National Accounts in a World of Naturally Occurring Data: A Proof of Concept for Consumption*

Gergely Buda
Barcelona School of Economics

Vasco Carvalho
University of Cambridge

Stephen Hansen
Imperial College and CEPR

Álvaro Ortiz
BBVA Research

Tomasa Rodrigo
BBVA Research

José V. Rodríguez Mora
University of Edinburgh

March 24, 2022

Abstract

This paper provides a first proof of concept that naturally occurring transaction data, arising through the decentralized activity of millions of economic agents, can be structured, organized and harnessed to produce national account-like objects. In particular, we deploy transaction-level data from Banco Bilbao Vizcaya Argentaria (BBVA), one of the largest banks in the world, to show (i) how one can reproduce current official statistics on aggregate consumption in the national accounts with a high degree of precision and, as a result of the richness of transaction data, (ii) produce novel, highly detailed distributional accounts for consumption.

*We thank seminar participants at the Bank of Canada, CEPR Workshop on New Consumption Data, and OECD. Carvalho gratefully acknowledges funding from the Leverhulme Trust and ERC Consolidator Grant 101001221. Hansen gratefully acknowledges financial support from ERC Consolidator Grant 864863.

1 Introduction

The workings of modern payment systems and financial institutions generate a complete ledger of everyday transactions. Every purchase, every debit, every transfer leaves behind a digital footprint, which is recorded in this ledger. This large, naturally occurring and unstructured transaction-level data, together with its associated rich meta-data, is increasingly available to researchers and holds the promise of reshaping economic measurement. Indeed, this promise has not gone unnoticed by academics, national statistics agencies and policy-makers alike, who all reaffirm that unstructured transaction data will necessarily play an increasingly prominent role in 21st century national accounting (see e.g. (Bean 2016)). And yet, despite recent advances – most noticeable in the profusion of real-time indicators that have surfaced during the recent COVID-19 crisis – national statistical agencies still rely on more traditional structured survey data and slow-moving censuses when compiling official national accounts.

Against this background, this paper provides a first proof of concept that, indeed, naturally occurring transaction data, arising through the decentralized activity of millions of economic agents, can be structured, organized and harnessed to produce national account-like objects. In particular, we deploy transaction-level data from Banco Bilbao Vizcaya Argentaria (BBVA), one of the largest banks in the world, to show (i) how one can reproduce current official statistics on aggregate consumption in the national accounts with a high degree of precision and, as a result of the richness of transaction data, (ii) produce novel, highly detailed distributional accounts for consumption. This is therefore a pure measurement paper demonstrating that such transaction data, when suitably organized via national accounting principles, can subsume current national accounts consumption methodologies and outputs and not simply, as already acknowledged in practice, serve as a source for useful coincident indicators or proxies.

Our data covers the universe of BBVA retail accounts in Spain by BBVA and yields an unprecedented granular ledger, allowing us to track expenditure as it flows out of these accounts, transaction by transaction, for a total of 3 billion individual transactions by 1.8 million BBVA customers, from 2016 to 2021. Based on this data, the current paper makes three contributions. First, we show how to construct a large, representative and highly detailed panel of household expenditure. Note that these transactions cover all BBVA households' debit and credit card transactions (both online and offline), all direct recurrent debits, all one-off transfers and individual payments as well as all cash withdrawals, which we assume are spent as consumption. Leveraging this comprehensive data, we further detail how to exploit meta-data associated with each account holder, transaction and means of payment to: (i) categorize transactions across harmonized consumption spending categories, (ii) filter out non-consumption expenditures (such as transfers to saving accounts, household-to-household transfers or tax payments), (iii) impute the consumption of housing services for all households, by exploiting information on actual rental, housing utilities, location and income for a subsample of BBVA households, and finally, (iv) construct a large sampling frame of households that is representative along demographic observables – in particular, gender, age and province cells – so as to mimic the characteristics of the Spanish adult population.

Second, leveraging this naturally occurring, large scale consumption survey, we show how to construct - from the bottom-up - a series for quarterly aggregate final consumption expenditures of domestic households and compare it against that in the Spanish Quarterly National Accounts, as compiled by Spain's National Statistics Institute (INE) under *"Gasto en Consumo Final de los Hogares"*. It is important to note that these two series follow different methodologies. Our series aggregates directly from a nationally representative large-scale real time expenditure survey, as described above. Instead, official quarterly national accounts consumption, are based on firm sales survey data, with subsequent imputations regarding who is consuming (e.g. distinguishing Spanish nationals vs. foreigners, households vs. firms) and what is being consumed (e.g. investment or intermediate goods vs. consumption by households). Despite these methodological differences, we show that our naturally occurring aggregate consumption matches the official INE series remarkably well, both in levels and in growth rates, thus providing a providing a first proof of concept that national accounts can feasibly rely on high-quality real time transaction level data rather than costly slow moving surveys.

