

## Research article

# Unbiasing the estimate of the role of income in carbon footprint of households: Analysis of the Spanish case as a pilot study

Rafael de Arce<sup>\*</sup>, Ramón Mahía

Applied Economics Department, Universidad Autonoma de Madrid, Calle Francisco Tomas y Valiente, 5, Facultad de CC.EE.EE. Modulo 14 D, 28049, Madrid, Spain

## ARTICLE INFO

## Keywords:

Carbon emissions reduction  
Carbon footprint income elasticity  
Determinants of carbon footprint  
Quantile regression

## ABSTRACT

Although the estimation of the elasticity of the household carbon footprint and income is a frequently analysed fact, unfortunately a fundamental aspect of this relationship has not been considered: it is not a constant factor for the whole population. To make an adequate estimate of this relationship, a Quantile Regression is proposed, obtaining significantly different results to those derived from the usual estimations using ordinary least squares (OLS), which have been carried out up to now. This fact is fundamental for the correct planning and evaluation of fiscal policies based on income taxation to reduce the carbon footprint. Our results confirm that the OLS estimation would overestimate the effects of income on CO<sub>2</sub> reduction by 26%.

## 1. Introduction

Several recent studies have estimated the effects of different socio-economic variables as key determinants on the CF of households produced, directly or indirectly, by their different consumption patterns, which have come to be referred to as demand-side solutions (see Refs. [1–3] or [4]).

Consumption is typically estimated to account for 60–70% of total CF [5]. Apart from the physical characteristics of elements of the construction of their dwellings or the systems used for heating/lighting, the literature has shown clear evidence on the relationship between CF and household income, the level of education of household members, the number of household members (household size), their ages, the geographical environment (surface of the dwelling, area/region of the country), the attitudes towards sustainability of household members, etc. (see Table 1).

The analysis of the elasticity of the Carbon Footprint (CF) with respect to income has been a recurring theme in the literature. The progressive increase in CF in relation to the increase in household income can be considered a stylized fact (see Ref. [6] or [7]) and, therefore, it is subject to reflection on how such income should be taxed to promote a potential reduction of CF.

In this context, an unbiased estimate of this elasticity is a key point for correctly analysing the real effects of income tax policies. The implicit reduction in disposable income of such policies, and their translation into a reduction in consumption, and, subsequently, in CF, can only be estimated based on a correct assessment of the income/CF elasticity.

In this research, we propose an alternative approach to the analysis for a better estimation of income/CF elasticity. Two differential elements are proposed in this investigation in comparison with the analyses usually performed: (i) a quantile regression (QR) analysis is carried out, instead of using the usual Ordinary Least Squares Regression (OLS) typically present in the literature (in a similar line,

<sup>\*</sup> Corresponding author.

E-mail address: [rafael.dearce@uam.es](mailto:rafael.dearce@uam.es) (R. de Arce).

see the recent research [8]), and (ii) a minimum household CF threshold is proposed above which to act in order to reduce these emissions (see Refs. [9,10]). These two contributions are specially interesting.

First, a minimum threshold is established below which there is no *a priori* family discretion to reduce/expand consumption and, consequently, their CF. This idea is similar to that anticipated in the study by Ref. [11] in which the responsibility for the reduction of CF (at a national level in their case) was placed on a “fairer principle” derived from the sum of the excess emissions of those “highest emitters” in each country, i.e., those that exceed a certain maximum threshold.

The setting of such a threshold must be, to certain point, arbitrary, and, in this study, has been determined based on a comparative analysis of a minimum relative to the region/area where the household is located. The idea behind the setting of this threshold is that the differences in prices in the different regions mark a different value of consumption in each type of product valued in euros, with a different valuation being necessary in this minimum threshold of consumption and, therefore, of CF. To certain extent, the well-known idea of parity purchasing power (PPP) would be replicated in setting this minimum.

Second, the non-linearity in the relationship between determinants and CF (see Ref. [6]) is poorly explained by the classical regression commonly used, i.e., ordinary least squares regression. As is well known, this regression pivots on the estimation of the mean. The estimation of these relationships through quantile regression can shed light by differentiating the parameters for different quantiles of the variable to be explained (in this case, the excess CF footprint in households).

As a result, the correct estimation of the elasticities of CF from its main drivers will allow a correct assessment of the effect of some policies aimed at reducing CF in future research and policy evaluation exercises.

The results confirm that the OLS estimate of the reduction in CF due to a cut in household income is clearly biased, since the quantile estimate shows that the income/CF elasticity is significantly different for each quantile. Aggregating the effects of these quantiles, the total impact would be 26% lower than the aggregated one shown by the single OLS regression.

This article is organised as follows. First, a brief review of the methodological features and data used is given. Subsequently, the regression results used are presented and discussed. Finally, some conclusions are drawn.

## 2. Methodology of analysis and data

As mentioned above, the aim of this research is to correctly estimate the elasticities of CF to changes in household income, controlling for the effect of the other determinants of the increase/decrease in CF.

The sources of information used are basically the Household Budget Survey 2019 (INE, 2021) - HBS -, and the E-MRIO tables for Spain for the same year. The methodology developed by Ref. [12] has been used to match the products/services of both sources of information.

Three main steps were followed: (i) to determine the CF embodied in the consumption of each product/service by each household; (ii) to estimate the minimum value of CF for each product/service attributable to each household, considering the characteristics of its geographical environment; and (iii) to estimate the different elasticities of CF by using quantile regression, which differ according to the level of CF achieved.

To determine the CF embodied in consumption, the Input-Output mechanism used by Refs. [1–3,13,14], among others, has been employed. The database for the application of the implicit Leontief model is taken from the E-MRIO database ([15,16]), and applied on the indicator of kilograms of CO<sub>2</sub> emissions per Euro consumed for each product, measured from the EDGAR indicator.

Based on the estimation of the embedded CF to each product/service consumed by the household, and the matching tables of [12], the CF attributable to each product/service consumed by Spanish households was obtained.

As described in the introduction, the next step would be to distinguish between the minimum (hardly reducible) CF from the remaining CF per capita in each household. To estimate this minimum threshold, it has been assumed that “excess per capita CF” will be considered as all that is above the value of 60% of the median CF in the household’s area of residence. The clear differences in parity purchasing power (PPP) between Spanish regions make it necessary to apply this calculation of the median at that level of available information (in our case, NUTS1 of the European Union classification). This will avoid distortions due to the profound price differences between regions.

The variable of interest in our study is the household CF per capita exceeding what could be considered the minimum emission for each household.

As discussed above, there is no standard value for the calculation of this minimum beyond what can be inferred from certain considerations about the actually observed CF data. Clearly, the consumption of the poorest households, and their consequent CF, may not reflect an “acceptable minimum”, but rather the value they can achieve with their disposable income.

It is conceivable that income restrictions may be “capping” the value of consumption at figures below those that would be produced by increasing income. See, for example, the case of what is known as “energy poverty”, where the value reached in consumption to heat the home in lower-income households is not the “acceptable minimum”, but surely the only one they can afford (see Ref. [10]). It is conceivable that, in the face of income increases in this quartile, the elasticity of consumption would be even higher than one.

Being aware of these limitations, we have calculated the value of the footprint that exceeds the 60% of the median value for the overall distribution observed for the total expenditure of the Spanish household population.

