

Households' Food Carbon Footprint.

Ondine Berland*¹

¹Paris-Saclay University, INRAE, AgroParisTech, Paris-Saclay Applied Economics.

28th February 2024

Abstract

Global food systems are a major contributor to greenhouse gas emissions, yet strategies to mitigate food-related emissions receive low public acceptance. To better understand this gap, this study pioneers a comprehensive analysis of household-level food carbon footprints using a representative panel survey for France from 2017-2019. Using machine learning techniques to match food purchases and environmental data, I unveil significant emission disparities across sociodemographic groups, with a notable portion of these differences attributable to unobserved heterogeneity. By segmenting households into quartiles based on their current emission levels, I estimate a structural demand model that reveals distinct profiles in consumption behavior and price sensitivity, particularly contrasting low and high-emission households. These findings underscore the need to consider households' heterogeneous reactions to price changes when designing climate-related food policies.

Preliminary results, please do not cite without permission.

Keywords: Food; Scanner Data; Demand Estimation; Carbon Emissions; France.

JEL Classification codes: D12; H31; I14; Q18; Q50; R22.

*I thank O. Allais, F. Caillavet, R. Chakir, Claire Chambolle, Stéphane De Cara, Hervé Dakpo, Thomas Delemotte, Fabrice Etilé, Maxence Gérard, Mathieu Lambotte, Hugo Molina, Artur Obminski, Katheline Schubert, Nicolas Treich and Stéphane Turolla for useful discussions, as well as participants to several conferences, seminars and workshops. I am especially thankful to Lisa Guo and Marie-Olive Thauray for help with the data processing and the great work during their internships, Emmanuel Paroissien, and Yannick Guyonvarch for their insights and time, and Guy Meunier for his support. Finally, I thank the Odalim team.
ondine.berland@inrae.fr.

1 Introduction

There is growing interest in food systems as a relevant scope for reducing anthropogenic greenhouse gas (GHG) emissions. This reflects that food systems are an important research dimension toward a more sustainable economic system (see, for example Clark et al., 2022). In this context, a better understanding of the drivers of food demand and how they affect the carbon footprints of households is needed. Food demand results from trade-offs conditioned by economic determinants beyond the food sector alone. It is a key adjustment margin for households in the event of a decrease in purchasing power (Griffith et al., 2016). To shed light on this topic, I investigate the heterogeneity in carbon footprints among households in France, and the impact of variations in their budget constraints on these footprints.

I start by presenting the construction of a unique database that matches food purchases with fine-grained environmental impact at the transaction level. I then highlight three key stylized facts. First, I show that the product intensity and the volumes purchased are two key determinants of total food emissions. Notably, the households emitting the most not only allocate a larger portion of their expenditure to animal-based products but also consistently buy larger quantities overall. In fact, households in the top emission quartile acquire volumes twice as substantial as those in the lowest polluting quartile. Second, I show significant heterogeneity in food emissions between households, with those in the top emission quartile singly contributing to over 40% of the aggregate food emissions. Lastly, I emphasize that age and education play significant roles as drivers of emissions, with older and less educated households exhibiting higher emission levels.

This paper contributes to the study of households' heterogeneity in reactions to disposable income shocks, particularly in relation to their baseline food emissions. To explore this, I develop and estimate a demand model for food products using detailed household-level scanner data from France, spanning 2017 to 2019. Utilizing these estimates, I simulate the composition and environmental impact of household food baskets under a hypothetical food price shock. Specifically, I model a scenario mirroring the actual food inflation rate experienced in France between the first quarters of 2022 and 2023, a period characterized by uneven price changes across food categories. For example, red meat prices remained relatively stable compared to dairy or starches. This price disparity potentially influences consumers towards higher-emission products. The structural demand model enables the prediction of such shifts in food consumption and their consequent environmental effects. This simulation exercise is informative to understand how disposable income shocks can affect households' food carbon footprints.

There are several reasons to be interested in households' food carbon footprints. First, food is a necessary good, consumed by everyone, and food-related emissions represent a large share of global emissions (estimated between a quarter Poore and Nemecek, 2018 and a third Crippa et al., 2021). Secondly, while food-related emissions exhibit a comparatively more even distribution than emissions from other goods (Büchs and Schnepf, 2013; Starr et al., 2023), I observe substantial heterogeneity. In my dataset, the quartile of households with the highest emissions pollutes 2.5 times more than the quartile with the lowest emissions. Moreover, food choices are mirrored in a number of health outcomes. Therefore, policies to reduce households' carbon footprint must be designed to limit nutritional inequalities.

In order to study differences in carbon footprints, I build an original database combining two datasets. For the purchase data, I use the Kantar WorldPanel. It records transaction-level information on volumes and expenses along with household characteristics. The database contains around two hundred thousand products, appearing and disappearing over time. Environmental data is obtained from Agribalyse database, which provides detailed information on standardized products¹. To bridge food purchases and carbon emissions, I exploit each dataset's product attributes and machine learning algorithms to find the best-fit correspondence between the two datasets for every product. This matching process covers approximately 98% of the Kantar data. It allows for the fine-grained study of heterogeneity in food emissions intensity, an important step in identifying margins to reduce food emissions.

I document differences in household food-related emissions: substantial differences in emissions driven by animal product consumption and total purchased volumes. Red meat is the most polluting food product category; one kilogram of red meat pollutes four times more than the same quantity of other meat types, which are the next most polluting product categories. I highlight that red meat consumption is strongly correlated with total emissions. However, all diets are polluting, and differences in purchased volumes also explain the heterogeneity of carbon footprints. These stylized facts justify the study of the entire food diet.

Accounting for heterogeneity in food-related emissions distribution, I analyze consumption patterns using a demand system for various food categories. In particular, I estimate the quadratic extension of Deaton and Muellbauer (1980)' Almost Ideal Demand System introduced by Banks et al. (1997), separately per emission quartile. I highlight important differences in household reactions across the emission quartile. The estimations of the demand model show significant

¹Environmental indicators are calculated for "medium standard" products representative of what is consumed in France. This does not allow comparison of food products of the same category (e.g. two yoghurts from competing brands). Therefore, these data reflect the "average" reality that allows food products of different categories to be compared, but not the same products of different production methods.

differences in food consumption patterns across emission quartiles. Notably, households in the highest emission quartile allocate a larger budget share to meat products and exhibit greater price sensitivity, especially for red meat, compared to those in the lowest quartile. The simulation, informed by recent inflation data, further illustrates these dynamics, indicating a 5% decrease in overall consumption volumes but a 7% increase in emissions. This is attributed to a shift from lower-emission foods like fruits and vegetables to higher-emission foods such as meat. These findings highlight the need for public policies that account for the diverse reactions of households to price changes, aiming to encourage sustainable dietary choices and minimize environmental impacts.

This study aims to contribute to two strands of the literature. First, it contributes to a growing literature highlighting heterogeneity in emissions within countries. Studies usually focus on aggregated emissions measured by financial flows per sector using an Environmental Multi-Region Input-Output (EMRIO) model such as Chancel (2022) and Starr et al. (2023). Instead, I use granular information at the transaction level and detailed environmental information to unpack heterogeneity in food carbon footprints. This is important because sector aggregates can hide substantial heterogeneity, especially within food systems, as highlighted by Poore and Nemecek (2018), Chanut (2021) and Clark et al. (2022). Moreover, the machine learning procedure I use to assemble the purchase and environmental datasets allows me to include barcoded and random-weight products in the analysis. Although the latter is frequently purchased in France, including categories with high emission intensity, such as meat, they are usually overlooked in other studies. I also describe emissions heterogeneity along several socio-demographic dimensions, such as age and education, while other studies usually focus on income-level differences. I find that in the case of food, these other dimensions are better predictors of carbon footprints than income. Additionally, the literature has investigated taxes on environmentally harming food products (Caillavet et al., 2016; Bonnet et al., 2018; Chanut, 2021; Funke et al., 2022), the assessment of the consequences of such policies within countries will benefit from a better understanding of the distribution of carbon footprints.

Second, I contribute to a strand of the literature that studies households' reactions to disposable income shocks. A growing number of research articles show that in reaction to the tightening of their budget, households are likely to cut food expenditures (Gicheva et al., 2010; Griffith et al., 2016; Nevo and Wong, 2019). Food accounts for a large share of households' total budgets.² This budget item represents an adjustment margin for households as they cannot

²Food represents 16% of total monthly expenditures in France (Households' Budget Survey, 2017), and 12% of households' disposable income in the US (Consumer Expenditure Survey, 2021).

costlessly adjust on committed expenditures (such as housing, transportation, or subscriptions).³ While the aforementioned articles study whether food expenditures react to income shocks, they do not model the structure of food preferences across product categories. This makes it difficult to derive the overall impact of changes in food consumption on the carbon footprints of households. Following Amano-Patiño (2019), I estimate a demand model of households' entire food diet. This flexible framework allows for studying the sensitivity of food purchases to disposable income shocks and measuring the total impact on food-related emissions.

The remainder of this paper is structured as follows. The next section 2 describes the data. Section 3 details the methodology and specification of the analysis before the results and interpretations are presented in Section 4. The final section summarises and concludes.

2 Data

2.1 Food Purchases.

I use detailed data from the *Kantar WorldPanel* for France (hereafter, KWP). This food-at-home purchase database spans from 2017 to 2019.⁴ The *homescan* data contain purchases' characteristics at the transaction level. Additionally, KWP includes adjustments for large purchase weights to address outliers and unrealistic observations, as well as period and household weights, which are applied in all data processing. The focus of the analysis is on 'active' households, defined as those with consumption for at least ten months annually. The study is restricted to these active households. Furthermore, the dataset encompasses socio-demographic details like household size, income level, age, and employment status. The sample used for estimation comprises an unbalanced panel of 7,540 households, with each household observed for an average duration of twenty months.

2.2 Environmental Impact of Food Products.

I use the 2017 version (3.0) of the *Agribalyse* database, initially developed in 2013 by the French Ministry of Environment. This database is a comprehensive resource, offering environmental impact data for 2,480 unique products that reflect the patterns of food consumption in France. It is structured based on the Life Cycle Analysis (LCA) methodology, which provides a detailed assessment of environmental impacts. *Agribalyse* covers various measures of environmental

³In most income-rich countries, the share of committed expenditure has increased. For example, in France, it has increased from 12% of the total budget of households in 1960 to 30% in 2019 (Accardo and Billot, 2020).

