to allocate emissions intensities to consumption categories, information from the World Input-Output Tables (WIOD) is used. The 2013 release of WIOD contains information on 35 production sectors in 40 countries for the years 1995 through 2009. Notably, WIOD publishes "Environmental Accounts", which include information on total yearly emissions per sector (represented by the vector **z**) and gross output per sector (represented by the vector **x**). In this paper, we make use of the information on 34 of the 35

[9] We are grateful to Arik Levinson and James O'Brien for kindly sharing their matching from UCC categories to IO codes used in their forthcoming paper and for answering our questions regarding their methodology. As there are many more UCC categories than IO sectors, the matching procedure applied by Levinson and O'Brien (2015) relies on a number of subjective judgements, which they outline in an online appendix to their forthcoming paper.
[10] Matching of CEX UCC codes to WIOD sectors to be provided as online appendix for eventual publication.

WIOD sectors (excluding production in "Private Households"). A list of the 34 WIOD sectors used and their estimated emissions intensities for the years 1996 and 2009 is provided in Table A.1. We use these to allocate to each sector a *direct emissions intensity* (kg of CO_2, CH_4, N_2O per \$ of total output)*:*

$$\boldsymbol{d} = \boldsymbol{z} \oslash \boldsymbol{x}$$

Here, $\oslash$ represents element-wise division. We make use of the input-output portion of WIOD to attribute to each sector a *total emissions intensity* (vector **e**). This total emissions intensity **e** is intended to capture the emission content of each unit of final demand **y** per industry. To arrive at a useful estimate of **e**, we need to incorporate the role of intermediate goods – output that is not used for final demand, but nevertheless requires economic activity and thus emissions. We exploit the global nature of the input-output tables to construct three types of emission factors based on different assumptions regarding trade: (a) Closed economy, (b) Global supply-chain, but no trade; (c) Global supply-chain and trade.

Closed economy:

We follow Leontief (1970), who proposed a linear relationship between the vector of total output in n sectors, **x**, and the final demand from those n sectors, **y**, of the form:

$$\mathbf{x} = \mathbf{Cx} + \mathbf{y}$$

Here, the $n \times n$ (n=34 under the closed economy assumption) matrix $\mathbf{C}$ is called the *Direct Requirement matrix* and has element c_{ij}, which stands for the dollar amount of input from industry i necessary for the production of a dollar output from production j. In order to take account of secondary and higher-order relationships between input and output sectors, the *Direct Requirement matrix* $\mathbf{C}$ can be converted into the *Total Requirement matrix* $\mathbf{T}$. This matrix gives the dollar amount of output necessary from each sector j for a dollar of consumption in each sector i, taking into account all intermediate steps in the supply chain *ad infinitum*:

$$\mathbf{x} = [\mathbf{I} - \mathbf{C}]^{-1}\mathbf{y} = \mathbf{Ty}$$

We then convert the *vector of emissions intensities* $\mathbf{d}$ into the *vector of total emissions intensities* $\mathbf{e}$:

$$\boldsymbol{e} = \mathbf{T}'\mathbf{d}$$

Global supply chain:

The above derivation of the emissions intensity for final demand by US consumers in 34 sectors, represented by the vector $\boldsymbol{e}$, is based on the assumption that the United States is a closed economy and that all final consumption as well as intermediate goods are produced by domestic sectors. We now introduce a global supply chain, which incorporates the fact that US sectors obtain intermediate goods from productive sectors around the world. We make use of data contained in WIOD on 40 countries (incl. the United States).

With $m = 41$ countries (including "Rest of the World") and $n = 34$ sectors, the *Direct Requirement matrix* $\mathbf{C}$ is now of dimension $(mn \times mn)$=(1394 × 1394). We again obtain the *Total Requirement matrix* $\mathbf{T} = [\mathbf{I} - \mathbf{C}]^{-1}$. The *vector of emissions intensities* $\mathbf{d}^{World}$ is now also of the dimension (1394 × 1) as is the *vector of total emissions intensities* $\boldsymbol{e}^{World} = \mathbf{T}'\mathbf{d}^{World}$.