Given the large differences in parity purchasing power between regions in Spain, and the fact that the HBS refers to the value in euros of purchases and not to quantities, the calculation of the minimum threshold above was carried out in a differentiated manner for each of the 19 Spanish Regions (Autonomous Communities).

$$\text{Threshold}_j = \text{median}(\text{CFPC}_{ij}) \times 0.6 \quad (1)$$

$$\text{EXPCW}_{ij} = \text{CFPC}_{ij} - \text{Threshold}_j \quad (2)$$

Where “j” refers to the region and “i” to the particular observation (family).

The decision to use this variable as the target of the analysis is in line with the most used definition of minimum consumption, which is set at 60% of the median (50%), generally accepted in the literature [17]. Of course, the resulting 30% threshold could be varied, but it can also be considered as a comparable frame of reference [18].

Finally, an estimation of the elasticities has been conducted using a Quantile Regression (QR) model in which the variables commonly analysed in the recent literature are included (see Table 1).

As common in the literature, the effect of different variables in the CF have been introduced in the regression model. As previously noted, the first one will be household income ([19–24] among others). As noted in Ref. [25], there is no clear sign in the influence of this variable on pollution produced as a function of disposable income. The empirical evidence shows that the wide degree of heterogeneity in the type of consumption produces notable differences in the estimated footprint, which must be specifically considered, redounding in our approach of the need to study different CF patterns depending on the level of consumption.

The remaining determinants commonly cited in the literature, such as socio-economic characteristics of the household ([26]), awareness of ecological issues ([28–31]), urban/rural environment implications ([1,32,33]) or those related to population density ([34–36]), show erratic behaviour in terms of their positive or negative influence. This behaviour is a new element indicating the need to consider different consumer profiles and multivariate analysis as a key element to identify the importance of certain drivers in the polluting pattern of subjects.

Other variables such as the type of household equipment (both in qualitative and quantitative terms) would be desirable information to improve our specification of the determinants of CF (see Refs. [37,38]). These authors demonstrate the clear relationship between CF and, on the one hand, the quality of available devices (beneficial effect) and, on the other hand, the quantity of devices used simultaneously in the household (negative effect). Unfortunately, this information is not available in the Household Budget Survey (HBS) which will be our main statistical source on the consumption of Spanish households.

As is well known, QR marks a clear difference from the usual Ordinary Least Squares (OLS) estimation, focusing, in the first case, on the estimation of different parameters for each of the chosen quantiles; unlike OLS, where the value of the mean is prioritised (see annex for a technical QR explanation).

Undoubtedly, the most widely used technique for estimating the CF/income elasticity in the literature has been linear regression, estimating the parameters by Ordinary Least Squares (OLS). Pottier [6] reviews more than 52 scientific articles in which an estimate of the elasticity of the carbon footprint of households to variations in disposable income is made (or can be estimated). In almost all of them the parameters are estimated using OLS.

A small number of authors (see Ref. [39]) have proposed some other estimation methods, focusing on the potential problems of not considering the nested character of the observed units. By using a Hierarchical Linear Model (also known as a Multilevel Model), these authors avoid potential biases in parameter estimation due to the presence of spatial autocorrelation when several observations share a geographical area, the main object of discrimination in their study. At the same time, they “isolate” information on sources of variation of the residual in their regressions.

As is well known, the introduction of dummy variables for each of the categorical variables in the model (and their interactions, if applicable), provides a solution like Multilevel Regression. Additionally, this method does not address the main objective of our research: to show that the influence of certain determinants on CF varies according to the percentile of this magnitude that we are considering. The multilevel analysis does not allow the estimation of parameters differentiated by quantiles, which is basic to the conclusions we present here.

In [40,41] is proposed a Structural Decomposition Analysis (SDA), widely cited in the literature for the aggregate analysis of CF drivers at the sectoral level in each country (often used for international comparisons and/or over time). This approach is clearly separate from the demand-side approach used in this paper, where the focus is on household consumption patterns as shown by the Household Budget Survey (HBS).

Finally, it is worth noting why a multiple regression model has not been used. It is common to find this type of specification in the literature when estimating the CF drivers for the different disaggregated consumption groups, and not for the overall household footprint, as is the case in this research. Certainly, when doing this disaggregation, it seems reasonable to consider the relationship between the different groups and their corresponding random shocks (see Ref. [42]). Econometrically, in this case it would be imperative to choose a more complex parametric estimation method than OLS to avoid biased coefficients, or (at least) to ensure their

**Table 1**

Frequently used variables to explain CF from a demand – side perspective in recent literature.

Drivers/Determinant variables	Papers
Income	[6,19–24]
Socio-economic aspects (household size, age structure, civil status ...)	[26]
Role of sustainability attitudes (educational aspects included)	[27–31]
Urban/rural location	[1,14,33]
Demographical density	[34–36]
Housing characteristics: Type of energy supply (heating, lighting ...). Dwelling type	[32,49,50]
Lifestyles - consumption patterns: food and leisure habits, transportation, clothing, Household use of technology	[37,38,47]

Table 2

Regression Estimates (Coefs. With p value &lt; 0.05) Global Estimation and Quantile Estimations.