⁴KWP does not provide information on food-away-from-home purchases nor self-consumption.

impact, including factors like climate change, water quality, air quality, and soil impact. I specifically concentrate on the aspect of climate change related to food consumption, quantified in terms of CO₂ equivalent emissions per kilogram of product. This particular metric is crucial as it evaluates the broader effects of food products on the global ecosystem, aligning with methodologies established in existing literature (see for example Poore and Nemecek, 2018).

The Agribalyse database, while a valuable resource, has certain limitations worth noting. First, it primarily focuses on standardized products and does not include detailed information about the origins of these products.⁵ Secondly, the LCA framework used in Agribalyse is still evolving and requires further refinement to fully capture all environmental aspects of the agricultural and food sectors. Despite these areas for improvement, the LCA framework in Agribalyse offers crucial insights into the environmental sustainability of the food sector.

2.3 Carbon Content of Food Purchases.

2.3.1 Matching and Random Forests Predictions

One of the aims of this research is to accurately measure the carbon footprint of food consumption by households. This requires a detailed understanding of the carbon content in food purchases, which I assess at the transaction level using product characteristics. The primary data source for this analysis is the Kantar WorldPanel (KWP) dataset, where products are defined by their brand and a set of attributes. For instance, a product might be described as 'Chocolate', 'Bar', and 'With Milk Filling'. However, a challenge arises as the KWP data lacks product barcodes, a crucial element for straightforward matching with environmental data.

To address this, I employ supervised machine learning techniques, specifically random forest algorithms, for systematic matching. This approach is chosen for its internal validity check and the ability to adapt the model as product identifiers in the KWP dataset evolve over time, reflecting changes in the food sector. For example, between 2017 and 2018, about a third of the products in the KWP database were new (45,000 products). The initial training of the classification model uses a correspondence table between the KWP for 2010 and Agribalyse.⁶ More information about the matching procedure can be found in Appendix A1.3.

⁵For instance, Agribalyse does not have environmental indicators for organic products. In this paper, I assume conventional production methods for food items. This might lead to a potential overestimation of the carbon footprint for high-income households, who may purchase more organic products than others. However, this concern is somewhat mitigated by the fact that organic purchases constitute only a small portion (2.4%) of total purchase volumes, and these purchases are mainly low-emission items like fresh fruits and vegetables.

⁶This correspondence dataset is detailed in De Mouzon and Orozco, 2011. Using only this correspondence dataset from 2010 and applying it to the 2017-2019 data, one would lose about 48% of the raw sample in terms of expenses, and 61% in terms of volumes for 2018.

The matching process is crucial for ensuring accuracy in the carbon content assessment. In the Agribalyse database, products are categorized by their labels, which must be matched with the descriptive attributes in the KWP data. Through the systematic matching process, I successfully identify the carbon content for 98% of the total transactions within the study period. Furthermore, the model demonstrates a prediction accuracy of 96% when evaluated against the validation set.⁷ Another advantage of this method is it allows for the inclusion of both barcoded and random-weight products in the analysis. This latter product type, although rarely accounted for, represents a significant portion of food consumption in France (in 2018, 30% of total sales).⁸

Overall, this matching exercise significantly improves precision compared to using averages across broad categories. Relying solely on aggregate emission intensity for each food category would yield annual emission estimates per household that are markedly different from those obtained through fine-grained matching. Using the average emission per category would result in a 15% overestimation of annual food emissions per household, while employing the median value would lead to a 10% underestimation. These discrepancies become even more pronounced when we analyze specific categories, especially those with high emissions. For example, in the case of dairy products, both the median and mean methods overestimate the impact by a factor of 2. Similarly, significant discrepancies are observed in the 'other meat' and 'white meat' categories. These variations are largely attributable to the degree of variation within categories, which is not effectively captured by aggregated approaches.

To facilitate estimation and analysis, transactions are grouped into quarterly periods and food products are organized into 13 distinct groups. These categories are differentiated based on their nutritional and carbon content, as detailed in Appendix section A1.1.⁹ While this categorization into fewer groups might obscure some heterogeneity in carbon emissions within these categories, a variance decomposition exercise (see Appendix A1.1) reveals that the majority of emission differences are attributable to variations between product categories. Additionally, I have addressed within-category heterogeneity at the household level by calculating specific emission factors for each category and household. As a result, households that consume less polluting products within a given food category are assigned a lower emission factor than

⁷It is important to note that while the prediction accuracy is high, it is not without flaws due to the presence of ambiguously labeled products in the environmental dataset. For example, items such as 'red meat, average product' do not have exact matches in the purchase data. However, this limitation is addressed by the procedure's design, which aims to minimize misclassifications within broader product categories, thereby reducing significant errors in emission predictions.

⁸Data on non-standardized products sourced from the French National Institute for Statistical and Economic Studies <https://www.insee.fr/fr/statistiques/4473482>.

⁹Alcoholic beverages have been excluded from this analysis. While alcohol may sometimes serve as a culinary aid, its primary function is not as a food source.

those with more emission-intensive consumption patterns. This approach ensures a nuanced understanding of household carbon footprints, reflecting both the diversity of food consumption and the environmental impact of different food categories.

2.3.2 Descriptive statistics.

Emissions per product category.

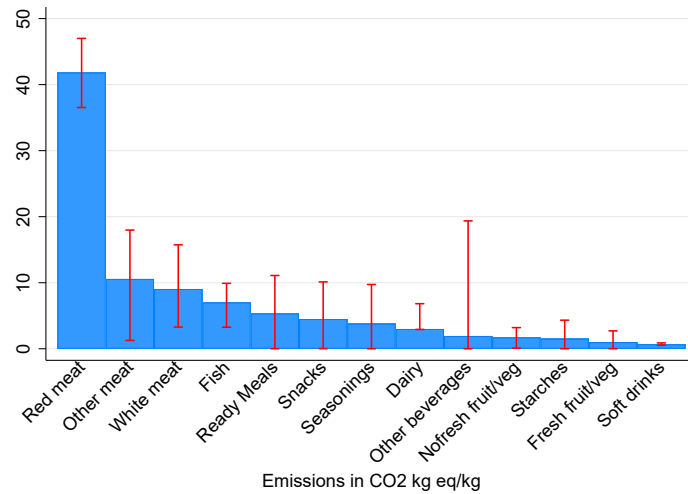
Using scanner data to measure the environmental impact of food purchases offers a significant advantage: it accurately accounts for the actual volumes purchased in each category. This approach contrasts with methods like input-output tables, which rely on financial flows (see Chancel, 2022), or consumer budget surveys. For example, consider two bottles of wine, each containing 0.75 liters but priced differently. Financial flow-based methods might inaccurately attribute a higher carbon footprint to the more expensive bottle, disregarding the fact that their actual environmental impacts could be similar.

Furthermore, it's essential to consider the actual volumes purchased, as consumption amounts vary significantly across different product types. For instance, when evaluating emissions on a per-kilogram basis, coffee appears to have a higher footprint than pork. However, this perspective changes when considering typical serving sizes: a standard coffee serving is about 15 grams, whereas a serving of white meat is between 150 and 200 grams. By focusing on the volumes bought for each category, we gain a more accurate understanding of the environmental impact of dietary choices.

Figure 1 presents the average environmental impact, proxied by greenhouse gas (GHG) emissions, across 13 product categories. The wide confidence intervals (represented in red) underscore the importance of accounting for diversity within each category. Consistent with the existing literature, the emissions from red meat are notably higher than those from other categories, with white meat being, on average, over four times less polluting. The emission trends observed in this study are qualitatively aligned with those reported by Poore and Nemecek (2018), where red meat (specifically beef herd, lamb, and mutton) is identified as having the highest emissions, followed by other animal products.¹⁰

¹⁰It is worth noting that there are differences in the magnitude of emissions when compared to Poore and Nemecek (2018). These differences can be primarily attributed to the scope of the two studies. While Poore and Nemecek (2018) assesses global impacts across 119 countries, this study is focused on food purchases by French consumers. Additionally, Poore and Nemecek (2018) bases their analysis on a smaller set of products (40 products), which could introduce a selection bias. Their data is derived from a meta-analysis that tends to include studies on products expected to be more polluting, whereas our dataset provides a broader and potentially more representative sample of food products.

Figure 1: GHG emissions per food category.



Average emissions per food product category. The red bars indicate 95% confidence intervals.
Source: Own calculations based on KWP 2017-2019 and Agribalyse data.

Household heterogeneity in emissions.

Figure 2 displays the cumulative distribution function of annual food emissions per household. The average annual emission is estimated at around 2.2 tons of CO₂ equivalent per consumption unit, a figure that is close to official French statistics.¹¹

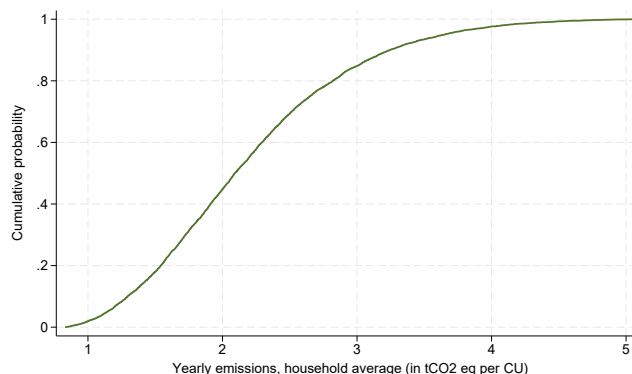
The distribution of yearly emissions is relatively concentrated around the mean but exhibits a log-normal shape with a pronounced right tail. This suggests that households in the top two deciles have significantly higher emissions than the rest of the population. Specifically, these households emit around 3 tons of CO₂ equivalent per consumption unit per year or more. This pattern is further illustrated in Appendix Figure 8, which presents the distribution of total food and red meat emissions as a Lorenz curve. This figure shows marked heterogeneity in emissions. For instance, 20% of French consumers are responsible for 38% of total food emissions. This disparity is even more marked in red meat emissions, where the top 20% of households account for half of all emissions related to red meat. There is a strong positive correlation (0.74) between total food and red meat volumes and total emissions.¹² While animal proteins, particularly red meat, are generally more polluting, the overall differences in food carbon footprints are influenced

¹¹The study by Barbier et al. (2019b) calculates the average food carbon footprint per person at approximately 2 tons of CO₂, including meals consumed both at home and outside. It is important to note that the number of consumption units in a household does not increase linearly with the number of members. In multi-member households, the total consumption units are typically fewer than the number of individuals, which is why the average annual emissions per household expressed in consumption units may appear relatively high. For a per capita perspective, see Appendix Figure 7, which presents the distribution of yearly food emissions per household member.