Third, we show that naturally occurring transaction data, beyond providing an alternative way of constructing traditional national account objects, immediately opens the door to the possibility of substantially more detailed - and informative - national accounting. In particular, due to its sheer size, real-time availability and abundant meta-data, we show that it is feasible to construct a variety of consumption sub-aggregates, be it (i) across consumption categories, thus going beyond the traditional durables vs. non-durables distinction in national accounts, (ii) across space, thus complementing currently sparsely populated regional accounts and making it feasible to construct, for example, zip-code level aggregate consumption, (iii) across demographics, allowing us to construct, for example, aggregate consumption by age group or gender, (iv) across means of payment, rendering it possible to capture both low and high frequency changes in expenditure channels (e.g. online vs offline aggregate consumption). Importantly, unlike for traditional consumption surveys, transaction data is plentiful and timely enough that it is possible to conduct analysis across all these dimensions, simultaneously and compare, for example weekly offline consumption of a particular category of goods by gender and across zip codes. This, in turn, opens the possibility of constructing distributional national accounts with unprecedented detail, allowing for an in-depth analysis of inequality of consumption (both in levels and growth rates). We intend to pursue this application in future versions of this paper.

This paper relates to three distinct literatures. First, our work is related to a small literature reviewing currents methods and sources in the compilation of national accounts, their shortcomings, as well as possible solutions in light of new data sources and methods (Bean 2016, Jarmin 2019, National Academies of Sciences, Engineering, and Medicine 2018). Recurring themes in this literature relate to the increasing costs of maintaining national accounts, declining response rates to traditional survey based sources underpinning national accounts and the increasing complex needs of data users and the demand for accurate, timely and granular measurement. This literature also invariably—and forcefully—suggests the use of unstructured data as a possible solution to alleviate such problems and concerns. Relative to this literature, our paper provides a first proof of concept that this suggestion is feasible and delivers high quality results.

Second, our paper relates to a fast growing literature leveraging from access to credit/debit card and financial app data in order to generate high frequency expenditure series; see (Gelman et al. 2014, Baker 2018, Aladangady et al. 2021, Olafsson and Pagel 2018). Given the increasing availability of such data and in face of societal demands for high frequency, granular tracking of the economy during the COVID-19 pandemic, this literature expanded rapidly over the past two years; see, for some early contributions, (Carvalho et al. 2021, Andersen et al. 2020, Chetty et al. 2020) and (Baker and Kueng 2021, Vavra 2021), for recent reviews taking stock of this literature. Relative to this literature, our main contribution is threefold. First, relative to papers based on card data alone, we expand the scope of consumption expenditure significantly, by additionally considering direct debits, one off transfers and cash withdrawals, thus providing a complete view of consumption expenditures across different means of payment. Second, relative to the literature based on data originating in financial apps, our sample is much larger and is therefore arguably less encumbered by sample selection and size issues, allowing us to both construct nationally representative aggregates and to offer micro-distributed series. Third, our focus on building national account objects via aggregation of transaction data is novel. As discussed above, our intent is not to create a high quality real time proxy for consumption. Rather, differently from this literature, it is to offer evidence that national accounting can be based on such data.

Finally, as described above, our bottom-up approach relies on constructing a large-scale, highly detailed consumption survey. Thus, this paper is related to a third literature, analysing the methods, biases and shortcomings associated to traditional consumption surveys; see for example, (Aguiar and Hurst 2013, Aguiar and Bils 2015, Attanasio et al. 2014, Barrett et al. 2014, Coibion et al. 2021, Passero et al. 2014, Pistaferri 2015, Koijen et al. 2014, Kreiner et al. 2014). In particular, papers in this literature stress the difficulties in either (i) reconciling the aggregate consumption series implied by these surveys with official national accounts aggregate consumption or (ii) analysing consumption inequality based on such data given biases in response rates along unobservables, heterogeneity in both the levels and dynamics in the coverage of particular consumption categories or peculiarities induced by particular forms of sampling. Relative to this literature, we show that our large scale consumption survey, as assembled via naturally occurring transaction data, is largely immune to such biases and criticisms. In particular, our survey tracks national accounts aggregates well (both in levels and growth rates) and provides an arguably more complete and unbiased record of expenditures across all categories, at all frequencies and across various demographic characteristics. Future versions of this paper will build on this to provide a granular understanding of both the levels and growth rates of consumption inequality in Spain.

2 From Transactions to Consumption Spending

Our transaction data originates from Banco Bilbao Vizcaya Argentaria (BBVA), a Spanish multinational bank operating in dozens of countries and ranked in the top 50 globally by total assets. It is the second-largest bank in Spain by total assets and market capitalization. We access the universe of BBVA retail financial accounts in Spain to build consumption spending. The data is held in a secure internal cloud environment that only BBVA employees and a limited number of non-BBVA individuals can access. Unlike previous papers that used BBVA transaction data (Carvalho et al. 2021, García et al. 2021), we go beyond debit and credit card transactions and consider all account outflows. To the best of our knowledge, this is the largest comprehensive spending dataset currently available for research.

In this section we describe how we filter and transform this vast database to build granular measures of consumption spending.[1] In the main text we provide a conceptual overview, and full details are provided in appendix A.