VARIABLES	OLS	QUANTILES								
Personal Characteristics (Household Head)	REGRESSION	Q = 0.1	Q = 0.2	Q = 0.3	Q = 0.4	Q = 0.5	Q = 0.6	Q = 0.7	Q = 0.8	Q = 0.9
AGE	<b>27.38</b>	22.46	24.49	22.95	24.88	26.24	26.15	27.02	20.37	21.28
AGE SQUARED	<b>−0.24</b>	−0.18	−0.20	−0.19	−0.21	−0.22	−0.22	−0.23	−0.19	−0.20
OCCUPATION = (Never have worked)	-	-	92.14	89.51	-	-	-	102.12	-	-
OCCUPATION = Managers	<b>375.27</b>	<b>115.92</b>	<b>159.65</b>	<b>173.98</b>	<b>229.13</b>	275.77	343.61	381.72	473.10	571.42
OCCUPATION =Technicians and professionals	<b>144.45</b>	<b>81.20</b>	117.54	114.12	114.67	124.67	139.16	174.05	189.04	218.99
OCCUPATION =Clerical-type and service and trade employees	<b>84.01</b>	40.19	69.88	55.67	48.22	63.02	92.83	91.19	105.62	-
OCCUPATION =Craftsmen and low skilled workers	<b>94.31</b>	68.10	67.36	57.50	61.31	65.33	79.05	82.63	89.65	-
<i>OCCUPATION_Reference = Non skilled workers</i>										
EDUCATION =Cannot read or write or went to school for less than 5 years	<b>−347.01</b>	-	−241.57	−299.81	−387.19	−371.11	−441.93	−381.96	−401.76	−345.09
EDUCATION =Primary education or attended school for at least 5 years	<b>−352.65</b>	−160.18	−250.76	−285.36	−359.86	−357.05	−434.59	−346.02	−340.77	−309.50
EDUCATION =Secondary education or lower secondary education	<b>−222.96</b>	-	−169.75	−191.07	−249.63	−252.21	−325.53	−255.11	−253.45	-
EDUCATION =High School, Vocational Training (Intermediate and Basic)	<b>−173.72</b>	-	−120.35	−144.92	−205.15	−209.45	−274.07	−211.56	−214.87	-
EDUCATION = Advanced vocational training	<b>−171.87</b>	-	−121.69	−143.14	−195.93	−204.35	−277.46	−212.80	−208.03	-
EDUCATION =Undergraduate Degree	-	-	−108.18	-	−140.35	−150.30	−209.16	-	-	-
EDUCATION =Bachelor's Degree	-	-	-	−113.71	−183.98	−187.39	−266.54	−198.26	-	-
<i>EDUCATION_Reference = Post Graduate education</i>										
COUNTRY OF BIRTH=Rest of European Union (EU 27)	-	-	-	-	-	-	-	-	-	-
COUNTRY OF BIRTH=Rest of Europe	-	−293.21	−195.70	-	-	-	-	-	-	-
COUNTRY OF BIRTH=Rest of the World	-	−80.80	−68.98	-	-	-	-	-	-	-
<i>COUNTRY OF BIRTH_Reference = Spain</i>										
Household & Family Characteristics		Q=0.1	Q=0.2	Q=0.3	Q=0.4	Q=0.5	Q=0.6	Q=0.7	Q=0.8	Q=0.9
HOUSEHOLD SIZE = Equivalent OECD scale	<b>−1717.93</b>	<b>−799.54</b>	<b>−901.19</b>	<b>−1029.01</b>	<b>−1102.90</b>	−1248.44	−1483.24	−1766.77	−2142.66	<b>−2641.59</b>
HOUSEHOLD SIZE (SQUARED) = Equivalent OECD scale	<b>299.24</b>	<b>121.75</b>	<b>156.64</b>	<b>175.45</b>	<b>205.89</b>	<b>243.05</b>	284.76	<b>329.55</b>	<b>395.96</b>	<b>474.86</b>
FAMILY WEALTH: Log INCOME (Exact amount of net monthly income of the main breadwinner)	<b>687.53</b>	<b>392.20</b>	<b>448.92</b>	<b>481.31</b>	<b>510.07</b>	<b>540.78</b>	<b>566.71</b>	<b>615.76</b>	656.93	697.67
FAMILY WEALTH= Other homes available	<b>733.74</b>	<b>428.77</b>	<b>476.66</b>	<b>517.44</b>	<b>569.62</b>	<b>603.71</b>	<b>669.00</b>	753.14	<b>829.81</b>	<b>916.49</b>
FAMILY HABITS = Total number of lunches and dinners	<b>2.76</b>	<b>1.33</b>	<b>1.43</b>	<b>1.22</b>	<b>1.48</b>	1.57	<b>1.25</b>	1.59	1.17	1.56
HOUSEHOLD TYPE = Person or couple (at least one of the partners) aged 65 years or older	<b>124.36</b>	138.80	127.13	129.54	125.47	150.08	119.76	140.12	112.37	158.13
HOUSEHOLD TYPE = Other households with one person or couple without children	<b>127.88</b>	110.84	123.31	131.23	115.76	125.45	143.89	171.29	169.78	<b>330.50</b>
HOUSEHOLD TYPE = Couple with children under 16 years old or adult with children under 16 yo	<b>−141.08</b>	−50.85	-	−91.05	−96.00	−85.96	−123.64	−128.90	−159.54	−173.59
HOUSEHOLD COMPOSITION= Number of household members 0–4 years old	<b>−414.34</b>	−248.69	−318.08	−303.97	−360.34	−415.60	−388.90	−389.20	−429.43	−446.02
HOUSEHOLD COMPOSITION= Number of household members 5–15 years old	<b>−450.31</b>	−253.73	−334.02	−325.77	−403.53	−457.41	−430.90	−438.45	−455.87	−469.93
HOUSEHOLD COMPOSITION= Number of household members 16–24 years old	<b>−472.02</b>	−213.07	−301.74	−313.84	−410.68	−467.10	−470.17	−455.91	−500.91	−511.34
HOUSEHOLD COMPOSITION= Number of household members 25–34 years old	<b>−531.86</b>	−213.04	−318.28	−328.29	−428.67	−487.72	−505.57	−513.03	−526.17	−525.14
HOUSEHOLD COMPOSITION= Number of household members 35–64 years old	<b>−634.39</b>	−254.46	−373.11	−389.33	−491.89	−561.82	−580.59	−587.30	−631.53	−674.04
HOUSEHOLD COMPOSITION= Number of household members 65–84 years old	<b>−693.72</b>	<b>−302.63</b>	−414.10	−429.72	−537.65	−630.42	−642.10	−655.43	−702.44	−712.20

(continued on next page)

Table 2 (continued)

VARIABLES	OLS	QUANTILES								
HOUSEHOLD COMPOSITION= Number of household members aged 85 years or over	−781.70	−360.24	−479.13	−497.68	−616.78	−701.52	−719.60	−731.80	−754.43	−802.75
HOUSEHOLD COMPOSITION=Number of dependent children	-	30.67	26.77	-	-	-	-	-	-	-
<b>House and Building Characteristics</b>		<b>Q=0.1</b>	<b>Q=0.2</b>	<b>Q=0.3</b>	<b>Q=0.4</b>	<b>Q=0.5</b>	<b>Q=0.6</b>	<b>Q=0.7</b>	<b>Q=0.8</b>	<b>Q=0.9</b>
HOUSING = Log Area	599.91	351.21	359.28	383.56	387.75	427.51	475.57	506.05	586.73	671.17
BUILDING TYPE = Detached single-family house	-	532.09	243.29	-	-	-	-	-	-	-
BUILDING TYPE = Semi-detached or semi-detached house	-	483.55	-	-	-	-	-	-	-	-
BUILDING TYPE = Building with more than one dwelling	-	485.31	-	-	-	-	-	-	-	-
BUILDING TYPE = Building with less than 10 dwellings	-	483.84	-	-	-	-	-	-	-	-
<i>BUILDING TYPE_Reference = Building with more than 10 dwellings</i>										
HOUSING= Number of rooms	27.52	15.60	18.58	19.65	22.79	27.45	27.50	22.69	22.84	35.78
TYPE OF HEATING = Electricity	206.94	125.67	120.82	130.20	140.06	172.76	191.51	198.91	213.88	254.63
TYPE OF HEATING= Natural Gas	250.80	111.56	129.22	153.59	177.20	208.39	231.08	268.30	290.59	314.53
TYPE OF HEATING = LPG	348.74	122.58	214.18	277.37	291.44	301.38	333.01	397.63	373.73	495.58
TYPE OF HEATING= Other liquid fuels	162.74	116.02	159.93	163.89	192.88	189.37	160.41	150.28	172.08	153.25
TYPE OF HEATING= Solid Fuels	-	-	-	-	-	-	-	-	-	-
TYPE OF HEATING= Other liquid fuels	-	-	-	-	-	-	-	-	-	-
<i>TYPE OF HEATING_Reference = NO Heating</i>										
HOT WATER SYSTEM = Electricity	-	-	-	-	-	-	-	-	-	-
HOT WATER SYSTEM = LPG	−111.32	−105.83	−116.63	−122.68	−113.50	−101.59	−80.20	-	-	-
HOT WATER SYSTEM= Other liquid fuels	-	-	-	-	-	-	122.77	142.43	194.30	-
HOT WATER SYSTEM= Solid Fuels	-	-	−176.46	−247.29	−218.85	−164.48	−182.12	-	-	-
HOT WATER SYSTEM= Other	-	-	-	-	−111.82	-	-	-	-	-
<i>HOT WATER SYSTEM_Reference = NO Hot water system</i>										
<b>Area Characteristics</b>		<b>Q=0.1</b>	<b>Q=0.2</b>	<b>Q=0.3</b>	<b>Q=0.4</b>	<b>Q=0.5</b>	<b>Q=0.6</b>	<b>Q=0.7</b>	<b>Q=0.8</b>	<b>Q=0.9</b>
AREA OF RESIDENCE = Urban luxury	587.86	169.07	148.49	207.89	362.01	497.62	565.82	745.81	743.06	1396.56
AREA OF RESIDENCE = Urban high	171.21	-	44.49	70.71	97.20	96.76	114.39	135.64	138.80	251.90
AREA OF RESIDENCE = Rural agrarian	-	-	-	-	-	-	-	-	-	-
<i>AREA OF RESIDENCE_Reference = Other</i>										
REGION=NORTHWEST	−182.02	−96.60	−102.19	−119.90	−146.06	−174.73	−178.40	−176.42	−161.56	-
REGION=NORTHEAST	−117.78	-	−44.01	−51.74	−80.31	−104.71	−104.53	−113.65	−127.92	-
REGION=CENTRAL	−166.14	−65.11	−53.82	−61.31	−84.77	−121.99	−124.09	−143.92	−165.91	−135.04
REGION = EAST	-	-	-	-	-	−54.97	-	-	-	-
REGION=SOUTH	-	-	-	-	-	-	-	-	-	135.67
REGION=CANARY ISLANDS	−99.06	-	−61.88	-	-	-	-	-	−120.17	-
<i>REGION_Reference = MADRID</i>										
POPULATION SIZE = 100,000 inhabitants or more	−94.88	-	-	-	-	−80.31	-	−95.19	−107.93	−199.69
POPULATION SIZE = 50,000 or more and less than 100,000 inhabitants	-	-	-	-	-	-	-	-	-	-
POPULATION SIZE = 20,000 or more and less than 50,000 inhabitants	−56.96	-	-	-	-	-	-	-	-	-
POPULATION SIZE = 10,000 or more and less than 20,000 inhabitants	-	-	-	-	-	-	-	-	-	-
<i>POPULATION SIZE_Reference = Less than 10,000 inhabitants</i>										
POPULATION DENSITY = Densely populated area	-	-	-	-	-	-	-	-	-	-
POPULATION DENSITY= Area with intermediate population density	-	-	-	-	-	-	-	−92.39	−111.47	−108.00
<i>POPULATION DENSITY_Reference = Sparsely populated area</i>										
Pseudo R-Squared	q = 0.10	q = 0.20	q = 0.30	q = 0.40	q = 0.50	q = 0.60	q = 0.70	q = 0.80	q = 0.90	
Mean Absolute Error (MAE)	0.277	0.296	0.308	0.319	0.33	0.341	0.352	0.362	0.373	
	946.966	795.91	708.407	658.425	639.981	660.173	728.637	881.248	1265.282	