In a final step, we then extract only the 34-element vector relating to the final demand of consumers in the United States, $\boldsymbol{e}^{US}$, which now incorporates the emissions of intermediate goods supplied by the 34 sectors in all 41 countries.

Trade in final goods:

In a final step, we incorporate the fact that some of the final demand by consumers in the United States will be met through final goods imported from other countries. To do so, we make use of information on "final consumption expenditure by private households" contained in the WIOD input-output tables. Starting from this, we construct a matrix $\mathbf{M}$, which has dimension $(m \times n)$=(41 × 34), where entry m_{ij} represents the share of final demand of US private households to sector j imported from country i (i.e. columns of $\mathbf{M}$ sum to 100%).

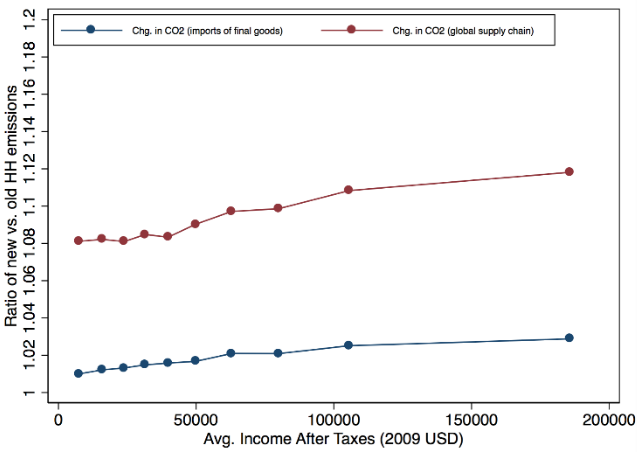
We then convert the *vector of total emissions intensities* $\boldsymbol{e}^{World}$ to a matrix $\boldsymbol{E}^{World}$ with dimensions $(n \times m)$=(34 × 41). The vector of emission intensities corresponding to final demand by US households, but incorporating the shares of final goods imported from other countries, is then given by:

$$\boldsymbol{e}^{Full} = \mathrm{diag}(\boldsymbol{E}^{World}\mathbf{M})$$

Figure A.1a represents adjustment factors when moving from the closed-economy assumption to a global supply chain and the inclusion of direct imports of final goods. Interestingly, the inclusion of trade has a larger relative impact on estimates of household carbon for those with higher incomes (e.g. an approximate 12% increase in CO2 for the top decile when considering global supply chains compared to an 8% increase for households at the bottom decile).

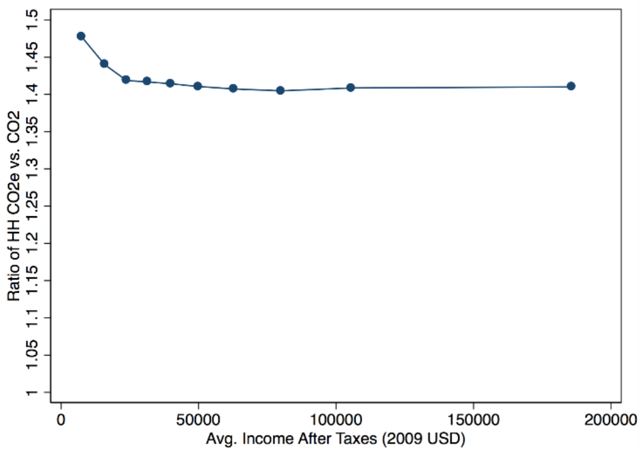
Figure A.1: Comparison of emission measures – 2009

Figure A.1a: Global supply chain & trade



Note: Red = Average ratio of household CO_2 emissions when including global supply chain vs. closed economy assumption; Blue = Average ratio of household CO_2 emissions when including direct imports of final goods vs. all final goods from US production. Both by income deciles, 2009 data.

Figure A.1b: CO2 vs. CO2e (incl. CH_4, N_2O)



Note: Average ratio of household total greenhouse gas emissions (CO_2e) vs. CO_2 emissions by income deciles. 2009 data.