2.1 Basic data structure

2.1.1 Types of transaction

The outflow transactions we observe fall into three broad groups each with their own associated metadata. First, there are debit and credit card transactions. Each transaction has an associated Merchant Client Code (MCC) that takes one of 838 unique values. MCCs are associated with retailers and provide information about the type of good bought in the transaction. They are a mix of generic international codes and Spanish-specific codes that refer to particular large retailers. A full list of codes is available at `https://www.dropbox.com/s/hroh7azjemtdh5x/mcc_to_coicop.csv`. We also generally observe the legal (tax) identity of the firm associated with the transaction which in Spain is termed a NIF code. One important exception is cash withdrawals from ATMs. These are recorded as card transactions with a dedicated MCC but the goods they are ultimately used to purchase are not available.

[1]All our database operations are GDPR compliant and have received additional legal approval from BBVA prior to execution.

Second, there are direct debit transactions. BBVA uses an internal classification procedure that assigns one of approximately 100 labels to each transaction, for example utility bill payment; council tax payment; and also more generic categories. In most cases, a NIF is available for the transactions in which the creditor is a firm, as is the firm's NACE sector code. If they choose, consumers can provide a free-text description of the transfer and in many instances this field is populated and searchable.

Third, there are non-recurring transfers which have the least associated metadata. For each beneficiary, we observe the International Bank Account Number (IBAN) of the receiving bank account and the name as a character string. In case the beneficiary is a BBVA client, we can use internal lookup tables to map the IBAN back into NIF and NACE codes. Otherwise, we resolve the firm by matching the beneficiary name with firms named in external commercial registry data.[2]

In appendix A we detail how we filter transactions to focus on those related to consumption rather than generic spending. The broad strategy is to use metadata on the beneficiary (MCC, NACE, etc.) to identify consumption-relevant spending categories. While all transactions originate from Spanish-domiciled bank accounts, beneficiaries can be located anywhere in the world. Transactions are available from 2015Q2 and are updated continuously. We stop the sample in 2021Q4 for this paper.

2.1.2 Sampling frame

Each transaction is linked to individual BBVA clients. While clients can hold multiple BBVA accounts, each unique account owner is assigned a customer ID that we can use to consolidate transactions originating from all that person's BBVA accounts. Information available about customers includes age, sex, and zip code of residence. In case a BBVA customer has co-signed a financial contract with BBVA, we also observe the customer IDs of all parties to the given contract. Finally, we observe whether clients have reported being self-employed and in this case we remove them from the data since we cannot distinguish their consumption purchases from input purchases for their firms.

In total there are 10,270,041 unique BBVA customers who conduct at least one consumption-related transaction in our sample. By way of comparison, the resident adult population of Spain in 2021 was 39,177,710. At the same time, many of these customers spend infrequently or in only a limited number of quarters. To define a consistent, stable sampling frame to measure consumption, we define *active* customers as those who make at least ten consumption-related transactions in each quarter in our sample. There are 1,831,040 such customers, after removing self-employed customers whose transactions might reflect production inputs instead of consumption.[3] Figure 1 shows the total number of unique customers observed per quarter compared to the number of active customers. One observes a steady growth in overall customer numbers. Our balanced panel ensures that any observed growth in aggregate spending is driven by spending increases within clients rather than a mechanical effect arising from increasing BBVA market share.[4]

Another important aspect of working with naturally occurring data is that—even after defining a balanced consumer panel—the demographic characteristics of the sample may not be representative of the Spanish population. Figure 2 compares the distribution of geographic location of residence, age, and

[2] We exclude common Spanish names from the list of potential firms. For example, if 'José González' appears as a beneficiary name in the transfer data, we do not treat this as consumption spending even if a legally registered firm exists with 'José González' in its name. This is because we assume transfers to private individuals are not related to consumption, and that transfers to accounts with common names are more likely received by individuals than firms.

[3] There are 181,918 such clients in the data that we drop. The 1,831,040 total is also after dropping a small number of outliers the procedure for which we describe below.

[4] To be fully consistent with how national statistical agencies conduct consumption surveys, we could re-sample the set of relevant customers at periodic intervals. Over the relatively short time series we use in this paper, we do not anticipate this to be important but over longer periods it would be.