(continued on next page)

**Table 2** (*continued*)

- 
- (i) In red, coefficients that are statistically different from those obtained in the overall regression. Differences are considered significant based on the 95% confidence intervals for the parameters of both regressions.
  - (ii) The column labelled “GLOBAL” shows the significant coefficients ( $p < 0.05$ ) obtained in the OLS regression after a process of selection of characteristics and examination of the different candidate explanatory variables. The columns labelled  $Q = 0.1$  to  $Q = 0.9$  show the coefficients of the same variables obtained in the quantile regression (provided they are significant at 95%).
  - (iii) Barslund (2007) stata subroutine has been used to assess robustness of the specification finally selected.
-

consistency. Even so, the use of very large samples has reduced the marginal utility of using complex estimation methods in multi-equation models using asymptotically consistent OLS. In any case, in this research, we have not disaggregated by groups, so the use of this type of model would not make sense.

### 3. Results and discussion

The following table (Table 2) shows and compares the coefficients significant at 95% confidence for both the overall regression and those corresponding to each of the quantiles. The results have been organised according to the type of explanatory variable, distinguishing between characteristics of the main breadwinner, characteristics of the household, characteristics of the dwelling and characteristics of the geographical area in which it is located. The variables/blocks chosen for the analysis correspond to those commonly used in the recent literature (see summary table in the annex I [25], or [30]).

A robustness analysis of the regression results has been carried out using the “checkrob” test ([43]). The results indicate robustness in terms of signs and significance for the most relevant variables. Some of the parameters are modified by progressively introducing covariates in the regression, which we must interpret as a necessary correction associated with avoiding the bias linked to the omission of relevant variables.

As previously mentioned, the usual “confounding variables” in the literature have been included in the regression, so that the specific isolated effect of income on the excess of CF “ceteris paribus” can be analysed. We will concentrate our comments on the variable of interest in our analysis: household income.

As in Refs. [19–24], income is obviously a powerful explanatory factor of consumption and thus of excess per-capita household pollution (see Ref. [25]). As in Ref. [7] and in the large body of experience reviewed in Ref. [6], our analysis specifies, as a stylized fact in the literature, a non-linear relationship between consumption and income.

The overall OLS regression indicates that a 10% increase in income generates increases in per capita excess pollution of around 3.6%. This result would be in line with those offered by Pottier [6] for the Spanish case derived from his own estimation of the articles by Refs. [22,44]. Both indicate bivariate elasticities lower than 0.41 and 0.54, respectively. In our analysis, the value would be around 0.36, but including the corresponding control variables (these authors did not include such variables).

However, the QR results are very revealing. Firstly, they show a progressively larger impact of income the higher the quantile of excess pollution is (see Fig. 1). But even much more interestingly, the illustration suggests that the single parameter estimate in the OLS would be clearly biased upwards, being an accurate representation only for households with a higher EXPCW. Fig. 2 shows, in effect, forecast lines with a slightly steeper slope as the level of the quantile increases.

Again, our results would coincide with those of [6] that confirmed notable differences between the lower quantiles (between 1.2 and 1.43) and the upper quantiles (from 0.47 to 0.65).

This relevant finding, only visible thanks to the use of QR, has enormous relevance in terms of fiscal policy aimed at reducing CF. Indeed, as can be seen in Fig. 3, if one wishes to compute the impact on the carbon footprint associated with a reduction in income (through a tax measure), the aggregate effect, quantile by quantile, would be clearly lower than that suggested by the single parameter obtained by the OLS regression. If we were to take the results of the OLS regression, the impact of a 1% reduction in income would be

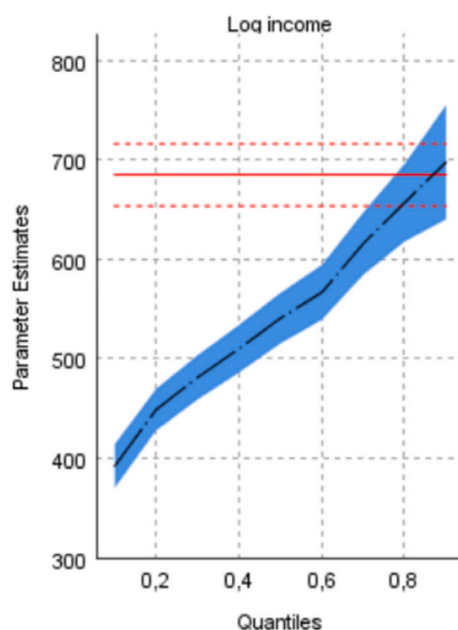


Fig. 1. Effect of Log Income (Main breadwinner) in EXPCW. (Global estimate Vs Quantile Estimations). Source: own elaboration.