¹²This correlation remains significant (0.64) even when trimming both variables at the 5% tails, suggesting it is not driven by outliers.

by both the emission intensity of the consumed products and their total volumes. Notably, households with the highest red meat emissions are not always the same as those with the highest overall food-related emissions. The Lorenz curve for total emissions, coupled with a relatively low Gini index (approximately 0.3), indicates that emissions are a common feature across all food diets. This observation underscores the importance of examining the entire food diet to understand patterns of substitutability and complementarity across different food categories.

Figure 2: Distribution of yearly emissions per household.



Source: Average yearly emissions per household. Emissions are trimmed at the bottom and top 1 percents. Own calculations based on KWP 2017-2019 and Agribalyse data.

Table 1 presents the socio-demographic characteristics of households across different quartiles of food emissions, with emission quartiles defined by average quarterly emissions per household. It appears that slightly poorer households tend to have more polluting diets, with the highest emitting quartile being 3.3% poorer in absolute terms and 8% poorer when adjusted for consumption units. Despite a significant difference in income per consumption unit at the 1% level, the correlation between income and food emissions is low (-0.03) after controlling for household size. This pattern is depicted in Appendix figures 9 and 10. Higher emissions are positively correlated with characteristics such as rurality, age, and a lower level of education.¹³ Notably, the highest emitters spend 45% more on food and purchase 53% more in terms of kilograms than the lowest emitting quartile. The low correlation between income and food emissions aligns with findings in other countries; for example, Reynolds et al. (2019) found only a 3% difference in food emissions between the lowest and highest income quartiles in the UK, including both in-house and outside food consumption. These findings are in line with other studies (Laursen et al., 2019; Koch et al., 2019) reporting that lowest levels of meat consumption are found among women, young and more educated persons .

¹³The t-statistics between the lowest and highest emitting quartiles are 75 for education and -52.15 for age.

The differences in sociodemographic characteristics might be partly due to the exclusion of food away from home (FAFH) in this analysis. FAFH and food-at-home (FAH) can have different carbon footprints for several reasons. First, the products consumed outside the house may differ in terms of carbon intensity, for example, if households consume more meat when in a restaurant. Second, the carbon intensity of home cooking and restaurant meals may differ, for example, due to packaging in take-out and deliveries. Additionally, households that spend more on FAFH might differ from those primarily relying on home cooking. Unfortunately, the KWP data does not include FAFH purchases, limiting the ability to test these mechanisms. However, in the French case FAFH is generally more polluting than FAH, as indicated by Barbier et al. (2019a) who estimate that FAFH emits twice as much as FAH due to more energy-intensive production and lack of economies of scale. In terms of eating behaviors, Biermann and Rau (2020) found significant differences in the German context, particularly among 'flexitarian' households who consume more meat in restaurants than at home.¹⁴ However, no significant differences were observed in the frequency of eating out among omnivorous households. These findings imply that studies focusing only on FAH might underestimate the total food carbon footprints, especially for households that underreport meat consumption.

Using the latest wave of the French household budget survey, I found that, on average, the largest share of food expenses is devoted to food at home (80%).¹⁵ Additionally, FAFH spending is not uniformly distributed across the French population; half of the population does not spend on FAFH, and the median expenditure on FAFH is only 10% of total food expenses. As expected, the proportion of food expenses dedicated to FAFH negatively correlates with age and positively correlates with income and education.

I complement the analysis using the Kantar data from 2020. The consumption trends from this survey period were likely heavily influenced by the Covid-19 pandemic (O'Connell et al., 2022), so this year is not part of the primary analysis. However, the shift in food consumption habits can help evaluate the significance of FAFH in our dataset. In fact, during the most stringent lockdown periods, FAFH access was greatly restricted. Appendix Table A5 shows similar age-related trends as before, but the influence of income becomes less significant, and the educational disparity across emission quartiles narrows.

This heterogeneity in emissions across product categories and households underscores the need to study how different groups react to price and expenditure changes. To address this, I

¹⁴Flexitarians are defined as individuals who primarily follow a vegetarian diet but occasionally eat meat or fish.

¹⁵While the 2017 Budget des Familles reports FAFH spending for a nationally representative sample, it does not detail the types of products purchased, making it impossible to calculate the carbon footprint of FAFH.

estimate a demand system for each quartile of emissions, exploring the varied responses to price and income changes.

Table 1: Households' characteristics by emissions quartile.

	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Monthly Income	2786.52	2056.76	2705.77	1850.86	2732.588	1538.07	2685.47	1598.09
Monthly Income/CU	1934.09	1341.24	1824.30	1135.51	1760.64	906.81	1780.40	961.58
Monthly Food Spending/CU	159.82	79.19	192.39	85.44	228.78	98.11	294.52	123.75
Monthly Food Volumes/CU	48.89	57.94	60.40	63.29	73.40	78.52	104.32	169.36
Age	51.25	17.42	54.70	17.49	58.00	15.20	61.68	14.19
Size of the HH	2.11	1.26	2.18	1.27	2.34	1.51	2.17	1.06
Rural (0/1)	0.15	-	0.17	-	0.17	-	0.21	-
Car owner (0/1)	0.85	-	0.87	-	0.91	-	0.91	-
House owner (0/1)	0.58	-	0.64	-	0.70	-	0.69	-
Education	2.76	0.91	2.58	0.87	2.43	0.84	2.23	0.74

Notes: *Income* is expressed in euros. *Income/CU* stands for income divided by the number of consumption units in the household, defined using the OECD modified scale. *Food Spending/CU* is the food expenditures in constant euros divided by the number of consumption units in the household. *Purchased Volumes/CU* stands for the volumes in kilograms divided by the number of consumption units in the household. *Age* indicates the age of the household head. *Education* is a categorical variable indicating the educational background of the household head. It takes values between 0 and 4, where 0 stands for no education, 1 stand for having a primary degree, 2 for a secondary degree, 3 for a bachelor and 4 for a master's degree or above.

3 Empirical Strategy

3.1 The Demand Model

I specify a continuous model of demand in the entire food market. As discussed in Nevo (2011), there are two main approaches to model households' demand. The first one is based on a random utility approach. This approach is best suited for modelling preferences between different goods within specific product categories, such as soft drinks or breakfast cereals. However, it does not address preferences across large product groups, that might be complements to one another. Another rationale for using continuous models is that they allow the study of income effects via the flexible modelling of Engel curves. Additionally, the stylized facts suggest that heterogeneity in food carbon footprints is not only explained by the intensity of the product but also by the volumes purchased, justifying the study of entire food diets.

I estimate the quadratic version of the Almost Ideal Demand System (AIDS) as developed in Deaton and Muellbauer (1980). The advantage over reduced-form equations is that demand systems are based on an underlying model of household consumption behavior for the entire food system, in which the sum of budget shares per category equals 1. It also estimates a system of joint equations instead of separate regressions. The budget shares per food category are functions of prices and expenditures. The level of polynomial in expenditures is determined empirically:

the quadratic version of the model is chosen over the linear one because it improves the fit of the model.¹⁶ Thus, I estimate the quadratic version of an AIDS model introduced in Banks et al. (1997).

The estimated model is expressed as follows:

$$s_{ijt} = \sum_{r=1}^2 \beta_{jr} (X_{it})^r + \sum_{j'} \gamma_{jj'} P_{ij't} + \Pi \mathbf{Z}_{it} + \epsilon_{ijt} \quad (1)$$

Where s_{ijt} is the share of the budget of the category of products j in period t for the household i and H is the number of households. Budget shares are calculated as the category expenditure divided by quarterly food expenditures. X_{it} is the log of household i 's total food expenditures at the period t . $P_{ij't}$ is the Stone price for product category j' faced by household i at period t .¹⁷ \mathbf{Z}_{it} is a vector of controls at the household-year level (socio-demographic variables and geographic fixed effects accounting for time-invariant regional characteristics and $\epsilon_{it,j}$ an error term.¹⁸ The demand model is first estimated for the full sample of households, then I estimate it for each emission quartile separately. This allows to investigate the heterogeneity in reactions to changes in the economic environment (prices and expenditures) along the emission distribution.¹⁹

It is worth denoting that the KWP *homescan* data does not provide the *store prices* i.e. the prices households face in the store. Instead, I have access to unit values per transaction, corresponding to the expenditures divided by the volumes. These unit values correspond to 'observed' prices', this raises potential endogeneity concerns that are addressed in the next subsection.

3.2 Endogeneity.

A key motivation for estimating demand systems is the derivation of expenditures and price elasticities of budget shares for each product category. In the model, budget shares are defined as functions of prices and period expenditures. However, both explanatory variables can be endogenous. Finding good instruments requires finding variables that explain food expenditures

¹⁶The fit is measured with the adjusted R^2 statistic, after simple regressions (without fixed effects) of budget shares on food expenditures for key product categories of interest.

¹⁷Formally, Stone indexes are defined per category as: $P_{ij} = \sum_k s_k \ln(p_{ik})$, with k the variety in product category j , s_{ij} the budget share for variety k in category j , and π_{ik} the unit price of variety k for household i . For each product category, the varieties form a partition.

¹⁸The model is estimated on Stata MP 16 using the *'aidsills'* package introduced in Lecocq and Robin (2015).

¹⁹One potential issue with this approach is the possibility of households shifting between emission quartiles over time. I address this concern by running robustness checks in which emission quartiles are based on a household's emissions during their initial appearance in the data, specifically in the first quarter of 2017. As discussed in the next section, with this approach the results remain very similar to the main ones. Results are presented in Appendix Table A7.

without directly affecting the budget share of the product category.

Quarterly food budget.

Food expenditures are instrumented with households income, following Banks et al. (1997). The identifying assumption, in this case, is that households face variation in income over time, while the rest of their characteristics, in particular their consumption environment, remain unchanged (Amano-Patiño, 2019). To better account for households' living standards, income is divided by consumption units. The first stage F-statistic is high (2538.29).

Prices.

When using scanner data, it's crucial to consider the method of price measurement. Specifically, if one aims to analyze the price trends of products through price indices, it's necessary to track the price of each product across different time periods. This is especially relevant in food categories, which consist of numerous products, each with a relatively small market share. As noted in the data section, the KWP database shows that products frequently emerge and phase out over time. To address this, I use the consistent product subcategories defined by KWP. By aggregating purchases at this subcategory level, I am able to calculate the prices for entire categories using Stone indexes.