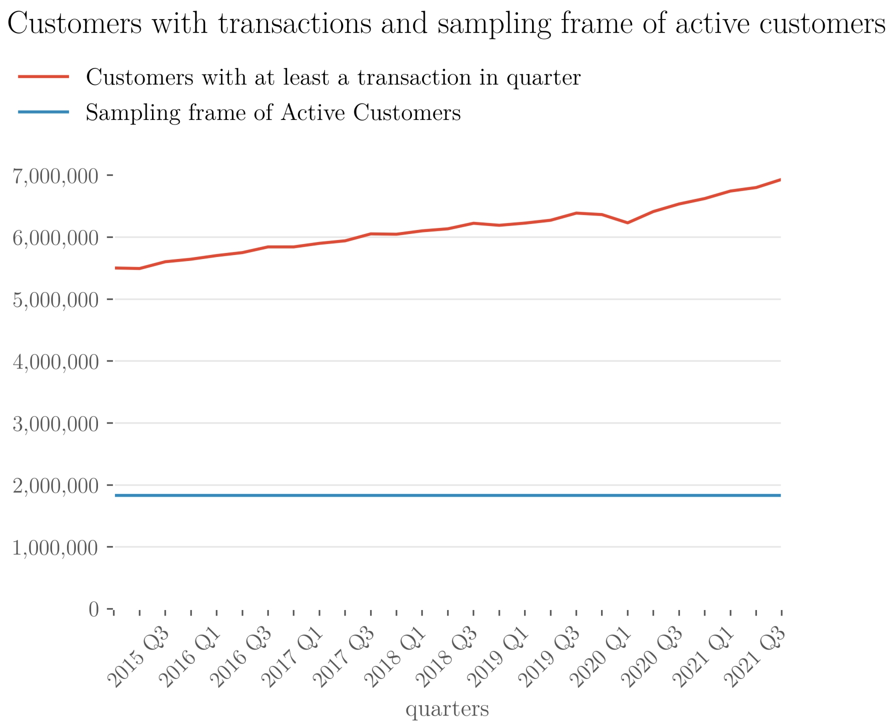

Figure 1: Frame of Active Customers

gender for the population of active clients against those for all Spaniards as recorded by the census.[5] While the distributions are clearly related, important discrepancies exist. BBVA active clients are over-represented in a particular region, among men, and among the middle aged. When we come to form aggregate consumption measures, we address these imbalances appropriately.

Table 1 tabulates the number and volume of transactions made by active clients in our sample that we classify as related to consumption, broken down by transaction type. We separate cash withdrawals from other transactions,[6] and do not include transfer payment related to rent, which we treat below as a special category. The total spending value is roughly 200 billion euros encompassing three billion total transactions. While card transactions make up a large majority of total transactions, their total value is comparable to that of direct debits.

Table 1: Consumption data volume of Active Customers

Spending Category	Volume of Transactions	Number of Transactions
Offline Card Transactions	60,319 million €	1,772 million
Online Card Transactions	11,858 million €	313 million
Direct Debits	66,036 million €	752 million
Cash Withdrawal	64,592 million €	359 million
Transfers excl. rent	11,148 million €	15 million

[5] We use the census year 2018 for the comparison. The geographic distribution is reported at the level of the 19 primary regions in Spain known as Comunidades Autónomas. For age, we only consider those at least 18 years old.

[6] There are two kinds of cash withdrawal in the data. The first is ATM withdrawals conducted with debit and credit cards. The second is cash extracted at BBVA branch offices from teller windows. The former is more prevalent in our data, especially in later years.