<u>*Direct emission factors for high-carbon goods:*</u>

To improve the precision of our estimates, we allocate emissions intensities to certain high-carbon consumption categories directly. We do so for expenditures on home electricity, heating oil, natural gas, gasoline for car (incl. Diesel and motor oil), and air travel. Data on end consumer prices for electricity, heating oil, natural gas, and gasoline are provided by the U.S. Energy Information Administration (2017). Emissions factors for gasoline, heating oil, natural gas, and kerosene are those used by the U.S. Environmental Protection Agency in guidelines for the Greenhouse Gas Inventory (EPA, 2009). Emission intensity of residential electricity is taken from the EPA's Emissions & Generation Resource Integrated Database (EPA, 2017). An overview of the resulting emission factors used is given in Table A.2.

We believe that this methodology improves significantly the precision of our estimates of household carbon embedded in consumption. The implementation of direct emission factors for these consumption categories increases aggregate household carbon by about 25% (from 25.0t on average with only WIOD factors to 31.0t with added direct emission factors in 2009).

Emission factors for methane (CH_4) and nitrous oxide (N_2O):

While carbon dioxide (CO2) is the most common greenhouse gas, especially when considering energy production based on fossil fuels, there are further greenhouse gases which contribute to global warming. Among those, we account for methane (CH_4) and nitrous oxide (N_2O), both of which are reported in the WIOD Environmental Accounts. We thus repeat the procedure described above for both CH_4 and N_2O. In a final step we then construct an aggregate measure for greenhouse gas content in consumption, converted into carbon dioxide equivalent scale, by multiplying emissions with their 100 year global warming potential multipliers[11]. Figure A.2b depicts adjustment factors of that process.

Consumption categories:

We follow closely the methodology of Heffetz (2011), building on Harris and Sabelhaus (2000), who assign UCC categories from the CEX survey to 109 categories (47 for consumption, 22 for income and 40 for other). We then assign expenditures to 29 of the consumption categories used by Heffetz (2011) (excluding from his original 31 categories those of expenditures on cell phones, and underwear).

[11] We use the 100 year global warming potential multipliers with climate-carbon feedbacks as reported in the IPCC AR5 report (Myhre et al., 2013) – namely 34 for CH_4 and 298 for N_2O.

Table A.1: List of WIOD Sectors used

WIOD Code	WIOD Name	CO_2 (kg/$, 1996)	CO_2 (kg/$, 2009)	CH_4 (g/$, 2009)	N_2O (g/$, 2009)
15t16	Food, Beverages and Tobacco	0.71	0.49	11.55	0.73
17t18	Textiles and Textile Products	0.91	0.75	8.58	0.34
19	Leather, Leather and Footwear	0.77	0.56	10.42	0.50
20	Wood and Products of Wood and Cork	1.20	0.85	10.43	0.55
21t22	Pulp, Paper, Paper , Printing and Publishing	0.69	0.47	2.21	0.06
23	Coke, Refined Petroleum and Nuclear Fuel	2.27	0.94	23.26	0.03
24	Chemicals and Chemical Products	1.15	0.68	5.02	0.18
25	Rubber and Plastics	0.94	0.62	4.62	0.13
26	Other Non-Metallic Mineral	3.21	1.94	6.17	0.05
27t28	Basic Metals and Fabricated Metal	1.50	0.85	4.77	0.04
29	Machinery, Nec	0.71	0.57	3.68	0.04
30t33	Electrical and Optical Equipment	0.64	0.42	2.87	0.04
34t35	Transport Equipment	0.55	0.38	2.24	0.03
36t37	Manufacturing, Nec; Recycling	0.71	0.55	4.80	0.13
50	Sale, Maintenance and Repair of Motor Vehicles and Motorcycles; Retail Sale of Fuel	0.32	0.17	0.94	0.01
51	Wholesale Trade and Commission Trade, Except of Motor Vehicles and Motorcycles	0.21	0.09	0.49	0.01
52	Retail Trade, Except of Motor Vehicles and Motorcycles; Repair of Household Goods	0.34	0.17	0.62	0.01
60	Inland Transport	1.07	0.79	9.63	0.03
61	Water Transport	2.94	1.98	5.20	0.10
62	Air Transport	1.77	1.48	4.95	0.07
63	Other Supporting and Auxiliary Transport Activities; Activities of Travel Agencies	0.45	0.44	2.04	0.02
64	Post and Telecommunications	0.23	0.18	1.32	0.01
70	Real Estate Activities	0.21	0.06	0.38	0.00
71t74	Renting of M&Eq and Other Business Activities	0.26	0.14	0.95	0.01
AtB	Agriculture, Hunting, Forestry and Fishing	0.73	0.49	36.88	2.61
C	Mining and Quarrying	1.29	0.57	34.90	0.02
E	Electricity, Gas and Water Supply	7.93	5.42	10.54	0.09
F	Construction	0.57	0.38	4.06	0.04
H	Hotels and Restaurants	0.57	0.30	2.29	0.10
J	Financial Intermediation	0.17	0.09	0.58	0.01
L	Public Admin and Defence; Compulsory Social Security	0.52	0.25	1.71	0.02
M	Education	0.56	0.35	1.17	0.03
N	Health and Social Work	0.36	0.17	0.85	0.02
O	Other Community, Social and Personal Services	0.43	0.18	8.59	0.04