Constructed as such, prices might reflect differences in product quality and, therefore, be correlated with households' preferences. These correlations are the source of potential bias. Endogeneity might also stem from prices being affected by local demand shocks. I use prices per category at the living zone level excluding the household observation as price proxies. The identifying assumptions are twofold: *i*) within the living zone, at the food category level: retailers do not react strategically to demand shocks, *ii*) measurement errors are independent across households within a living zone conditional on the control variables. Although the second assumption likely holds, the first one may not be satisfied in every living zone. Therefore, the preliminary estimates presented in the next section might suffer an under-estimation bias and represent a lower bound for households' reactions.

3.3 Censored Demand.

[In progress] I restrict the sample to household-period pairs (i, t) (where a period spans three months) in which food expenditures are positive. This translates into having a positive budget share for at least one of the product categories I consider. This sample is referred to as the 'full'

sample, and it represents over 99.99% of the raw sample.

Households do not consume all the food categories in each period. In this situation, referred to as censoring, information about the category price is missing. This is due to the aforementioned limitation of the data: the prices in the stores households visited are not observed. I call *positive* sample the sample of household-period pairs (i, t) (where a period spans three months) with all positive budget shares. After aggregation at the quarterly level, 63% of the 'full' sample household-period observations satisfy this property. Table A6 in Appendix displays a balance table for the full and the positive sample. The positive and full samples differ mainly by household size, and larger households are more frequent in the positive sample. Households also differ by income, but the difference fades after accounting for consumption units. *In the current version of this study, the estimation sample is restricted to the positive sample, and I do not account for censoring.* To the best of my knowledge, few studies have dealt with censoring. Future versions of this research should account for censoring following Zhen et al. (2014) and Amano-Patiño (2019).

4 Results

4.1 Demand estimates and elasticities

Table 2 presents the budget shares, expenses elasticities and uncompensated price elasticities obtained after estimating the demand model. The expenditure elasticity is defined as: $\eta_E = \frac{\partial s_j}{\partial X} \times \frac{1}{s_j} + 1$ and the uncompensated price elasticity is defined as: $\eta_{P_{ij}} = \frac{\partial s_j}{\partial P_{jj'}} \times \frac{1}{s_j} - \delta_{jj'}$ where $\delta_{jj'}$ is the Kronecker delta.²⁰ The estimates are for the entire household sample and five key product categories. Taken together these categories represent over half of households' total food emissions.

Looking at the second column categories such as fish and fruits and vegetables are more elastic to changes in total food expenditures. When focusing on price elasticities, red meat stands out as being more price-sensitive compared to other products. This suggests that while policy measures targeting red meat prices could lead to a reduction in emissions for top emitters they might have adverse effects in other emission groups. Furthermore, both the price and expenditure elasticity estimates observed here are in line with those reported by Amano-Patiño (2019) in the context of the United States, and by Caillavet et al. (2016) in the French context.²¹

While the initial estimation using the full sample provides insights at an average level, it may

²⁰ $\delta_{jj'}$ is equal to 0 if $j = j'$, in which case $\eta_{P_{jj'}}$ is the own price elasticity, and 1 if $j \neq j'$, in which case $\eta_{P_{jj'}}$ is the crossed price elasticity between j and j' .

²¹For a detailed comparison of these estimates, particularly for meat and dairy, refer to Appendix Table A11.

not fully capture the heterogeneity in households' responses along the emissions distribution. To address this, I conducted separate model estimations for each quartile of emissions. I thus run the same model estimation separately for each quartile of emissions. Table 3 displays the estimates for the two extreme emission quartiles, focusing on the same five product categories.²² The estimated budget shares, presented in columns 1 and 4, differ largely across emission quartiles. Households in the top-emitting quartile spend more (40%) on meat products (red meat and white meat) than the ones in the bottom-emitting quartile. Marked differences in budget shares are also observed in the estimated elasticities, with households in the top-emitting quartile being more reactive to price changes for most product categories. Differences are especially marked for red meat. A potential rationale could be that households in the lowest emission quartile aim to maintain red meat intake (one could think of the example of households who want to consume red meat once a week to maintain iron levels) making them less responsive to price fluctuations. In contrast, households in the highest emission quartile, who typically consume more red meat, have greater flexibility to adjust their consumption in response to price changes. These disparities in households' reactions to shocks suggest that price shocks will affect households differently along the emissions distribution, with consequences for households' food carbon footprint.

These results highlight the heterogeneity of economic behaviors along the emission distribution. They also underscore the nuanced ways in which price shocks can influence the carbon footprint associated with household food consumption.

²²For better readability, I only show here the model for the bottom and top emission quartiles. Results for the four emission quartiles are presented in Appendix Table ??.

Table 2: Demand estimates for the total sample.

	(1)	(2)	(3)
	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$
Red meat	0.057*** (0.000)	1.005*** (0.051)	-1.327*** (0.015)
White meat	0.075*** (0.000)	1.088*** (0.042)	-0.815*** (0.008)
Fish	0.064*** (0.000)	1.694*** (0.048)	-0.862*** (0.008)
Dairy	0.111*** (0.000)	0.915*** (0.027)	-0.943*** (0.007)
Fresh fruits & veg	0.112*** (0.000)	1.697*** (0.040)	-1.083*** (0.015)
Obs.	50789		

Notes: Results after estimating a quadratic AIDS model, for the full sample. $\hat{\sigma}$, $\hat{\eta}_E$, $\hat{\eta}_P$: respectively represent the estimated budget shares, expenditure elasticity and uncompensated price elasticity. The number of observations is the number of households times the number of periods on which the model is estimated. ***, **, *, indicate significance at the 1%, 5%, and 10% levels, respectively. Standard errors are presented in parentheses.

Table 3: Demand estimates for the top and bottom emission quartiles of emissions.

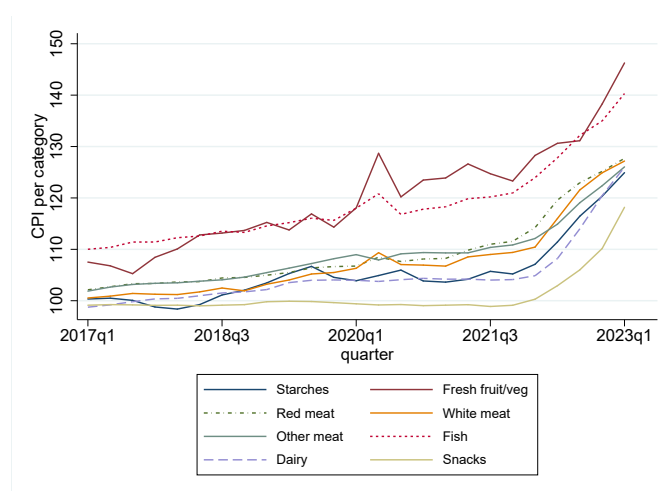
	Bottom emitting quartile			Top emitting quartile		
	(1)	(2)	(3)	(4)	(5)	(6)
	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$
Red meat	0.038*** (0.001)	0.762*** (0.123)	-1.068*** (0.024)	0.072*** (0.002)	0.945*** (0.069)	-1.409*** (0.022)
White meat	0.063*** (0.002)	1.125*** (0.084)	-0.722*** (0.029)	0.093*** (0.002)	0.778*** (0.069)	-1.001*** (0.022)
Fish	0.072*** (0.002)	1.380*** (0.073)	-0.854*** (0.011)	0.053*** (0.002)	1.952*** (0.151)	-0.946*** (0.026)
Dairy	0.110*** (0.002)	0.932*** (0.060)	-0.853*** (0.026)	0.106*** (0.001)	1.077*** (0.056)	-1.073*** (0.015)
Fresh fruits & veg	0.140*** (0.003)	1.541*** (0.056)	-0.961*** (0.029)	0.094*** (0.002)	1.601*** (0.101)	-1.408*** (0.040)
Obs.	12702			12697		

Notes: Results after estimating a quadratic AIDS model. The demand models are estimated separately per quartile of emissions. $\hat{\sigma}$, $\hat{\eta}_E$, $\hat{\eta}_P$: respectively represent the estimated budget shares, expenditure elasticity and uncompensated price elasticity. ***, **, *, indicate significance at the 1%, 5%, and 10% levels, respectively. Standard errors are presented in parentheses.

4.2 Predicting changes in emissions after a price shock.

Using the estimates of the structural parameters, I study an alternative economic environment in which households face a purchasing power shock. I use the recent inflation data (from INSEE) to compute changes in total prices and differential price changes.²³ This analysis specifically leverages the significant price fluctuations observed between the first quarters of 2022 and 2023. Figure 3 presents the evolution of food prices for the main product categories of interest. There is a noticeable evolution in food prices across key product categories, as prices increase by 13% on average in one year. Before 2022, most categories experienced a modest upward price trend, with notable variability in certain items like fresh fruits, vegetables, and fish. However, the onset of the Ukrainian/Russian war in early 2022 marked a pronounced increase in price variations across these categories, each exhibiting different inflation rates. For example, the inflation rate for red meat was relatively lower, showing an 11 percentage point change over the year, compared to a 17 percentage point change in dairy products. The differential inflation rates, coupled with the necessity for maintaining sufficient food intake, may prompt households to alter their food basket composition, potentially leading to shifts in food-related emissions.

Figure 3: Time series, food CPI per category (Year 2015=100).



Notes: Evolution of the Consumer Price Index disaggregated per food product.

This simulation exercise is based on the following assumptions:

- *Price sensitivity is fixed.* This assumption cannot be tested with the 2017 and 2019 data. Nevertheless, if the price sensitivity varies in the same direction for all the food categories the effect should not matter for the predictions.

²³This requires aggregating INSEE 59 food groups into 14 food categories, by taking averages of the price changes.

- *No composition effect within categories.* The current version of this article does not allow for composition effects within categories. If households mitigate the rise in prices by turning to cheaper products then our predictions for the change in the budget share will represent an upper bound.
- *Nominal food expenses remain constant.* This assumption is related to wages not being indexed on consumer price indexes, at least in the short run. Future versions of this research will challenge this assumption by allowing nominal expenditures to vary and setting real expenditures fixed.