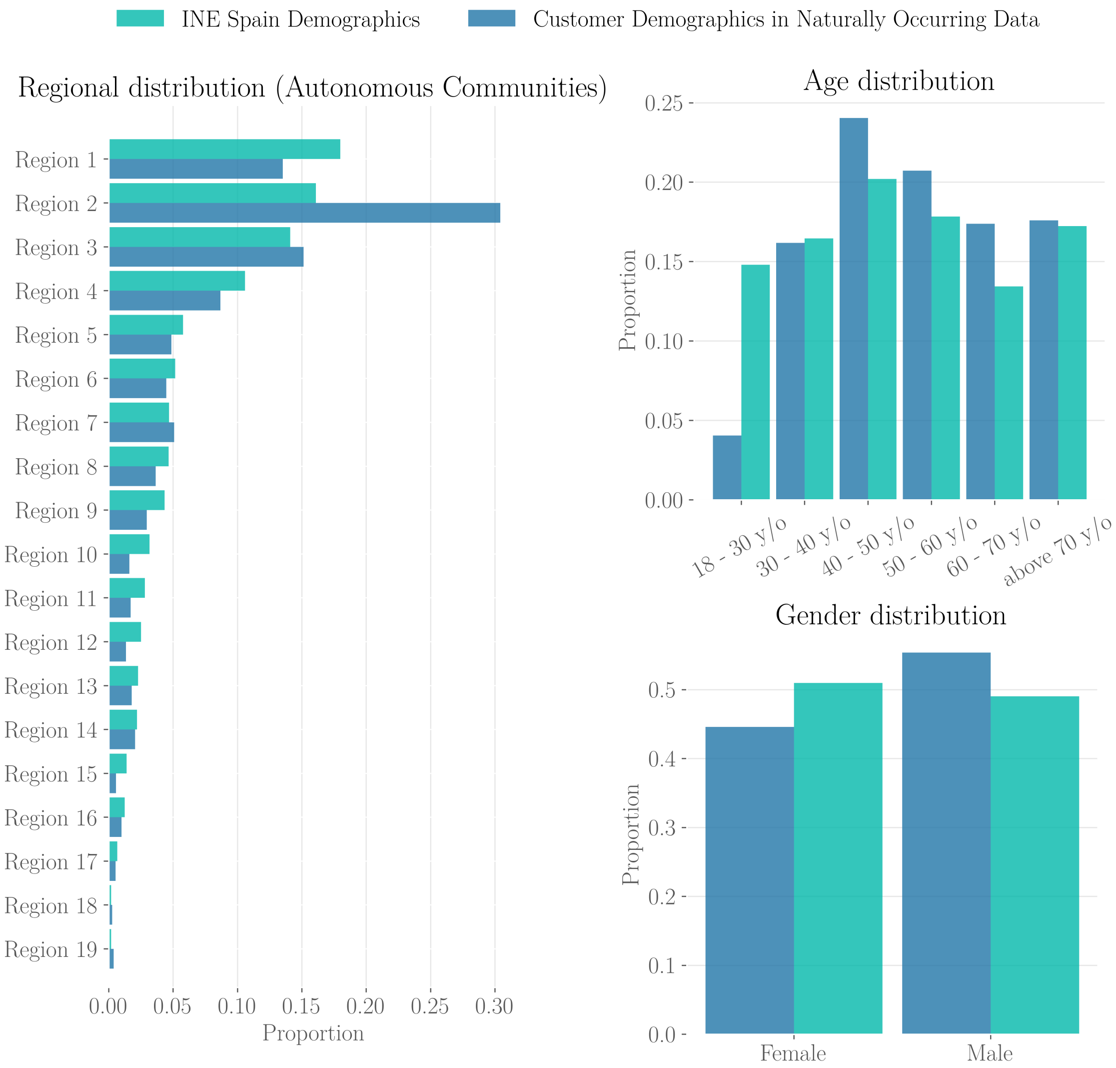

Figure 2: Demographics of Active Customers

2.1.3 Household structure proxy

At several points in our construction, it is useful to consider the unit of consumption as a household rather than an individual. BBVA does not directly record information on its clients' household structure, and so we proxy it from available information.

We first identify—for each active client—the set of BBVA customers who have both co-signed a financial contract (e.g. are co-owners of a bank account, jointly liable for a loan, etc.) at any point in the sample and reside in the same postal code at the end of the sample. This creates an initial estimate of the number of people in each active client's household besides himself. In cases where active clients appear in each other's sets, they are joined together into a single household. This procedure creates 1,592,356 household groups.

In cases where an active client remains unmatched to any other BBVA client but is listed as married, we assume s/he resides with one other person, e.g. a spouse.

Finally, BBVA records for each client the number of dependent adults in the household. If after the above steps an active client is grouped with fewer individuals than appear as dependent adults, we recode the number of additional household members as equal to the number of dependent adults.

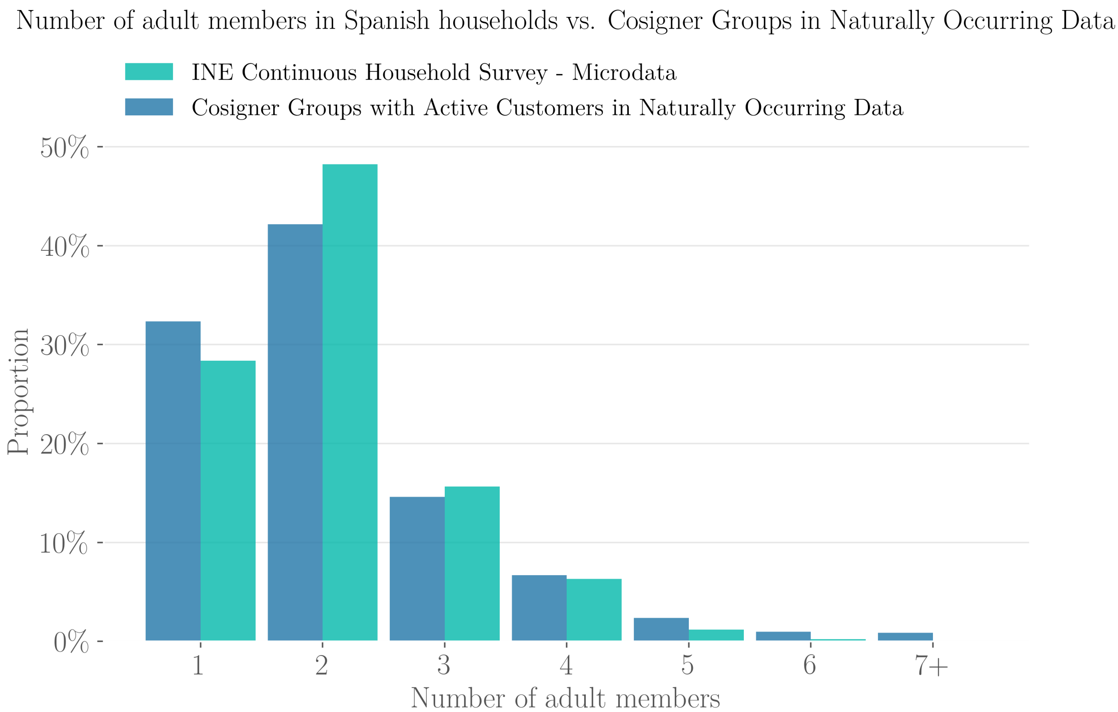

Figure 3: Household Proxy vs Official Data

Figure 3 compares the resulting distribution of household sizes according to our grouping procedure against official data. Since our households size estimates are based on adults, we use INE's *Continuous Household Survey*[7] and extract from each surveyed household the number of adults. While there are some discrepancies in the two distributions, overall they track each other quite closely which suggests our grouping procedure is a viable estimate of household size in the absence of direct data.