Notes: List of 34 out of 35 WIOD sectors (excluding “Private Household”). Estimates for kg CO2 content per USD output according to methodology described in Section 3 (1996 and 2009).

Table A.2: List of WIOD countries

Code	Country	Code	Country
AUS	Australia	JPN	Japan
AUT	Austria	KOR	Korea
BEL	Belgium	LVA	Latvia
BRA	Brazil	LTU	Lithuania
BGR	Bulgaria	LUX	Luxembourg
CAN	Canada	MLT	Malta
CHN	China	MEX	Mexico
CYP	Cyprus	NLD	Netherlands
CZE	Czech Republic	POL	Poland
DNK	Denmark	PRT	Portugal
EST	Estonia	ROM	Romania
FIN	Finland	RUS	Russia
FRA	France	SVK	Slovak Republic
DEU	Germany	SVN	Slovenia
GRC	Greece	ESP	Spain
HUN	Hungary	SWE	Sweden
IND	India	TWN	Taiwan
IDN	Indonesia	TUR	Turkey
IRL	Ireland	GBR	United Kingdom
ITA	Italy	USA	United States
RoW	Rest of World		

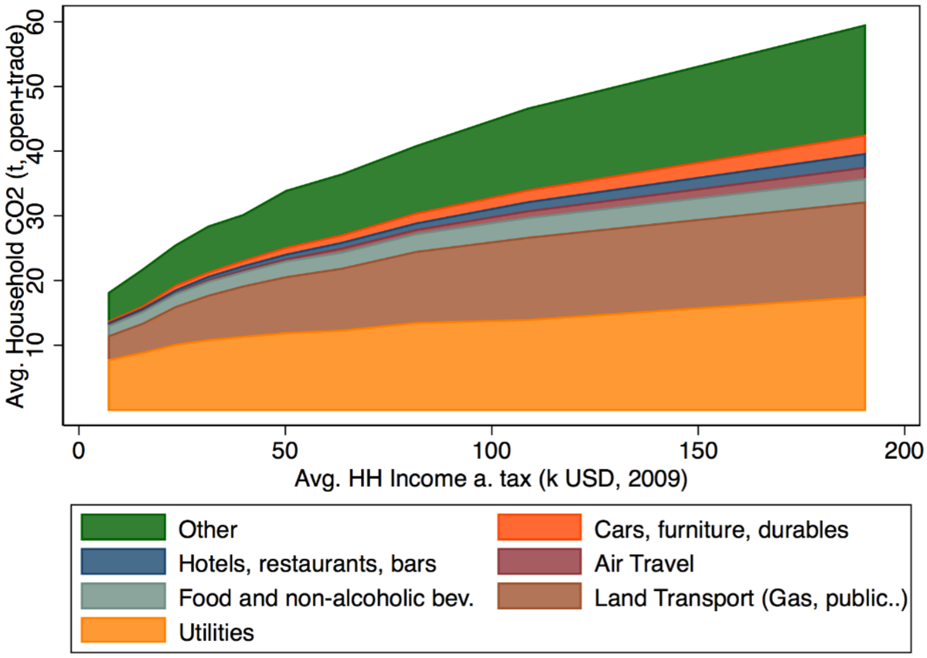
Notes: List of 41 WIOD countries (including "Rest of World").