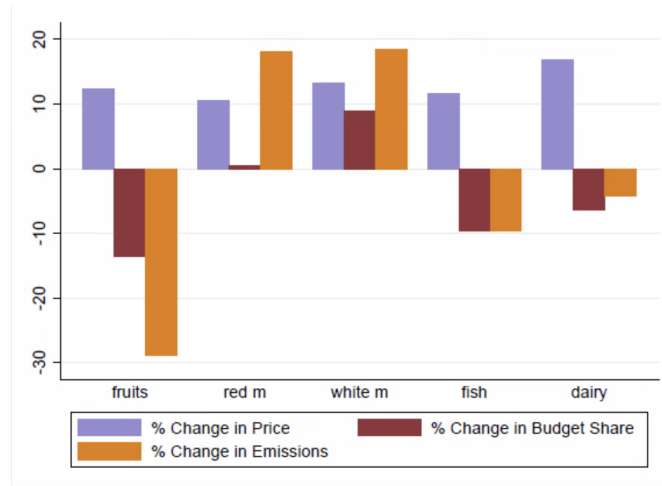
Given these assumptions, the simulation exercise uses the prediction from the structural demand estimation (indicated by letters with a hat) and allows variation in the prices of each food category (P_{ij}) as shown in equation ??.

$$s_{ijt} = \sum_{r=1}^2 \hat{\beta}_{jr} (X_{it})^r + \sum_{j'} \hat{\gamma}_{jj'} P_{ij't} + \hat{\Pi} \mathbf{Z}_{it} \quad (2)$$

Figures 4 and 5 illustrate the outcomes of the simulation exercise. The results indicate an average decrease of 5% in consumption volumes, accompanied by a 7% increase in emissions. This counterintuitive outcome primarily stems from the relatively smaller price increases in red meat compared to other products, despite red meat's significantly higher pollution impact. The data suggests a shift in household consumption patterns, moving away from fruits and vegetables towards more emission-intensive foods like meat. Figure 4 details these trends for the entire sample of households, showing that even a modest increase (less than 2 percentage points) in the budget share allocated to red meat leads to a substantial rise in emissions. Figure 5 disaggregates these changes by quartile, revealing a consistent increase in emissions across all quartiles. Notably, the smallest increase occurs in the top-emitting quartile, which exhibits greater price inelasticity regarding red meat consumption.

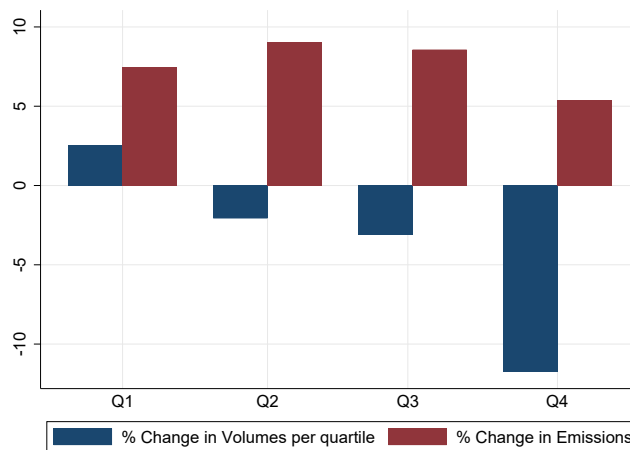
These findings, particularly the significant reduction in fresh fruits and vegetables and the increased consumption of red meat in this initial simulation, highlight the critical need for public policies that consider both price differentials and the varied responses of households. Such policies are essential for promoting sustainable diets and mitigating environmental impacts.

Figure 4: Average change in prices, budget shares and emissions by product category.



Notes: Results for the household full sample after the simulation of an inflation shock.

Figure 5: Predicted volumes and emissions changes by emission quartile.



Notes: Results per emission quartile after the simulation of an inflation shock.

5 Discussion

This study contributes to the debate on households' food consumption, which explains a substantial share of their total carbon footprint. To explore variations in carbon footprints, I build a unique database by merging two datasets. I combine purchase data on France for the period 2017-2019 and environmental. This process, involving machine learning algorithms, is made at the product level, enabling a nuanced examination of emission intensity diversity. Notably, significant discrepancies in emissions are driven by animal product consumption and total purchased volumes. Red meat emerges as the most environmentally taxing food category, emitting four times more than other meats. Despite this, all diets contribute to emissions, with

variations in purchase volumes playing a role.

To address the variance in emissions distribution, I analyze consumption patterns using a demand system across diverse food categories. Specifically, I estimate the quadratic version of Deaton and Muellbauer (1980)'s Almost Ideal Demand System introduced by Banks et al. (1997), separately per emission quartiles. This analysis exposes noteworthy disparities in how households respond across quartiles. Particularly, I identify that households with higher emissions are more responsive to changes in red meat prices than the rest of the population. I then run simulations to explore consumption shifts in the event of a purchasing power shock and its effect on food-related emissions. In particular, I use the food inflation observed between the first quarters of 2022 and 2023, with an average 13% increase in food prices. The simulation, grounded in assumptions of fixed price sensitivity and unchanged nominal food expenses, reveals a 5% reduction in consumption volumes and a 7% increase in emissions, primarily due to a shift from fruits and vegetables to more emission-intensive foods like meat. These findings highlight the necessity for public policies that consider both price differentials and the diverse responses of households to foster sustainable diets.

References

- Accardo, Jérôme and Sylvain Billot**, “Plus d’épargne chez les plus aisés, plus de dépenses contraintes chez les plus modestes,” Technical Report 2020.
- Amano-Patiño, Noriko**, “The Role of Prices, Income, and Preferences,” Technical Report 2019.
- Banks, James, Richard Blundell, and Arthur Lewbel**, “Quadratic Engel curves and consumer demand,” *Review of Economics and statistics*, 1997, 79 (4), 527–539.
- Barbier, Carine, Christian Couturier, Prabodh Pourouchottamin, Jean-Michel Cayla, M Sylvestre, and Ivan Pharabod**, “Energy and carbon footprint of food in France,” *IDDRI, Club Ingénierie Prospective Energie et Environnement: Paris, France*, 2019.
- , –, –, –, –, **Marie Silvestre, and Ivan Pharabod**, “L’empreinte énergétique et carbone de l’alimentation en France,” *Club Ingénierie Prospective Energie et Environnement, Paris, IDDRI, 24p*, 2019.
- Biermann, Gesa and Henrike Rau**, “The meaning of meat:(Un) sustainable eating practices at home and out of home,” *Appetite*, 2020, 153, 104730.
- Bonnet, Céline, Zohra Bouamra-Mechemache, and Tifenn Corre**, “An environmental tax towards more sustainable food: empirical evidence of the consumption of animal products in France,” *Ecological economics*, 2018, 147, 48–61.
- Büchs, Milena and Sylke V Schnepf**, “Who emits most? Associations between socio-economic factors and UK households’ home energy, transport, indirect and total CO2 emissions,” *Ecological Economics*, 2013, 90, 114–123.
- Caillavet, France, Adélaïde Fadhuile, and Véronique Nichèle**, “Taxing animal-based foods for sustainability: environmental, nutritional and social perspectives in France,” *European Review of Agricultural Economics*, 2016, 43 (4), 537–560.
- Chancel, Lucas**, “Global carbon inequality over 1990–2019,” *Nature Sustainability*, 2022, 5 (11), 931–938.
- Chanut, Nicolas**, “The heterogeneous carbon content of food consumption and its implication for policy,” Technical Report 2021.
- Clark, Michael, Marco Springmann, Mike Rayner, Peter Scarborough, Jason Hill, David Tilman, Jennie I Macdiarmid, Jessica Fanzo, Lauren Bandy, and Richard A Harrington**, “Estimating the environmental impacts of 57,000 food products,” *Proceedings of the National Academy of Sciences*, 2022, 119 (33), e2120584119.
- Crippa, Monica, E Solazzo, D Guizzardi, F Monforti-Ferrario, FN Tubiello, and AJNF Leip**, “Food systems are responsible for a third of global anthropogenic GHG emissions,” *Nature Food*, 2021, 2 (3), 198–209.
- Deaton, Angus and John Muellbauer**, “An Almost Ideal Demand System,” *American Economic Review*, 1980, 70 (3), 312–326.
- Funke, Franziska, Linus Mattauch, Inge van den Bijgaart, H Charles J Godfray, Cameron Hepburn, David Klenert, Marco Springmann, and Nicolas Treich**, “Toward Optimal Meat Pricing: Is It Time to Tax Meat Consumption?,” *Review of Environmental Economics and Policy*, 2022, 16 (2), 219–240.

- Gicheva, Dora, Justine Hastings, and Sofia Villas-Boas**, “Investigating Income Effects in Scanner Data: Do Gasoline Prices Affect Grocery Purchases?,” *American Economic Review Papers & Proceedings*, 2010, *100* (2), 480–84.
- Griffith, Rachel, Martin O’Connell, and Kate Smith**, “Shopping around: how households adjusted food spending over the great recession,” *Economica*, 2016, *83* (330), 247–280.
- Koch, Franziska, Thorsten Heuer, Carolin Krems, and Erika Claupein**, “Meat consumers and non-meat consumers in Germany: A characterisation based on results of the German National Nutrition Survey II,” *Journal of Nutritional science*, 2019, *8*, e21.
- Laursen, Ulla Bach, Martin Berg Johansen, Albert Marni Joensen, Cathrine Juel Lau, Kim Overvad, and Mogens Lytken Larsen**, “Educational level and living arrangements are associated with dietary intake of red meat and fruit/vegetables: A Danish cross-sectional study,” *Scandinavian Journal of Public Health*, 2019, *47* (5), 557–564.
- Lecocq, Sébastien and Jean-Marc Robin**, “Estimating almost-ideal demand systems with endogenous regressors,” *The Stata Journal*, 2015, *15* (2), 554–573.
- Mouzon, Olivier De and Valérie Orozco**, “Nutrixconso: recherche et appariement de données d’achats et de données nutritionnelles,” *Cahier des Techniques de l’INRA*, (74), 2011, pp. 42–83.
- Nevo, Aviv**, “Empirical Models of Consumer Behavior,” *Annual Review of Economics*, 2011, *3* (1), 51–75.
- **and Arlene Wong**, “The elasticity of substitution between time and market goods: Evidence from the Great Recession,” *International Economic Review*, 2019, *60* (1), 25–51.
- O’Connell, Martin, Kate Smith, and Rebekah Stroud**, “The dietary impact of the COVID-19 pandemic,” *Journal of Health Economics*, 2022, *84*, 102641.
- Poore, Joseph and Thomas Nemecek**, “Reducing food’s environmental impacts through producers and consumers,” *Science*, 2018, *360* (6392), 987–992.
- Reynolds, Christian J, Graham W Horgan, Stephen Whybrow, and Jennie I Macdiarmid**, “Healthy and sustainable diets that meet greenhouse gas emission reduction targets and are affordable for different income groups in the UK,” *Public Health Nutrition*, 2019, *22* (8), 1503–1517.
- Starr, Jared, Craig Nicolson, Michael Ash, Ezra M Markowitz, and Daniel Moran**, “Assessing US consumers’ carbon footprints reveals outsized impact of the top 1%,” *Ecological Economics*, 2023, *205*, 107698.
- Zhen, Chen, Eric A Finkelstein, James M Nonnemaker, Shawn A Karns, and Jessica E Todd**, “Predicting the effects of sugar-sweetened beverage taxes on food and beverage demand in a large demand system,” *American Journal of Agricultural Economics*, 2014, *96* (1), 1–25.