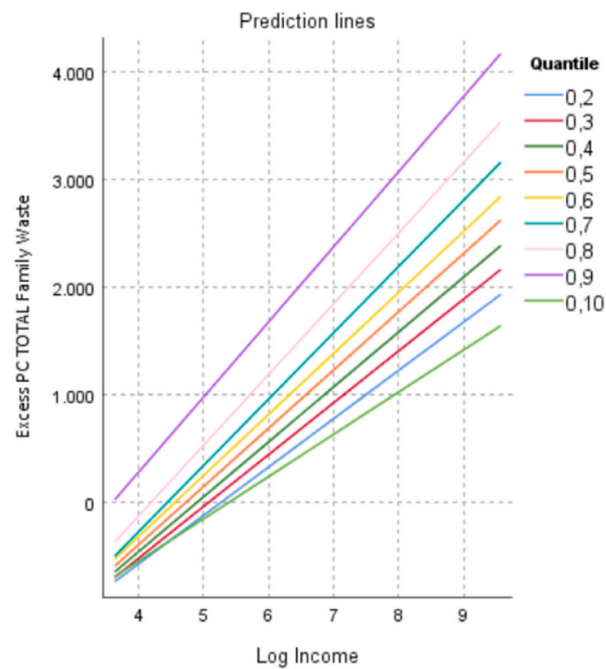


Fig. 2. Prediction of Impact of Log INCOME in EXPCW. (Global estimate Vs Quantile Estimations). Source: own elaboration.

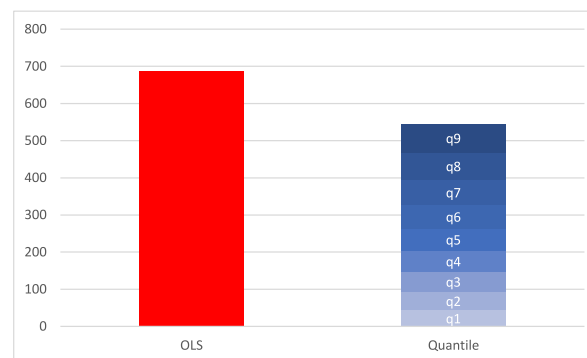


Fig. 3. Comparison of OLS impacts – Quantile Regression. Source: own elaboration.

687.6 kg of CO<sub>2</sub>, i.e., 26% higher than that obtained using the quantile regression (545.6 kg).

Although the positive correlation between income and excess pollution is always true, its importance is much lower than previously suggested and also varies in a wide range suggesting that this variable should not be the only or the main variable used to stimulate

Table 3

Mean of Income by Excess CF quantiles.

		Net monthly income of the main breadwinner (€)		
		Mean	95,0% Lower CL for Mean	95,0% Upper CL for Mean
Excess PC TOTAL Family Waste (Binned)	≤ 502,94	916	883	948
	502,95–783,98	1151	1115	1188
	783,99–1035,45	1246	1201	1290
	1035,46–1278,25	1308	1265	1351
	1278,26–1537,48	1328	1282	1374
	1537,49–1837,77	1497	1440	1553
	1837,78–2219,39	1486	1428	1544
	2219,40–2753,37	1561	1503	1619
	2753,38–3638,79	1653	1584	1723
	3638,80+	1979	1895	2063



fiscal policies to reduce CF.

As can be seen from some of the conclusions in Ref. [5], the result of an extensive study on the different climate policy measures with the potential to influence household emissions, the authors conclude that the voluntary measures that consumers would be willing to take have little real effect, thus opting for coercive measures as the most efficient. Among these, tax policy based on personal income tax seems to be an ideal instrument but having in mind that the sound impact suggested by the OLS regression is in fact only true for the highest/polluting quantiles. Thus, an increase in income-based taxation is even more effective the more it is concentrated in these quantiles.

However, these highest polluting quantiles are also the highest income quantiles, which allows for the identification of the population on which income-based taxation measures are appropriate. As Table 3 shows, the average monthly income of the most polluting quantile is almost €2,000, more than double the income of the first quantile and 20% higher than even the previous one. There are some authors, such as [45], who, based on this evidence, propose some measures that could be considered extreme, such as “capping” maximum wages.

As can be seen in Fig. 4, the inequality by deciles in the concentration of CF, looking at both income and population, is remarkable. This inequality is represented by the concentration curve as proposed by Kakwani [46].

Recognising the fundamental role of income as a driver of an increase in the CF and correctly estimating its elasticity, a progressive direct tax on carbon footprint through the household income tax is proposed. This tax could be modulated with deductions on investments in elements that reduce emissions (household infrastructure such as photovoltaic panels -or others-, more efficient thermal enclosures - doors, windows -, purchase of non-polluting vehicles, ...) financed directly with the proceeds of this tax.

A successful “green tax rate” can deepen the implementation of policies in favour of introducing economic incentives, adopting renewable energy, changing household behaviour, improving national infrastructure, enhancing stakeholder engagement, and promoting sustainability education (see in Ref. [47] or [25] a detailed definition of these measures referred in several literature papers).

As far as a semi-federal system such as the Spanish one allows a different taxation by different geographical areas; it is of special interest to observe the regional differences when dealing with income and CF relationship. For this purpose, the concentration curve, previously showed for the whole Spanish territory is now displayed in Fig. 5 for different regional areas.

This concentration curve can be complemented by calculating the standardised Gini inequality coefficient for the excess CF estimated above, as shown in Fig. 6. The Gini index is calculated, in our case, as an aggregate metric describing the proportion of areas on the concentration curve. As in Pottier [6,48], the concentration curve depicts the concentration of carbon footprint with income representing, on the x-axis, the cumulative share of population ranked by increasing income, and, on the y axis, its cumulative share of the carbon footprint whereas.

At the level of disaggregation allowed by the available data (NUTS1), we can distinguish three large groups: the North (East and West) and the Centre, which, with differences, show a prominent level of linear proportionality in the evolution of the curve. A second group, the South, made up of the Canary Islands, Andalusia, and Extremadura. Here the values of “inequality” could be summarised as average (with a normalised index around 0.5). Finally, a third group would include both the Madrid region and the Spanish Levante (east coast).

These three groups have quite different climatological characteristics: cold and humid in the first group, dry and hot in the second, and typically Mediterranean weather in the third (except for Madrid, the capital of Spain). Given that the concentration of income and excess CF varies significantly by region, it is straightforward to assume that a tax measure targeting income at the most polluting quantiles would have different effects in different geographical areas. Thus, once again, we insist that the implementation of this type of income measures must simultaneously consider several aspects, some as relevant as territory.

Finally, it is worth considering an important limitation usually pointed out in this type of HBS data analysis as mentioned in Ref. [6]. It should be noted that increases in the monetary value of products do not necessarily have to be related to increases in the number of units consumed of the same product. The same quantity of product might be priced higher when it is of higher quality, or refers to certain “luxury” brands, or, even, when it has an explicit “green product” label, which often also results in a price increase for those people who are susceptible to the problem and are willing to pay more if the product is less damaging to nature.