[7]https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176952&menu=resultados&idp=1254735572981

2.2 Categorizing consumption categories

The metadata associated with each transaction associates it with the type of good being purchased. As noted above, though, different transactions have differently structured labels and none have labels that readily map into consumption categories that appear in national accounts. We now describe how we harmonize all transactions and map them where possible into categories defined by European Classification of Individual Consumption According to Purpose (ECOICOP) at the two-digit level, which we detail in table 2. We adopt the ECOICOP classification because it is used by INE to report consumption categories and so facilitates comparisons to national statistics.[8]

Table 2: ECOICOP Consumption Categories (Two-Digit)

Category	Description
01	Food and Non-Alcoholic Beverages
02	Alcoholic Beverages, Tobacco, and Narcotics
03	Clothing and Footwear
04	Housing, Water, Electricity, Gas, and Other Fuels
05	Furnishings, Household Equipment, and Routine Household Maintenance
06	Health
07	Transport
08	Communication
09	Recreation and Culture
10	Education
11	Restaurants and Hotels
12	Miscellaneous Goods and Services

For all categories but 04 we rely on observed transactions. Housing is a unique category in that it makes up a large portion of household consumption and has an imputed component for the notional rental value of owner-occupied housing. We discuss below how we perform this imputation using BBVA data.

Finally, it is worth emphasizing that many transactions do not have a clear categorization. The most prominent example is cash, which as shown in table 1 makes up a significant portion of total spending. We therefore designate a separate category of *non-classifiable consumption* for transactions we assume are related to consumption but for which a clear mapping into ECOICOP is not available. We treat cash in this way, but also other types of transactions, for example card transactions with the MCC corresponding to *shopping clubs*.

2.2.1 Non-housing consumption

The mapping from card transactions to ECOICOP categories is primarily achieved through directly manual labeling of all 838 MCC codes in the dataset. We provide this mapping at `https://www.dropbox.com/s/hroh7azjemtdh5x/mcc_to_coicop.csv`, which should be useful for other researchers as well who wish to organize card transactions into official consumption categories. The main complication is that a limited number of MCC codes refer to the sales of multi-product retailers such as supermarkets. In these cases, we use published statistics on the distribution of sales across ECOICOP categories to allocate shares of a transaction's value appropriately. Full details are provided in appendix A.

For direct debit transactions, we begin by manually classifying the BBVA-provided labels into ECOICOP categories. Because the labels are proprietary, we do not provide the mapping publicly. Again,

[8] The ECOICOP is very similar to the international COICOP scheme. The main difference is that the latter has two separate categories *Insurance and financial services* and *Personal care, social protection and miscellaneous goods and services* which in ECOICOP are merged into a single *Miscellaneous Goods and Services* category.

though, we obtain a mix of ECOICOP categories (with the same treatment of multi-product retailers as for cards) and non-classifiable consumption. For the non-classified transactions, we attempt to assign a consumption category using alternative strategies. First, we retrieve the creditor firm's tax ID and search for a match in the set of card transactions. When there is a match, we use the MCC-to-ECOICOP mapping to generate a consumption category. If the transaction remains non-classified, we retrieve the NACE code of the creditor firm and apply our own manually constructed NACE-to-ECOICOP mapping which is available at `https://www.dropbox.com/s/9lcab2zajijxltn/nace_to_coicop.csv`

For non-regular transfers, there are 17 counterparties that we classify as consumer credit firms based on their NIF. Twelve of these are non-classifiable and five are related to cars and so are assigned to category 07. For the other cases, use the counterparty's tax ID to first attempt a match into the set of firms that receive a direct debit. In case of a match, we use the manual mapping from direct debit labels. If this fails to produce a category, we next attempt a match into the set of counterparty firms for the card data and use the manual mapping of MCCs. If this too fails we use our manual mapping of NACE codes as a final attempt. Finally, if no category is assigned after these steps, the transfer is left unclassified.