Appendix

A1 Data Appendix

A1.1 Aggregation of Food Products into Product Categories

Food products are grouped into 13 categories that reflect differences in both the nutritional and carbon content of food. The 13 categories are defined as follows.

1. *Starchy food*: pasta, bread, semolina, cook-type cereals, potatoes.
2. *Fresh fruits and vegetables*: fresh fruits, fresh vegetables.
3. *Red meat*: beef, mutton, lamb.
4. *White meat*: chicken, turkey, veal, rabbit, pork, poultry.
5. *Other meat*: cold cuts, mix of meat, ham, game meat, frogs.
6. *Fish*: fish, seafood, fish eggs, surimi, mix of fish.
7. *Dairy*: goat cheese, cow cheese, sheep cheese, melted cheese, milk, yogurts.
8. *Snacks*: sugary biscuits, jam, honey, chocolate spread, cereal/granola bars, chocolate, pastries, breakfast drink preparation, breakfast cereals, ice cream, deserts, dry fruits and seeds, salty biscuits, olives, crisps.
9. *Ready meals*: prepared dishes, frozen dishes, canned dishes.
10. *Soft drinks*: sodas, syrups, juices.
11. *Other beverages*: coffee, tea, infusions, water, chicory.
12. *Seasonings*: spices, oil, vinegar, butter, croutons, breadcrumbs, raw pastry, confectionery flavourings, flour, cornflour, prepared crust, coconut milk, cream, lemon juice, fresh herbs, dry herbs, garlic, onion.
13. *Non-fresh fruits and vegetables*: canned fruits, canned vegetables, frozen fruits, frozen vegetables, lyophilised vegetables, beans, dried vegetables, packaged vegetables.

An additional category composed of very heterogeneous products, such as baby foods and food supplements, has been excluded from the analysis. The 180 subgroups mentioned in Table A4 are not listed. They roughly correspond to the components of the aforementioned categories.

Variance Decomposition

In my analysis of carbon intensity per product, weighted by the volumes purchased, I observe that the variance component between product categories is larger than that within product categories. This finding, which is consistent even at higher levels of product aggregation, is presented in Table A4.²⁴ This result contrasts with the findings of Chanut (2021), who reports that the within-category variance often exceeds the between-category variance. The key difference stems from the fact that the aforementioned study groups meat, eggs, and fish into a single, broad category. However, such aggregation combines items with significantly different emissions intensities, like red and white meat. This highlights the necessity of forming well-defined product categories that reflect the emission variations across different types of products.

The notable variance observed between product categories justifies conducting the analysis at a higher level of product aggregation. By doing so, the model estimation is facilitated, as it minimizes the number of cross-elasticities that need to be considered. This approach not only simplifies the analytical process but also ensures that the variations in emissions across different product types are adequately captured and reflected in the results.

Table A4: Decomposition of the variance of emission intensity.

Share of variance for different aggregation levels	Within	Between
Product subgroup (N=180)	17.3	82.7
Product category (N=13)	43.2	56.8

Notes: This table presents the decomposition of carbon intensity variation between group variation and within-group variation for two levels of product aggregation.

A1.2 French Geography

- **Departements** correspond to territorial collectivities. They have an administrative status and are smaller than regions. There are 97 *departements* in Metropolitan France. The analysis excludes the two *departements* of Corsica because of differences in food prices and consumption patterns.
- **Living zones** correspond to the smallest area where residents have access to the most common facilities and services. There are 1666 living zones in metropolitan France and 1320 in my estimation sample. Living zones do not correspond to administrative units. Hence, they can spread over several *departements*.

²⁴The 180 product subgroups are defined in the KWP data, the 13 product categories are defined in Appendix A1.1.

A1.3 Matching Procedure

The KWP data allows a thorough study of the evolution of eating habits, encompassing over ten million purchases annually. However, the frequent introduction and discontinuation of products pose a challenge in finely matching this database with others, such as environmental databases. Despite this, the product-defining characteristics, like fat content or type of meat, remain relatively consistent. The goal of this matching procedure is to align purchase data with environmental data at the product level.

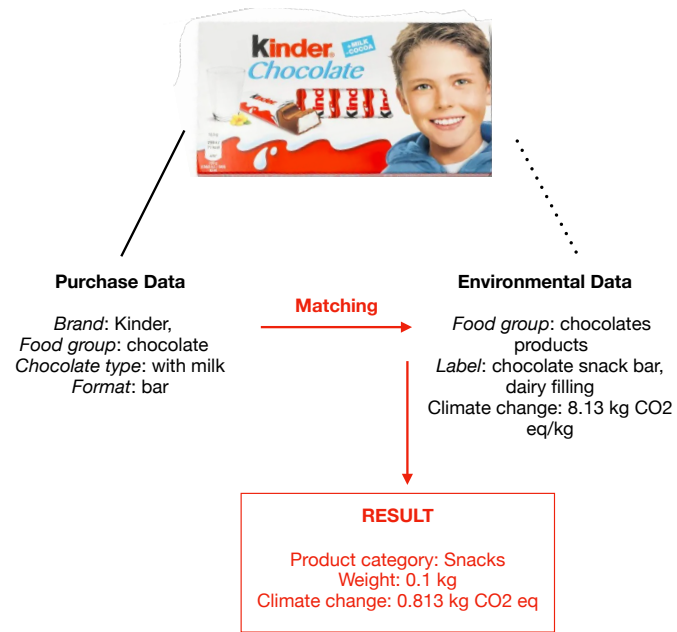
In practice, I use existing correspondences between the two databases and employ a machine learning algorithm to predict the most accurate product matches. In the KWP database, each purchase is identified by a unique product identifier, defined by various attributes such as batch size, type of animal for meats, and flavor for yogurts. The Agribalyse 2017 database, which we aim to match with KWP data, lists standardized food products without specifying brands or retailers, and quantities are expressed in kilograms. An example of the matching outcome for a snack product is illustrated in Figure 6.

Leveraging the fact that KWP and Agribalyse databases were previously matched for the year 2010,²⁵ I use a supervised classification algorithm to match every KWP purchase with a standardized Agribalyse product, thereby retrieving environmental impact data. A limitation of this method is its reliance on the initial correspondence table, which matches KWP with an older version of Agribalyse (version 2013) containing 1500 products, compared to the latest version's 2500 products. Consequently, the algorithm training is confined to these 1500 matched products, excluding newer Agribalyse products from the analysis. While supervised algorithms cannot overcome this limitation, the impact on information loss is expected to be minimal. This is because the product type is the most significant factor in determining a food's carbon content. Additionally, 'new' products are infrequent, and many are simply variations of existing products, such as yogurts with different berries, thus minimizing information loss.

The Random Forest algorithm is initially tested on a dataset containing all previously matched products, divided into a training set (70%) and a testing set (30%). This algorithm integrates the outputs of multiple decision trees to predict a single outcome. Each tree begins with the product label and makes a series of yes/no decisions based on one or two features. With 1000 decision trees and 10 random samplings, the algorithm achieves an accuracy score of 0.96. Out of 1500 Agribalyse products, only 13 have an F1-score below 0.6.

²⁵For more detail, see De Mouzon and Orozco, 2011.

Figure 6: Matching example in the snacks category.

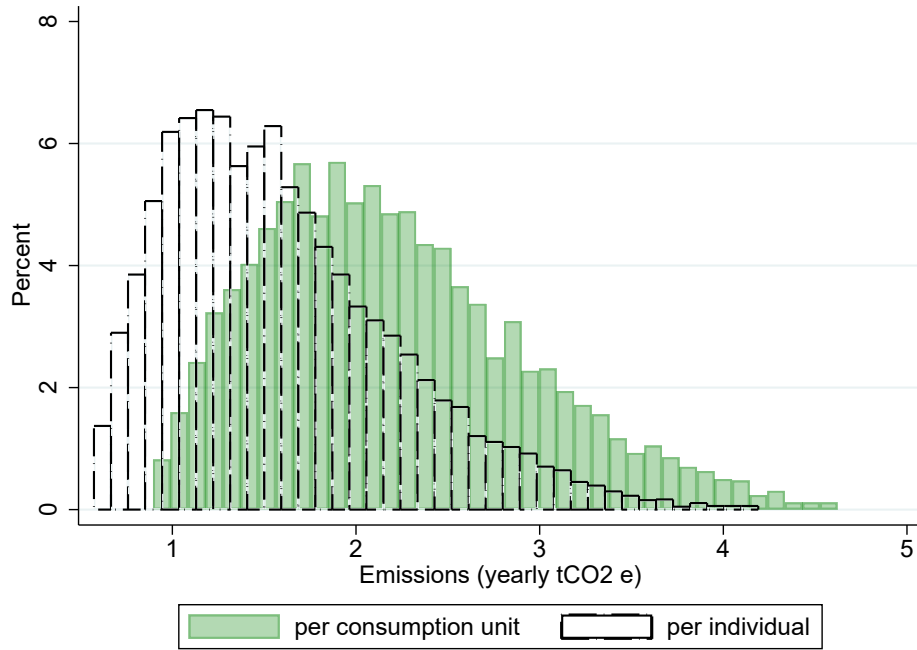


Notes: Simple example taken from the real-life setting.
The variable and product names are not extracts from the KWP dataset.