Undoubtedly, this is an essential element to consider when penalising consumption (in monetary value) through income. In any

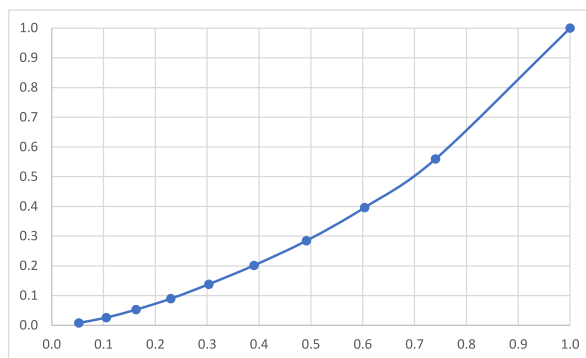


Fig. 4. Concentration Curve of Excess CF by Income. Source: own elaboration.

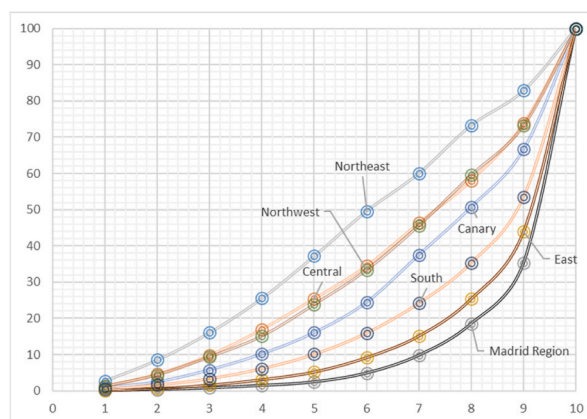


Fig. 5. Concentration Curve of CF by Geographical Areas (NUTS1, Spain). Source: own elaboration.



Fig. 6. Standardised gini inequality index by Spanish regions.

case, if the products in question are of higher quality, it is presumable that they have undergone more careful production processes and are therefore more polluting. In the second case (green products), their current low weight in total household expenditure means that the proposed valuation mechanism does not lose its generality.

#### 4. Conclusions

The main proposal of this research refers to the correct estimation of the CF/income elasticity as a fundamental first step in the design and evaluation of tax policies for the reduction of carbon emissions. As shown, using the usual OLS-based estimation produces a significant bias in the valuation of these elasticities and, therefore, a flawed instrument for planning and subsequent evaluation of the fiscal policies to be used for an effective reduction of carbon dioxide emissions.

As previously highlighted, although the positive correlation between income and excess pollution is always true, its importance is much lower than previously suggested and varies in a wide range (coefficients between 400 and 700 approximately).

This relevant finding, only visible thanks to the use of QR, has enormous relevance in terms of fiscal policy aimed at reducing CF. If one wishes to compute the impact on the carbon footprint associated with a reduction in income (through a tax measure), the aggregate effect, quantile by quantile, would be clearly lower than that suggested by the single parameter obtained by the OLS regression. If we were to take the results of the OLS regression, the impact of a 1% reduction in income would be 687.6 kg of CO<sub>2</sub>, i.e., 26% higher than that obtained using the quantile regression (545.6 kg).

The results indicate that the highest polluting quantiles are also the highest income quantiles, which allows for the identification of the population on which income-based taxation measures are appropriate. Indeed, a 3.5% income reduction in the two most polluting brackets alone would generate a CO<sub>2</sub> reduction equivalent to 1% for the entire population.

In this framework, it is seen as absolutely key to employ an estimation method such as the one used in this research: quantile regression (QR). This avoids the bias of important dimensions in the evaluation/planning of policies based on income taxation. As in Ref. [5], we consider that this instrument, correctly used, can be the most effective tool to move towards the objectives of reduction of

CF. Of course, such a tax policy must be duly complemented by measures that contribute to behavioural change at household level, to the adoption of renewable energy, to the improvement of national infrastructures, to the promotion of coherent sustainability education, etc.

### Author contribution statement

Rafael de Arce: Conceived and designed the experiments; Performed the experiments; Analysed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Ramon Mahía: Conceived and designed the experiments; Performed the experiments; Analysed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

### Data availability statement

Data will be made available on request.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

## ANNEX 1. QUANTILE REGRESSION

We use Quantile Regression as the proper technique to execute the analysis. There are, at least four key elements that make this technique the most appropriate for our case: (i) the likely existence of discontinuities between CO2 emissions and explanatory variables, (ii) the presence of outliers that may distort the results of a plain OLS regression, (iii) the expected presence of heteroscedasticity in the residuals, and (iv) the non-normal distributions.

The main advantage of Quantile regression is to provide a different estimate of the coefficients for each one of the different quantiles considered for the dependent variable. If we want to corroborate (or reject) that the importance of the independent variables for CF is not homogeneous, Quantile Regression emerges as natural appropriate technique.

Quantile Regression estimator is based also in minimization of errors as the OLS does, but it considers the median (or another selected quantile) instead differences with the average.

$$|e_i| = \sum_i |y_i - \text{median}| \quad (3)$$

As Koenker and Basset (1978) showed, in the above equation, the equal weight of both the left and right sides of the endogenous variable derives an accurate estimate of the median. Therefore, by weighting each tail of the distribution by the desired quantile and minimizing the previous function, we can estimate the specific coefficients for any other quantile (called  $\tau$ ):

$$\text{Quantile}(\tau) = \sum_i \rho_\tau |y_i - q| \quad (4)$$

where  $(\rho\tau)$  is:

$$\rho_\tau(x) = \begin{cases} -x \cdot (1 - \tau) & x < 0 \\ x \cdot \tau & x \geq 0 \end{cases} \quad (5)$$

In OLS regression, the estimated value of the endogenous variables corresponds to the expectancy of the mean given the set of variables and the explanatory parameters-variables  $X\beta$ :

$$\hat{y} = \mu = E(y|X\hat{\beta}) \quad (6)$$

What is equivalent to:

$$\hat{y} = q = \text{quant}(y|X\hat{\beta}_\tau) \quad (7)$$

Then, the coefficients for each quantile can be estimated by using the following equation:

$$\min \sum_i \rho_\tau |y_i - X\hat{\beta}_\tau| \quad (8)$$

Rewriting the previous equation, we have:

$$\min \left\{ \sum_{y_i \geq X\hat{\beta}_\tau} \rho_\tau |y_i - X\hat{\beta}_\tau| + \sum_{y_i < X\hat{\beta}_\tau} (1 - \rho_\tau) |y_i - X\hat{\beta}_\tau| \right\} \quad (9)$$

Therefore, Quantile Regression procedure uses a different weight for positive and negative errors, producing different coefficients for each percentile/quantile.

As commented above, another additional advantage of this estimation method is that it avoids the so-called “Heckman selection bias” (Heckman, 1976). As demonstrated by this author, the commonly used sample trimming produces biased parameters, and invalidates their later applicability. In the quantile regression, the total sample is always used, although conveniently weighted, and then, non-biased estimates are obtained.

For the coefficient covariance matrices and standard errors estimates, we have used direct method under i. i.d. Error assumption and Bofinger bandwidth specification (see Koenker, 2005 for an overview).

## ANNEX 2. CARBON FOOTPRINT CALCULATION

In this research, the Leontief pricing model is used to derive the kilograms of CO2 emissions per Euro consumed for each product, measured from the EDGAR indicator - Emissions Database for Global Atmospheric Research -, collected in the E-MRIO database (Lenzen et al.31 and Kanemoto et al.32).

The use of the CO2 EDGAR indicator follows the recommendations of the Intergovernmental Panel on Climate Change (IPCC-2014) where CF was considered the most important and urgent climate change threat (even though it is known the negative side effects of reducing CO2 in terms of increasing other pollutants).