Table A.3: Direct emission factors (kg CO_2 per USD)

Year	Electricity	Gasoline	Heating fuel	Natural gas	Air travel
1996	8.67	7.14	9.26	7.82	2.14
1997	8.72	7.14	9.46	7.31	2.11
1998	8.61	8.29	11.09	7.32	1.99
1999	8.58	7.56	10.69	7.42	2.07
2000	8.45	5.84	6.85	6.40	1.81
2001	8.07	6.09	7.66	5.50	1.99
2002	8.16	6.41	8.26	6.35	2.07
2003	7.86	5.54	6.73	5.13	1.89
2004	7.61	4.69	5.65	4.68	1.92
2005	7.03	3.84	4.46	3.97	1.80
2006	6.30	3.39	4.23	3.85	1.65
2007	6.07	3.13	3.64	3.83	1.59
2008	5.57	2.69	3.26	3.46	1.51
2009	5.28	3.69	4.03	4.22	1.63

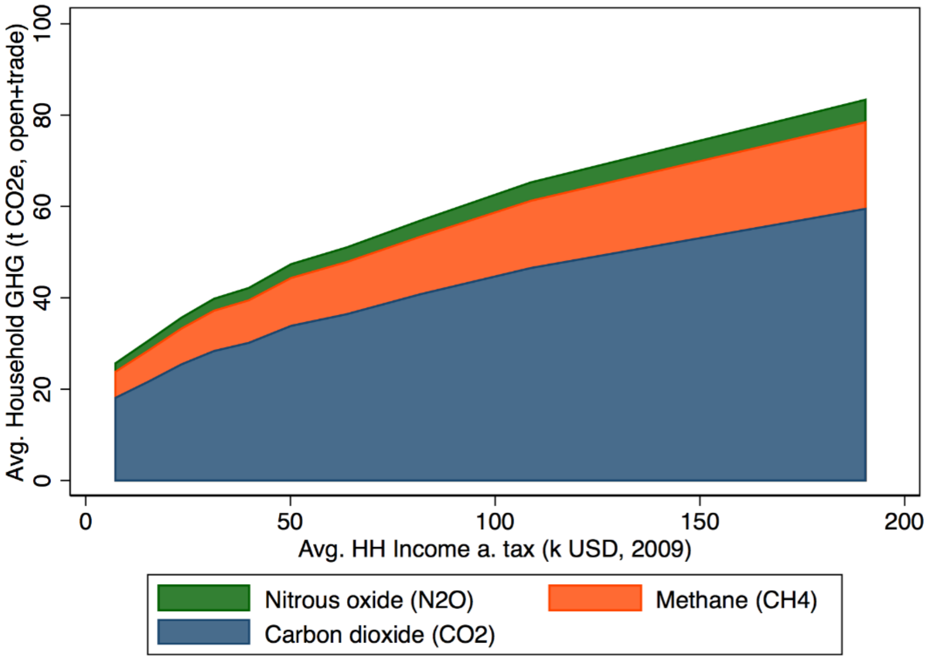
Notes: Based on annual average price data in the United States for residential electricity, gasoline, heating fuel, and natural gas (EIA); data on average air fares, passenger miles, and fuel consumption by US domestic airlines with revenue above $20m (BTS); constant CO2 emission factors for gasoline, heating fuel, natural gas, and kerosene (EPA); yearly average emission intensity of electricity generation (EPA eGRiD).