A2 Additional stylized facts

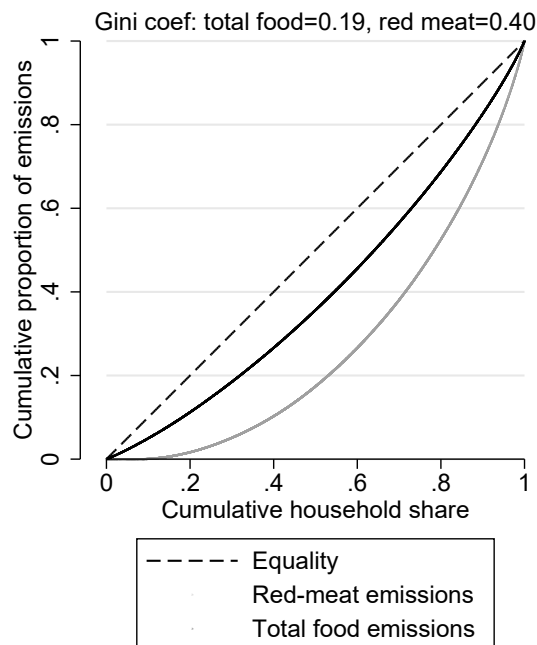
A2.1 Distribution of emissions

Figure 7: Distribution of yearly emissions per household.



Source: Average yearly emissions per household divided by consumption unit and the number of household members. Emissions are trimmed at bottom and top 1 percent. Own calculations based on KWP 2017-2019 and Agribalyse data.

Figure 8: Lorenz curves: total emissions and red meat emissions.

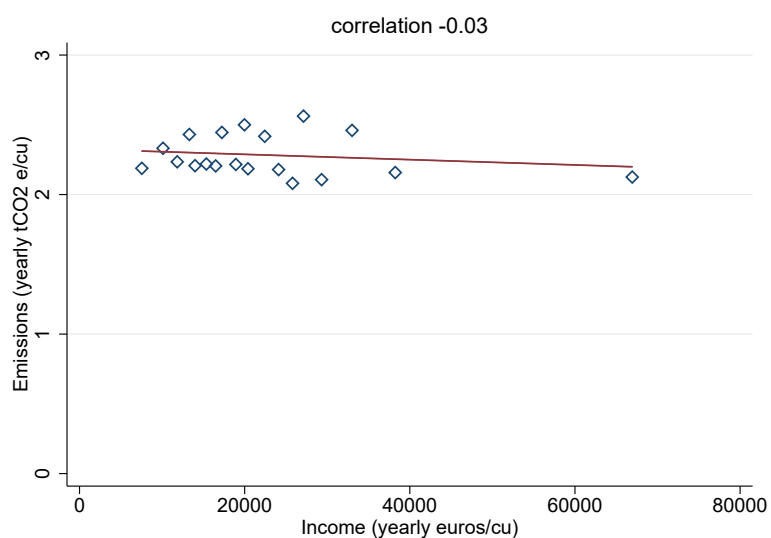


Distribution of food related emissions.

Source: Own calculations based on KWP 2017-2019 and Agribalyse data.

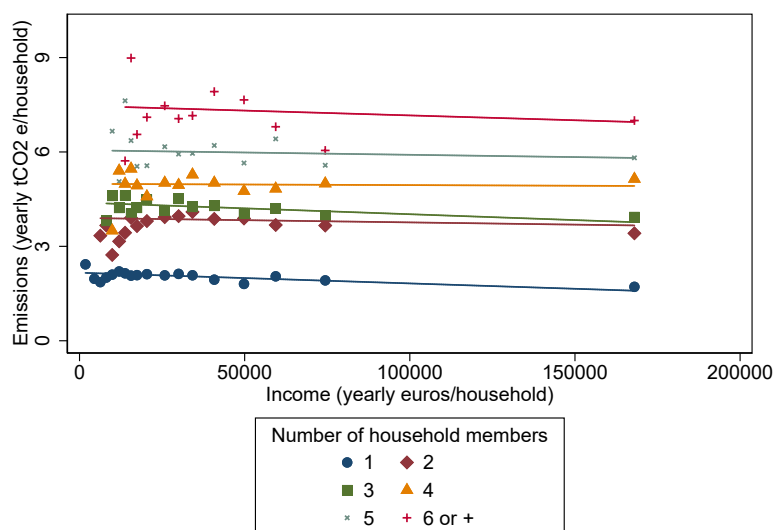
A2.2 Households' income and food carbon footprints

Figure 9: Income and food emissions weighted by consumption units.



Binscatter of income and food emissions (20 equally sized bins).
Source: Own calculations based on KWP 2017-2019 and Agribalyse data.

Figure 10: Income and food emissions for different households' sizes.



Binscatter of income and food emissions (20 equally sized bins) per household size.
Source: Own calculations based on KWP 2017-2019 and Agribalyse data.

A2.3 Distribution of emissions during Covid-19 Lockdown

Table A5: Households' characteristics by emission quartile (averages over April and November 2020).

	Quartile 1		Quartile 2		Quartile 3		Quartile 4	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Income/CU	1806.13	954.93	1882.61	1237.85	1863.20	1096.23	1862.78	1078.70
Age	51.05	19.20	54.92	17.01	55.83	15.08	56.40	14.38
Size of the HH	1.94	1.19	2.02	1.18	2.24	1.21	2.35	1.18
Rural (0/1)	0.14	0.35	0.17	0.38	0.20	0.40	0.19	0.39
Car owner (0/1)	0.84	0.37	0.86	0.34	0.91	0.29	0.91	0.29
House owner (0/1)	0.59	0.49	0.65	0.48	0.68	0.47	0.67	0.47
Education	2.73	0.92	2.62	0.89	2.53	0.87	2.46	0.84

Notes: Distribution of key socio-demographic variables by monthly average emissions quartiles. *Income* is expressed in euros. *Income/CU* stands for income divided by the number of consumption units in the household, defined using the OECD modified scale. *Food Spending/CU* is the total monthly food expenditures in euros (deflated) divided by the number of consumption units in the household. *Age* indicates the age of the household age. *Education* is a categorical variable indicating the educational background of the household head. It takes values between 0 and 4, where 0 stands for no education, 1 stands for having a primary degree, 2 for a secondary degree, 3 for a bachelor and 4 for a master's degree or above.

A3 Censoring

Table A6: Balance table of descriptive statistics across full and positive samples.

Variable	(1) Full sample Mean/SE	(2) Positive sample Mean/SE	T-test Difference (1)-(2)
Income	2704.819 (25.774)	2818.297 (25.796)	-113.478***
Income/CU	1829.786 (17.187)	1788.915 (16.399)	40.872***
Age	52.372 (0.265)	52.534 (0.261)	-0.162
Household size	2.189 (0.016)	2.415 (0.017)	-0.227***
Rural (0/1)	0.178 (0.005)	0.189 (0.006)	-0.011***
Pesco-vegetarian (0/1)	0.008 -	0.000 -	0.008***
Vegetarian (0/1)	0.003 -	- -	0.003***
N	70543	51036	
Clusters	8641	8067	
F-test of joint significance (F-stat)			1458.587***
F-test, number of observations			121579

Notes: The value displayed for t-tests are the differences in the means across the groups. *, **, ***, indicate significance at the 1%, 5%, and 10% levels, respectively. Standard errors are presented in parentheses. The value displayed for F-tests is the F-statistic.

A4 Additional results.

A4.0.1 Results by ex-ante heterogeneity

Table A7: Demand estimates top and bottom emission quartiles, as defined by the first period of appearance in the data (first quarter of 2017).

	Bottom emitting quartile			Top emitting quartile		
	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$
Red meat	0.040*** (0.002)	0.554** (0.192)	-1.093*** (0.034)	0.070*** (0.001)	0.973*** (0.107)	-1.441*** (0.029)
White meat	0.059*** (0.002)	0.705*** (0.144)	-0.693*** (0.040)	0.092*** (0.002)	0.769*** (0.084)	-0.915*** (0.023)
Fish	0.067*** (0.002)	1.236*** (0.121)	-0.820*** (0.023)	0.055*** (0.001)	2.193*** (0.164)	-0.917*** (0.034)
Dairy	0.119*** (0.002)	1.184*** (0.070)	-0.851*** (0.017)	0.106*** (0.001)	1.013*** (0.064)	-1.025*** (0.017)
Fresh fruits & veg	0.141*** (0.003)	1.679*** (0.081)	-0.966*** (0.050)	0.097*** (0.002)	1.844*** (0.119)	-1.298*** (0.046)

Notes: Results after estimating a quadratic AIDS model, per household emission quartile. Uncompensated price elasticities.

A4.0.2 Estimation for all the quartiles

Table A8: Demand estimates by emission quartiles.

	Q1			Q2			Q3			Q4		
	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$	$\hat{\sigma}$	$\hat{\eta}_E$	$\hat{\eta}_P$
Red meat	0.038	0.762	-1.068	0.047	0.772	-1.175	0.058	1.121	-1.274	0.072	0.945	-1.409
White meat	0.063	1.125	-0.722	0.071	1.173	-0.762	0.076	1.046	-0.826	0.093	0.778	-1.001
Fish	0.072	1.380	-0.854	0.068	1.717	-0.824	0.062	1.809	-0.845	0.053	1.952	-0.946
Dairy	0.110	0.932	-0.853	0.111	0.834	-0.923	0.112	0.846	-0.957	0.106	1.077	-1.073
Fresh fruits & veg	0.140	1.541	-0.961	0.121	1.668	-1.040	0.105	2.094	-1.007	0.094	1.601	-1.408

Notes: Results after estimating a quadratic AIDS model, per household emission quartile. Uncompensated price elasticities.

A4.0.3 Full tables estimation

A5 Comparisons with estimates from the literature.

Table A9: Uncompensated cross-price elasticities, households in the **first quartile** of emissions.