E-MRIO database provides an estimation of each industry CO2 footprint, incorporating both the intermediate consumption of the economy and that resulting from imports from different countries. The original E-MRIO framework presents commodities-industries information. For the conversion to the traditional industries-industries input-output framework, a common attribution technique has been employed using the origin and destination coefficient matrices as follows.

1. Starting from the origin ( $C_{comm \times ind}$ ) and destination ( $D_{comm \times ind}$ ) matrices (commodities in rows and industries in columns), the coefficient matrices are calculated using their inner cells  $C_{ij}$ .

$$Origins\ Coef \rightarrow C_{comm \times ind} = \frac{C_{ij}}{\sum C_i} \quad (10)$$

$$Destinations \rightarrow D_{comm \times ind} \quad (11)$$

To achieve the IO (industries x industries) framework, the following matrix product is performed:

$$IO \rightarrow C'D \rightarrow (ind \times comm) \times (comm \times ind) \rightarrow ind \times ind \quad (12)$$

2. From the input-output framework, the well-known Leontief price model is applied to derive the total effect associated to an increase in demand

$$(I - A)^{-1}FD = P \quad (13)$$

where A is the so-called technical coefficient matrix, FD the final demand matrix and P the final production matrix.

In this way, we can then compute the direct and indirect chain effects after a unit increase in final demand of each product, considering the industrial structure of the country (national and imported) and the intermediate consumption among different industries.

Once the total increase in production associated with the final demand for each product is known, we can estimate the CO2 emissions (using EDGAR indicator of CO2 emissions) per euro of expenditure through the “environmental extension” built into the E-MRIO system.

3. In the next stage, we linked the 119 CPA products considered in E-MRIO dataset to the 316, 5 digits ECOICOP categories available in the Spanish Household Budget Survey. The Spanish HBS 2018 has been used as primary data source. This survey contains data for a geographically representative sample of 23,116 households for Spain as a whole (5-digit product detail) and at the level of the 19 regions (4-digit product detail).

As pointed out in Cazcarro et al.<sup>30</sup>, “Consumption data is usually available in classifications of expenditure according to purpose such as ECOICOP (Classification of Individual Consumption by Purpose). However, multisectoral multi-industry models follow a classification of products aligned with an industry classification such as CPA (Classification of Products by Activity, Eurostat, 2019a). Thus, it is necessary to bridge both. In this article, authors refer to conversion from ECOICOP to CPA products, focusing the model process on Input Output structures (IO) or Computable General Equilibrium (CGE) proposals. In our case, we followed the exact opposite path: from CPA to ECOICOP data.

4. Once CPA and ECOICOP products have been conveniently paired, the next step was to compute the family CO2 emissions considering the expenditure of each household in the different products. As an intermediate step, the corresponding Consumer

Price Index (CPI) supplied by the Spanish Statistics Institute (INE) have been used to update the values of the 2018 current prices of the Household Budget Survey to the current values of the E-MRIO tables available for Spain (2015).

## References

- [1] F. Creutzig, J. Roy, W.F. Lamb, I.M.L. Azevedo, W. Bruine de Bruin, H. Dalkmann, O.Y. Edelenbosch, F.W. Geels, A. Grubler, C. Hepburn, E.G. Hertwich, R. Khosla, L. Mattauch, J.C. Minx, A. Ramakrishnan, N.D. Rao, J.K. Steinberger, M. Tavoni, D. Ürge-Vorsatz, E.U. Weber, Towards demand-side solutions for mitigating climate change, *Nat. Clim. Change* 8 (2018) 260–263.
- [2] A. Beylot, B. Boitier, N. Lancesseur, J. Villeneuve, The waste footprint of French households in 2020. A comparison of scenarios of consumption growth using input-output analysis, *J. Ind. Ecol.* 22 (2017) 2.
- [3] M. Christis, K. Breemersch, A. Vercauteren, E. Dils, A detailed household carbon footprint analysis using expenditure accounts e Case of Flanders (Belgium), *J. Clean. Prod.* 228 (2019) 1167–1175.
- [4] M. Salo, H. Savolainen, S. Karhinen, A. Nissinen, Drivers of household consumption expenditure and carbon footprints in Finland, *J. Clean. Prod.* 289 (2021) 125607.
- [5] G. Dubois, B. Sovacool, C. Carlo Aall, M. Nilsson, C. Barbier, A. Herrmann, S. Bruyère, C. Andersson, B. Bore Skold, F. Nadaud, F. Dorner, K.R. Mober, J.P. Ceron, H. Fischer, D. Amelung, M. Baltrusiewicz, J. Fischer, F. Benevise, V. Louis, R. Sauerborn, It starts at home? Climate policies targeting household consumption and behavioural decisions are key to low-carbon futures, *Energy Res. Social Sci.* 52 (2019) 144–158.
- [6] A. Pottier, Expenditure elasticity and income elasticity of GHG emissions: a survey of literature on household carbon footprint, *Ecol. Econ.* 192 (2022) 107521.
- [7] Y. Shigetomi, K. Kanemoto, Y. Yamamoto, Y. Kondo, Quantifying the carbon footprint reduction potential of lifestyle choices in Japan, *Environ. Res. Lett.* 16 (2021) 064022.
- [8] T.S. Adebayo, E.B. Agyekum, S. Kamel, Hossam Zawbaa, M. Mehmet Altuntaş, Drivers of environmental degradation in Turkey: designing an SDG framework through advanced quantile approaches, *Energy Rep.* 8 (2022) 2008–2021.
- [9] R. Mahía, R. De Arce, Quantifying the Excess Carbon Footprint and its Main Determinants in Spanish Households, *Energy and Environment*, 2022, <https://doi.org/10.1177/0958305X2211405>, 2022, december 2022.
- [10] R. De Arce, R. Mahía, Drivers of electricity poverty in Spanish dwellings: a quantile regression approach, *Energies* 12 (2019) 2089.
- [11] S. Chakravarty, A. Chikatur, H. de Coninck, S. Pacala, R. Socolow, M. Tavoni, Sharing global CO<sub>2</sub> emission reductions among one billion high emitters, *Proc. Natl. Acad. Sci. USA* 106 (2009) 11884–11888.
- [12] I. Cazarro, A. Amores, I. Arto, K. Kratena, Linking multisectoral economic models and consumption surveys for the European Union, *Econ. Syst. Res.* 34 (1) (2020) 22–40, 2020.
- [13] R. Roman, J.M. Cansino, J.M. Rueda-Cantuche, A multi-regional input-output analysis of ozone precursor emissions embodied in Spanish international trade, *J. Clean. Prod.* 137 (2016) 1382–1392.
- [14] M. Tomas, L.A. Lopez, F. Monsalvo, Carbon footprint, municipality size and rurality in Spain: inequality and carbon taxation, *J. Clean. Prod.* 266 (2020) 121798.
- [15] M. Lenzen, D. Moran, K. Kanemoto, A. Geschke, Building eora: a global multi-regional input-output database at high country and sector resolution, *Econ. Syst. Res.* 25 (1) (2013) 20–49, 2013.
- [16] K. Kanemoto, D. Moran, E. Hertwich, Mapping the carbon footprint of nations, *Environ. Sci. Technol.* 50 (19) (2016) 10512–10517, 2016.
- [17] T. Atkinson, B. Cantillon, E. Marlier, B. Nolan, Social Indicators: the EU and Social Inclusion, Oxford University Press, Oxford, 2002, 2002.
- [18] A. Hagenaars, K. de Vos, M.A. Zaidi, Poverty Statistics in the Late 1980s: Research Based on Micro-data, Office for Official Publications of the European Communities, Luxembourg, 1994.
- [19] G. Vita, J.R. Lundstrom, E.G. Hertwich, J. Quist, D. Ivanova, K. Stadler, R. Wood, The environmental impact of green consumption and sufficiency lifestyles scenarios in Europe: connecting local sustainability visions to global consequences, *Ecol. Econ.* 164 (2019) 106322.
- [20] J. Buhl, C. Liedtke, J. Teubler, K. Biengen, N. Schmidt, Measure or management? Resource use indicators for policymakers based on microdata by households, *Sustainability* 10 (12) (2018), 2018.
- [21] M. Dutta, P. Gupta, Relating emissions of carbon to characteristics of consumption in India, *J. Sustain. Dev. Energy Water Environ. Syst.* 6 (2) (2018) 255–275, 2018.
- [22] R. Duarte, A.J. Mainar-Causape, J. Sanchez-Choliz, Domestic GHG emissions and the responsibility of households in Spain: looking for regional differences, *Appl. Econ.* 49 (53) (2017) 5397–5411, 2017.
- [23] M. Rakić, B. Rakić, Sustainable lifestyle marketing of individuals: the base of sustainability, *Amfiteatru Econ* 17 (40) (2015) 891–908, 2015.
- [24] J. Zhang, B. Yu, J. Cai, Y.-M. Wei, Impacts of household income change on CO<sub>2</sub> emissions: an empirical analysis of China, *J. Clean. Prod.* 157 (2017) 190–200.
- [25] O. Shittu, Emerging sustainability concerns and policy implications of urban household consumption: a systematic literature review, *J. Clean. Prod.* 246 (2020) 119034.
- [26] J. Minx, G. Baiocchi, T. Wiedmann, J. Barrett, F. Creutzig, K. Feng, M. Förster, P.-P. Pichler, H. Weisz, K. Hubacek, Carbon footprints of cities and other human settlements in the UK, *Environ. Res. Lett.* 8 (2013) 035039.
- [27] W. Al-Marri, A. Al-Habaibeh, M. Watkins, An investigation into domestic energy consumption behaviour and public awareness of renewable energy in Qatar, *Sustain. Cities Soc.* 41 (2018) 639–646.
- [28] M. Herrmann, D. Brumby, T. Oreszczyn, Watts your usage? A field study of householders' literacy for residential electricity data, *Energy Effic* 11 (7) (2018) 1703–1719, 2018.
- [29] C. Longo, A. Shankar, P. Nuttall, It's not easy living a sustainable lifestyle: how greater knowledge leads to dilemmas, tensions, and paralysis, *J. Bus. Ethics* 154 (3) (2019) 759–779, 2019.
- [30] A. Druckman, T. Jackson, Understanding households as drivers of carbon emissions, in: *Taking Stock of Industrial Ecology*, Springer International Publishing, 2016, 2016.
- [31] S. Moser, S. Kleinhückelkotten, Good intents, but low impacts: diverging importance of motivational and socioeconomic determinants explaining pro environmental behaviour, energy use, and carbon footprint, *Environ. Behav.* 50 (2018) 626–656.
- [32] T. Sole, C. Wagner, Understanding domestic fuel use practices in an urban township, *Build. Res. Inf.* 46 (2) (2018) 220–230, 2018.
- [33] M. Jalas, J.K. Juntunen, Energy intensive lifestyles: time use, the activity patterns of consumers, and related energy demands in Finland, *Ecol. Econ.* 113 (2015) 51–59.
- [34] B. Gill, S. Moeller, GHG emissions and the rural-urban divide. A carbon footprint analysis based on the German official income and expenditure survey, *Ecol. Econ.* 145 (2018) 160–169.
- [35] J. Ottelin, J. Heinonen, J. Nässen, S. Junnila, Household carbon footprint patterns by the degree of urbanisation in Europe, *Environ. Res. Lett.* 14 (2019), 114016.
- [36] D. Wiedenhofer, B. Smetschka, L. Akenji, M. Jalas, H. Haberl, Household time use, carbon footprints, and urban form: a review of the potential contributions of everyday living to the 1.5 C climate target, *Curr. Opin. Environ. Sustain.* 30 (2018) 7–17.