Figure A.2: Carbon Consumption Breakdown – 2009



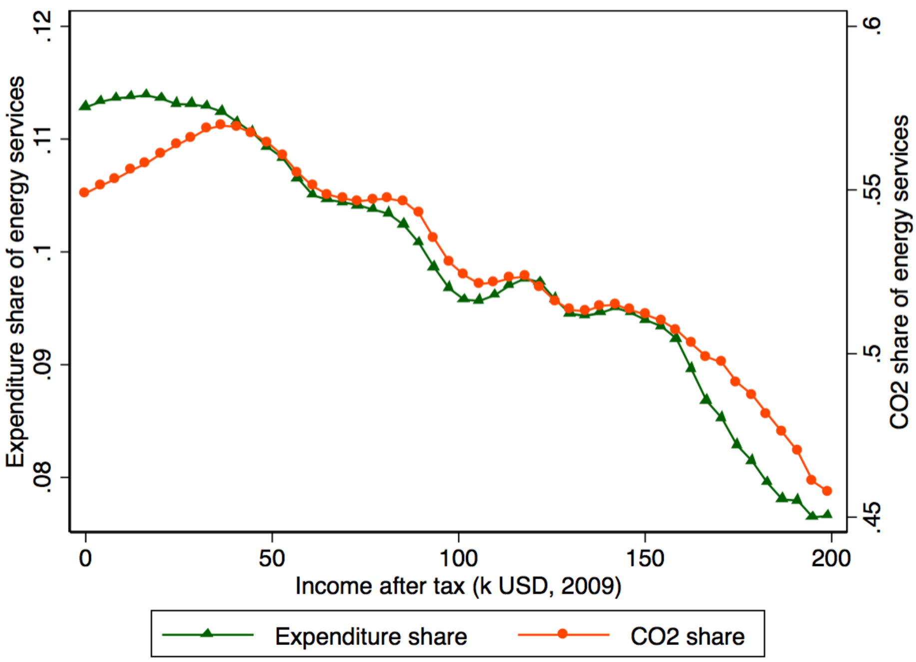
Notes: Decile averages of household income after tax (2009 USD) and estimated CO2-content of consumption (current technology). Household weights as provided by CEX sample. Households with reported after-tax income below 0 USD and above USD 400 k excluded.

Figure A.3: Greenhouse Gas Breakdown – 2009



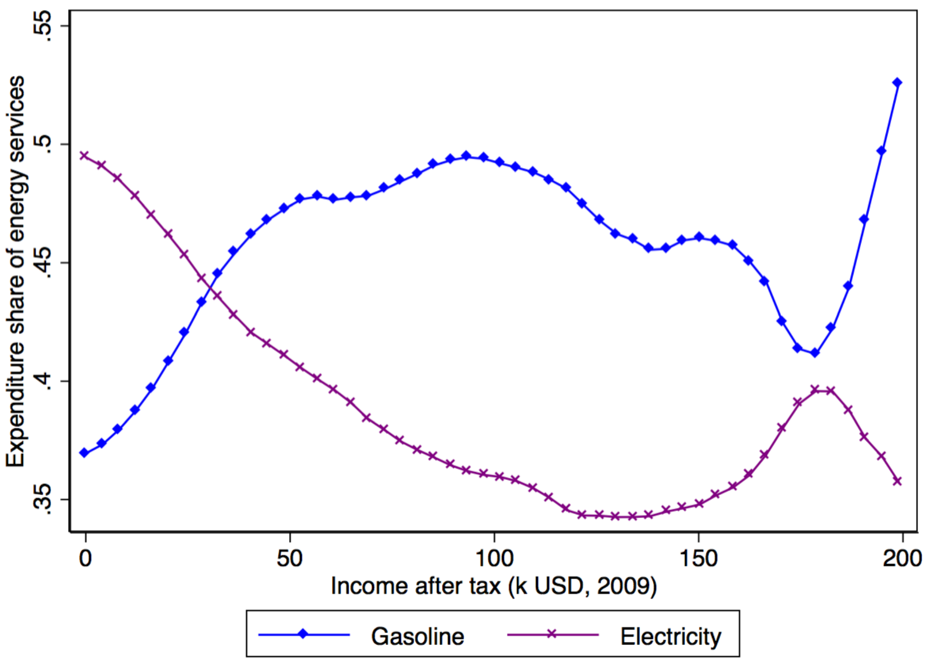
Notes: Decile averages of household income after tax (2009 USD) and estimated GHG-content of consumption (current technology). Household weights as provided by CEX sample. Households with reported after-tax income below 0 USD and above USD 400 k excluded.