Category	Starches b/se	Fresh fruits/veg b/se	Red meat b/se	White meat b/se	Other meat b/se	Fish b/se	Dairy b/se	Snacks b/se	Ready meals b/se	Soft drinks b/se	Other drinks b/se	Seasonings b/se	Canned fruits/veg b/se
Starches	-0.680*** (0.022)	0.016 (0.031)	0.044* (0.020)	-0.100*** (0.024)	-0.102*** (0.025)	-0.011 (0.017)	-0.108** (0.036)	-0.194*** (0.056)	-0.076** (0.029)	-0.078** (0.030)	-0.024 (0.028)	0.013 (0.011)	-0.049** (0.017)
Fresh fruits/veg	-0.039* (0.018)	-0.961*** (0.029)	-0.009 (0.019)	-0.046* (0.022)	-0.138*** (0.022)	0.046** (0.017)	-0.039 (0.032)	-0.272*** (0.046)	-0.130*** (0.025)	-0.071** (0.026)	-0.027 (0.024)	0.008 (0.012)	-0.032* (0.015)
Red meat	-0.021 (0.030)	0.008 (0.040)	-1.068*** (0.024)	0.192*** (0.034)	0.064 (0.036)	-0.013 (0.021)	0.105* (0.052)	-0.046 (0.082)	0.017 (0.040)	0.163*** (0.042)	0.018 (0.040)	-0.012 (0.013)	0.068** (0.024)
White meat	-0.089*** (0.022)	0.046 (0.033)	-0.117*** (0.025)	-0.722*** (0.029)	-0.049 (0.028)	0.021 (0.018)	-0.174*** (0.039)	-0.356*** (0.060)	-0.124*** (0.030)	0.032 (0.032)	-0.021 (0.030)	0.004 (0.012)	-0.029 (0.019)
Other meat	-0.049** (0.017)	-0.012 (0.020)	0.023 (0.016)	0.032 (0.018)	-0.606*** (0.021)	-0.026** (0.009)	-0.009 (0.029)	-0.102* (0.045)	-0.020 (0.020)	-0.014 (0.024)	-0.002 (0.022)	-0.009 (0.006)	0.007 (0.013)
Fish	-0.044* (0.020)	0.022 (0.025)	-0.013 (0.020)	-0.016 (0.022)	-0.037 (0.024)	-0.854*** (0.011)	0.011 (0.035)	-0.290*** (0.055)	-0.042 (0.024)	0.019 (0.029)	-0.029 (0.027)	0.003 (0.007)	0.017 (0.016)
Dairy	0.003 (0.015)	0.028 (0.017)	0.001 (0.014)	-0.025 (0.016)	0.018 (0.017)	-0.018* (0.008)	-0.855*** (0.026)	-0.101* (0.039)	-0.040* (0.017)	-0.048* (0.021)	0.065** (0.020)	-0.014** (0.005)	-0.038** (0.012)
Snacks	0.031* (0.013)	-0.008 (0.017)	-0.005 (0.013)	-0.035* (0.015)	-0.015 (0.016)	-0.053*** (0.009)	-0.074** (0.023)	-0.403*** (0.035)	0.046** (0.017)	0.013 (0.018)	-0.071*** (0.017)	-0.012* (0.006)	-0.037*** (0.011)
Ready meals	-0.024 (0.027)	-0.042 (0.033)	0.046 (0.025)	-0.022 (0.030)	0.008 (0.031)	-0.019 (0.015)	0.046 (0.047)	-0.035 (0.074)	-0.665*** (0.033)	0.020 (0.039)	-0.271*** (0.036)	-0.013 (0.009)	0.053* (0.022)
Soft drinks	0.041 (0.036)	-0.087 (0.057)	0.023 (0.037)	0.065 (0.043)	0.096* (0.046)	-0.050 (0.030)	0.197** (0.065)	0.319** (0.101)	0.178*** (0.052)	-0.968*** (0.041)	0.082 (0.049)	-0.056** (0.020)	0.043 (0.030)
Other drinks	-0.020 (0.015)	-0.033 (0.018)	0.019 (0.014)	-0.069*** (0.016)	-0.074*** (0.017)	0.009 (0.009)	-0.006 (0.026)	-0.104** (0.040)	-0.079*** (0.017)	-0.021 (0.021)	-0.573*** (0.023)	0.000 (0.005)	-0.011 (0.012)
Seasonings	0.032 (0.021)	-0.033 (0.032)	0.053* (0.022)	-0.016 (0.024)	-0.033 (0.026)	-0.019 (0.016)	0.070 (0.037)	0.037 (0.057)	0.007 (0.030)	0.085** (0.031)	0.007 (0.028)	-0.882*** (0.016)	0.002 (0.018)
Canned fruits/veg	0.024 (0.026)	-0.027 (0.032)	0.020 (0.025)	-0.060* (0.028)	-0.143*** (0.031)	0.009 (0.016)	-0.006 (0.045)	-0.044 (0.070)	0.059 (0.031)	0.060 (0.037)	0.035 (0.035)	-0.007 (0.010)	-0.692*** (0.023)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Demand estimation for the lowest quartile of emissions.

Table A10: Uncompensated cross-price elasticities, households in the fourth quartile of emissions.

Category	Starches b/se	Fresh fruits/veg b/se	Red meat b/se	White meat b/se	Other meat b/se	Fish b/se	Dairy b/se	Snacks b/se	Ready meals b/se	Soft drinks b/se	Other drinks b/se	Seasonings b/se	Canned fruits/veg b/se
Starches	-0.808*** (0.013)	-0.032 (0.025)	-0.005 (0.015)	-0.026 (0.018)	-0.034* (0.015)	0.017 (0.015)	-0.128*** (0.018)	0.033 (0.028)	-0.040 (0.029)	-0.034** (0.012)	-0.006 (0.017)	0.013* (0.006)	0.014 (0.012)
Fresh fruits/veg	-0.019 (0.017)	-1.408*** (0.040)	0.155*** (0.027)	0.128*** (0.029)	-0.051* (0.024)	-0.015 (0.023)	0.184*** (0.029)	-0.044 (0.042)	0.080 (0.044)	0.090*** (0.019)	0.146*** (0.023)	-0.037*** (0.011)	0.010 (0.018)
Red meat	-0.025 (0.017)	0.181*** (0.039)	-1.409*** (0.026)	0.050 (0.028)	0.041 (0.021)	0.079*** (0.023)	0.103*** (0.026)	-0.218*** (0.042)	-0.126** (0.042)	0.077*** (0.017)	-0.062* (0.024)	0.009 (0.011)	0.005 (0.018)
White meat	0.006 (0.014)	-0.006 (0.028)	-0.075*** (0.017)	-1.001*** (0.022)	0.048** (0.016)	-0.002 (0.017)	0.038 (0.020)	-0.062 (0.032)	-0.066* (0.033)	0.020 (0.013)	0.032 (0.019)	0.005 (0.007)	-0.029* (0.014)
Other meat	-0.002 (0.012)	0.019 (0.024)	-0.015 (0.015)	-0.014 (0.018)	-0.831*** (0.013)	-0.007 (0.014)	-0.008 (0.017)	-0.071** (0.027)	0.035 (0.029)	-0.023* (0.011)	0.009 (0.016)	-0.017** (0.006)	-0.005 (0.012)
Fish	-0.042 (0.023)	0.078 (0.048)	0.056 (0.029)	-0.049 (0.035)	0.005 (0.027)	-0.946*** (0.026)	0.213*** (0.036)	-0.297*** (0.056)	-0.119* (0.056)	0.058** (0.022)	0.051 (0.031)	-0.021 (0.013)	-0.033 (0.024)
Dairy	0.039*** (0.010)	0.032 (0.022)	0.070*** (0.014)	0.011 (0.016)	-0.015 (0.012)	-0.021 (0.013)	-1.073*** (0.015)	-0.026 (0.024)	0.012 (0.026)	0.021* (0.010)	0.028 (0.014)	-0.006 (0.006)	0.003 (0.011)
Snacks	-0.013 (0.009)	0.063*** (0.018)	0.041*** (0.011)	-0.027* (0.013)	-0.070*** (0.011)	-0.016 (0.011)	-0.102*** (0.013)	-0.655*** (0.020)	-0.002 (0.021)	-0.052*** (0.008)	-0.099*** (0.012)	-0.000 (0.005)	-0.013 (0.009)
Ready meals	-0.077*** (0.017)	0.032 (0.035)	-0.015 (0.020)	-0.063* (0.025)	0.048* (0.021)	-0.000 (0.021)	-0.026 (0.026)	-0.066 (0.039)	-0.921*** (0.046)	-0.044** (0.017)	-0.178*** (0.024)	0.023* (0.009)	0.064*** (0.017)
Soft drinks	-0.006 (0.019)	0.086* (0.041)	0.093*** (0.026)	0.028 (0.031)	-0.036 (0.023)	-0.020 (0.024)	-0.051 (0.029)	0.095* (0.044)	0.121* (0.047)	-1.161*** (0.019)	-0.049 (0.027)	-0.046*** (0.011)	0.002 (0.020)
Other drinks	0.032** (0.010)	0.004 (0.020)	-0.031* (0.012)	-0.025 (0.015)	-0.017 (0.011)	0.013 (0.012)	-0.036* (0.015)	-0.039 (0.023)	-0.043 (0.024)	-0.020* (0.009)	-0.767*** (0.013)	0.005 (0.005)	-0.036*** (0.010)
Seasonings	-0.014 (0.020)	0.104* (0.042)	0.128*** (0.026)	-0.055 (0.030)	-0.147*** (0.025)	-0.015 (0.024)	0.004 (0.029)	-0.131** (0.046)	-0.015 (0.048)	0.048* (0.019)	-0.038 (0.028)	-0.875*** (0.010)	-0.042* (0.020)
Canned fruits/veg	0.044** (0.014)	-0.117*** (0.029)	0.019 (0.019)	0.028 (0.022)	-0.004 (0.017)	-0.037* (0.018)	-0.047* (0.022)	0.018 (0.033)	0.093** (0.035)	0.035* (0.014)	0.018 (0.020)	0.004 (0.008)	-0.884*** (0.014)

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Demand estimation for the households in the highest emission quartile.

Table A11: Estimates from the literature (product space models).

	Amano-Patiño (2019)			Caillavet et al. (2016)		
	$\hat{\eta}_E$	$\hat{\eta}_{CP}$	$\hat{\eta}_{UP}$	$\hat{\eta}_E$	$\hat{\eta}_{CP}$	$\hat{\eta}_{UP}$
Processed meat	0.5413	-1.2145	-1.2161			
Fresh meat and seafood	0.3741	-1.1864	-1.1874			
Fruit and Vegetables	0.6012	-0.9787	-0.9810			
Dairy	0.3572	-0.6985	-0.7012	0.952	-1.347	-
Beef				1.002	-1.113	-
Country	US			France		

Notes: $\hat{\eta}_E$ the estimated expenditure elasticity, $\hat{\eta}_{UP}$ (resp. $\hat{\eta}_{CP}$) the estimated uncompensated price elasticity (resp. compensated price elasticity).