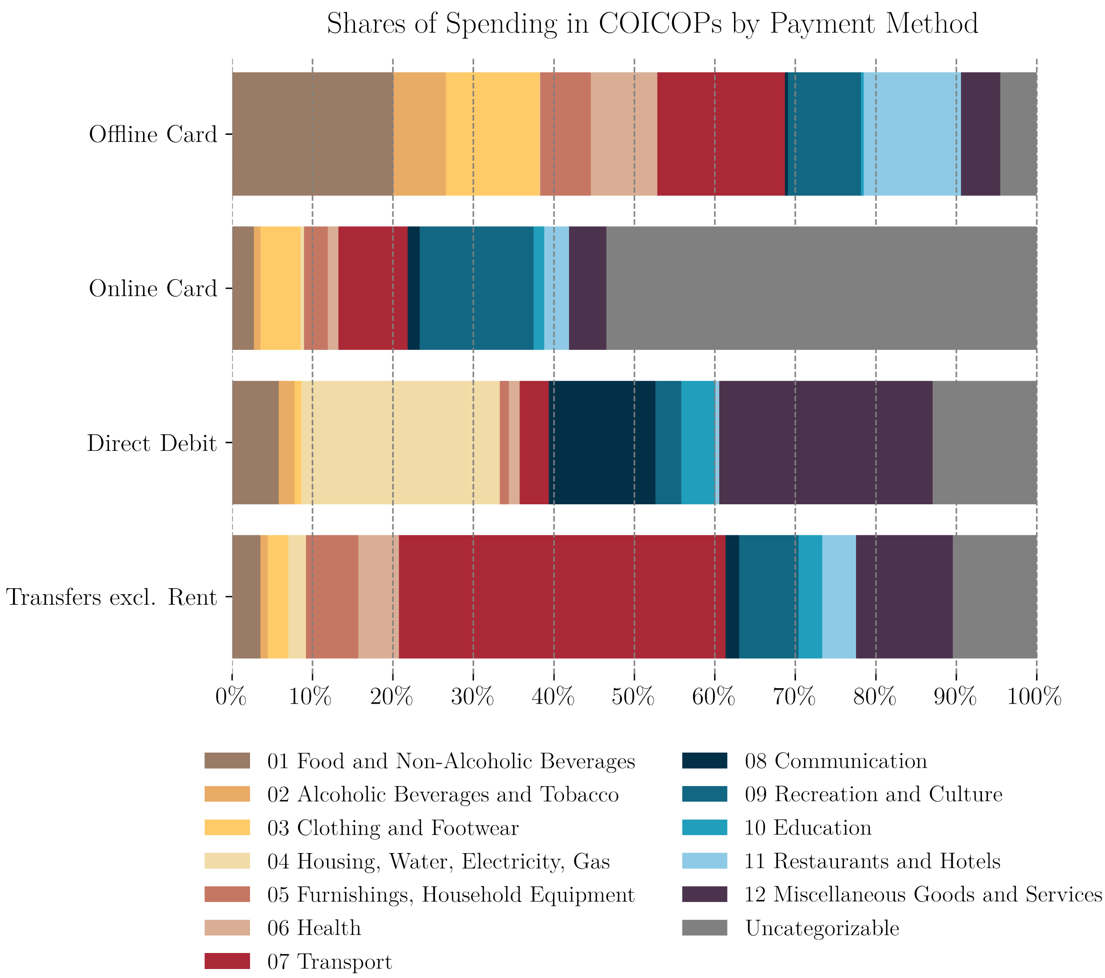

Figure 4: ECOICOP Shares by Payment Method

Figure 4 shows the distribution across consumption categories by payment methods. One observes substantial heterogeneity across methods. Food spending makes up a substantial part of offline card spending, but less of other modes'. Transportation makes up nearly half of irregular transfers, while utility payments mainly come via direct debits.

2.2.2 Housing consumption

Housing presents a challenge for consumption accounting because many people do not pay rent for the homes in which they reside, primarily because they own them. In this case, statistical agencies must impute the rent for owner-occupied housing and this alone can account for a sizable part of aggregate consumption.

In the BBVA setting, we face this challenge but also an additional one: BBVA clients who do rent their residences but whose rental payments we cannot retrieve from the transaction database. This also requires imputation of missing data.

We begin by extracting those rental payments we can from the direct debit and transfer tables. The natural unit of analysis for housing consumption is a household, so we search for payments made by all individuals who make up households whether or not they are active clients. We identify rental payments based on the free text field that customers can use to describe the payments they make. The search terms we use are variants of 'rent' or 'rental' in Spanish and other regional languages. We *exclude* transactions that additionally include terms that suggest the rental payment is for a non-housing asset, like a garage, parking space, or car. We also impose a minimum value of 100 euros for a transaction to be considered rent. This allows us to construct a rental total for each household and month in the sample. 437,307 households have at least one

Our next goal is to estimate a regression model of household rent based on household observables that we can use to impute rent for households with no or sparse rental information. To avoid noise arising from households with few monthly rental observations, in our estimation sample we limit attention to households with non-missing rental payments in at least 70 of the 81 total months in our sample. There are 32,127 such households.

The household covariates we use to predict monthly rent are income, utility payments, and geographic location. For income we rely on an auxiliary BBVA data table that records monthly income from wages, government benefits, and pensions. We use this to compute six-month rolling average household income. Utility payments are computed from the direct debits table and expressed as rolling three-month totals. We only keep households in the estimation sample that have at least one month of observed utility payments and income. This reduces the number of households to 16,977. Table 3 provides summary statistics for household-level observables in this set.

Table 3: Summary Statistics of Train Sample for Rent Regression

	Rent	**3 Month Total Utility Expenditure**	**6 Month Average Income**
Mean	551.1 €	293.0 €	2385.5 €
SD	259.1 €	240.0 €	1918.3 €
25%	400.0 €	148.9 €	1411.4 €
50%	500.0 €	236.5 €	2010.6 €
75%	650.0 €	365.5 €	2850.0 €

For geographic location, we seek to define spatial units that are sufficiently well populated with households that fixed effects can be reliably estimated. To do so, we apply the following algorithm within each of the 52 Spanish provinces:

1. Begin with regional units defined by the set of postal codes present in the observation sample.
2. Iterate as follows until each regional unit has at least 30 households or until the entire province has been consolidated, whichever occurs first:

(a) Identify the regional units with the fewest number of households.

(b) Combine these regional units with the closest regional units based on Haversine distance computed between centroids, which forms a new set of regional units.

Figure 5 illustrates the final result of the algorithm for the province of Madrid. The original units are the distinct postal codes, and the colored blocks represent our final unique regions. Whereas we begin with 2,687 unique postal codes in the observation sample, our algorithm produces 327 regions for the regression model. The average number of households in each region is 52.0, and the average number of postal codes is 8.2.