Figure A.4: Energy services – Share in expenditure / CO_2 emissions – 2009



Notes: Household total expenditure on energy services (air travel, electricity, gasoline, heating fuel, natural gas) as share of total expenditures (left axis) and CO2 emissions related to energy services as share in CO2 emissions in total consumption expenditures (right axis); both as a function of income after tax (2009 USD). Kernel-weighted local polynomial fit (Epanechnikov, bandwith=7.52). Households with reported after-tax income below 0 USD and above USD 200 k excluded.

Figure A.5: Electricity & gasoline – Share in energy expenditure – 2009



Notes: Household expenditure on individual energy services (electricity and gasoline) as share of total expenditure on energy services (air travel, electricity, gasoline, heating fuel, natural gas); both as a function of income after tax (2009 USD). Kernel-weighted local polynomial fit (Epanechnikov, bandwith=7.94). Households with reported after-tax income below 0 USD and above USD 200 k excluded.

Appendix A.2: Oaxaca-Blinder decomposition – Difference in means

In this paper we use Oaxaca-Blinder decomposition to decompose the change in average emission content of household consumption over time. The methodology was initially suggested to decompose wage differentials between population groups (Oaxaca, 1973; Blinder, 1973).

The decomposition method relies on coefficient estimates from a multiple linear regression analysis. It is assumed that expected emissions of household i in any year $m = 1996, \dots, 2009$ have a linear form in k covariates:

$$\mathrm{y}_{\mathrm{i}}^{\mathrm{m}} = \beta_0^m + \beta_1^m x_{1i}^m + \cdots + \beta_k^m x_{ki}^m + \varepsilon_i^m$$

The difference in means between two years, 2009 and 1996, can then be expressed as:

$$\begin{aligned}\bar{y}^B - \bar{y}^A &= (\beta_0^B - \beta_0^A) + (\beta_1^B \bar{x}_1^B - \beta_1^A \bar{x}_1^A) + \cdots + (\beta_k^B \bar{x}_k^B - \beta_k^A \bar{x}_k^A) \\ &= G_0 + G_1 + \cdots + G_k\end{aligned}$$

Here, then G_k is the contribution to the difference in means by the k$^{\text{th}}$ covariate. The contribution by each covariate k can then be further decomposed into three effects:

$$\begin{aligned}G_k &= (\beta_k^B \bar{x}_k^B - \beta_k^A \bar{x}_k^A) = (\beta_k^B - \beta_k^A)\bar{x}_k^B + \beta_k^A(\bar{x}_k^B - \bar{x}_k^A) \\ &= \Delta\beta_k \bar{x}_k^B + \beta_k^A \Delta\bar{x}_k \\ &= \Delta\beta_k \bar{x}_k^A + \beta_k^A \Delta\bar{x}_k + \Delta\beta_k \Delta\bar{x}_k \\ &= C + E + CE\end{aligned}$$

Here, C represents the difference due to changes in the coefficient of the k$^{\text{th}}$ covariate, E represents the difference due to the difference in covariate means, and CE represents the interaction effect.

Appendix A.3: Factor decomposition of inequality

In this paper, we decompose the inequality in household carbon budgets using the regression-based approach suggested by Fields (2003) and building on factor decomposition initiated by Shorrocks (1982).

It is assumed that the expected carbon budget of household i in year m, y_i^m, is linear in k covariates:

$$\mathrm{y}_\mathrm{i}^\mathrm{m} = \beta_0^m + \beta_1^m x_{1i}^m + \cdots + \beta_k^m x_{ki}^m + \varepsilon_i^m$$

The variance of household carbon budgets, $\sigma^2(y)$, can then be written as:

$$\sigma^2(y) = \sum_{j=1}^{k} cov[\beta_k x_k, y]$$

We then define the *relative factor inequality weight* of covariate k, $s_k(y)$, as:

$$s_k(y) = \frac{cov[\beta_k x_k, y]}{\sigma^2(y)}$$

This weight describes the contribution of the variation in the covariate k, in the variance of household emission budgets, $\sigma^2(y)$.

Shorrocks (1982) has shown that under a number of assumptions, this decomposition will not only hold for the variance, but for any inequality measure $I(y)$ that is continuous, symmetric, and has $I(\mu, \mu, \dots, \mu) = 0$.

The decomposition is carried out using the STATA module from Fiorio and Jenkins (2007).