- [37] B.M. Cai, B.B. Liu, B. Zhang, Evolution of Chinese urban household's water footprint, *J. Clean. Prod.* 208 (2019) (2019) 1–10.
- [38] M. Pothitou, R.F. Hanna, K.J. Chalvatzis, ICT entertainment appliances' impact on domestic electricity consumption, *Renew. Sustain. Energy Rev.* 69 (2017) (2017) 843–853.
- [39] I. Gough, S. Adbdallah, V. Johnson, J. Ryan-Collins, C. Smith, The Distribution of Total Greenhouse Gas Emissions by Households in the UK, and Some Implications for Social Policy (LSE STICERD Research Paper No. CASE152), Centre for Analysis of Social Exclusion, London School of Economics, and New Economics' Foundation, London, 2011.
- [40] J. Lan, A. Malik, M. Lenzen, D. McBain, K. Kanemoto, A structural decomposition analysis of global energy footprints, *Appl. Energy* 163 (2016) (2016) 436–451.
- [41] A. Owen, K. Steen-Olsen, J. Barret, T. Wiedman, M. Lenzen, A structural decomposition approach to comparing MRIO databases, *Econ. Syst. Res.* 26 (2014) (2014) 3.
- [42] P. Zsuzsa Lévy, J. Vanhille, T. Goedeméa, G. Verbist, The association between the carbon footprint and the socio-economic characteristics of Belgian households, *Ecol. Econ.* 186 (2021) (2021), 107065.
- [43] M. Barslund, CHECKROB: Stata Module to Perform Robustness Check of Alternative Specifications," Statistical Software Components S456837, Boston College Department of Economics, 2007 revised. (Accessed 6 June 2018).
- [44] L.A. López, G. Arce, M. Morenate, F. Monsalve, Assessing the inequality of Spanish households through the carbon footprint: the 21st century great recession effect, *J. Ind. Ecol.* 20 (3) (2016) 571–581.
- [45] B. Sovacool, A perspective on treaties, maximum wages, and carbon currencies: innovative policy instruments for global decarbonization, *Energy Pol.* 160 (2022) (2022), 112702.
- [46] N. Kakwani, Measurement of tax progressivity: an international comparison, *Econ. J.* 87 (345) (1977) 71–80, 1977.
- [47] P. Allen, M. Robinson, E. Butans, L. Varga, Innovations for sustainable lifestyles: an agent-based model approach, *Sustain. Sci.* 14 (2) (2019) 341–354, 2019.
- [48] J. Yan, J. Yang, Carbon pricing and income inequality: a case study of Guangdong Province, China, *J. Clean. Prod.* 296 (2021) (2021), 126491.
- [49] P. Gianniou, X. Liu, A. Heller, P. Sieverts, S. Nielsen, C. Rode, Clustering-based analysis for residential district heating data, *Energy Convers. Manag.* 165 (1) (2018) 840–850, 2018.
- [50] C. Liao, S. Wang, J. Fanga, H. Zheng, J. Liu, Y. Zhanga, Driving forces of provincial-level CO2 emissions in China's power sector based on LMDI method, *Energy Proc.* 158 (2019) (2019) 3859–3864.