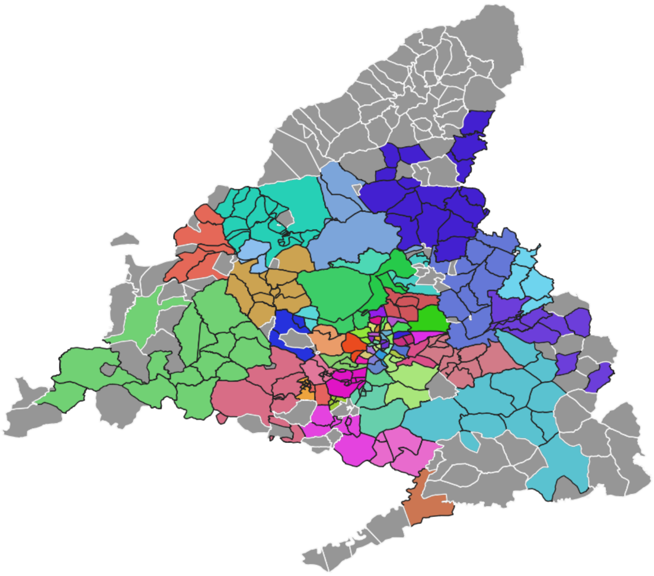

Figure 5: Merged Regions in Madrid

Finally we regress monthly rental payments by household on region fixed effects, income, and utility payments. Table 4 displays the results. Although simple, our model explains 40% of the variation in rental payments and both continuous covariates are highly significant and contribute to high within-region R^2. The estimated coefficients imply that a one standard deviation change in income shift rental payments by 70 euros a month, or roughly one third of the IQR of the overall rental payment distribution. The impact of utilities is more muted, with a one standard deviation change shifting rent by 21 euros.

Table 4: Regression for Rent

Variable	**Model**	**Test set**
Spending on House Utilities	0.0884	
	(0.0008)	
Income	0.0362	
	(0.0011)	
N of Contract Groups	16,977	15,512
N of Observations	1,134,735	15,512
R^2	0.3911	
Adjusted R^2	0.3765	
Within R^2	0.1200	
Root MSE	204.6144	221.64

We then use our estimated rental regression to impute monthly rents to all households not in our estimation sample (the vast majority). Where a household lies outside the regions defined for the estimation sample, we assign it to the closest region based on centroid distance. Where no income or utility information is available to form a given month's rolling average, we use the household average over all months if any record is available. Otherwise we assign the region average.

To form an initial estimate of out-of-sample accuracy, we consider the 15,512 households for which we observe between 50 and 70 monthly rental payments and compute the root mean squared error of the imputed rent with respect to actual rent for a randomly drawn month for each household. The RMSE rises only slightly with respect to that of the estimation sample, which suggest that our rent model, while simple, generalizes well out-of-sample. The averages also line up well: the actual average rent is 551 euros and the imputed average rent is 538 euros.

2.3 Household and demographic weighting

So far we have described how we select the set of BBVA clients to form our naturally occurring consumption survey, and how we tabulate transactions into consumption categories. The final step in deriving our consumption measures is to adjust spending totals by household and demographic structure.

Accounting for household structure is important because part of each active clients' spending is potentially undertaken on behalf of others. On the other hand, since we do not tabulate the spending of non-active clients (except for housing) we are missing the part of their spending that benefits active clients. To balance these effects, we adopt the following weighting scheme. Suppose active client i is grouped to a household with a_i active clients (including himself) and n_i non-active clients. Then if c_i^{raw} is the raw consumption spending of i, we use $c_i^{\text{hh}} = \frac{c_i^{\text{raw}}}{a_i+0.5n_i}$ as the weighted measure. Hence if a household is made up exclusively of active clients, we assume all spending by any client is shared equally among household members. On the other hand, the spending that spills over to non-active clients is down-weighted by 0.5.

In much of analysis below, we aggregate individual spending into larger units. To do this, we define cells at the gender, age, region level using the categories displayed in figure 2. Let $c_{g,a,r}^{\text{hh}}$ be the total household-adjusted spending of all active clients with gender g, age a, and residing in region r. Also, let $x_{g,a,r}^{\text{BBVA}}$ be the total count of active clients in cell (g, a, r) and $x_{g,a,r}^{\text{INE}}$ be the total count of Spanish adults according to census data. Then our consumption spending measure for each cell is

$$c_{g,a,r} = c_{g,a,r}^{\text{hh}} \left(\frac{x_{g,a,r}^{\text{INE}}}{x_{g,a,r}^{\text{BBVA}}} \right)$$

From here one can form arbitrary data aggregate by summing over the cells. Aggregate consumption is the sum over all cells; regional consumption for r is the sum over all gender and age categories holding r fixed; and so on. Category-specific consumption is derived by only considering c_i^{raw} that correspond to ECOICOP of interest. In the next section, we discuss properties of such derived series.

3 Properties of Consumption Measures

3.1 Time Series of Aggregate Consumption.

The Spanish Statistical Institute (INE) computes aggregate consumption at a quarterly frequency. Our naturally occurring data can be aggregated in multiple manners both in the time domain and across people. In this section we show that the aggregation into national level quarterly data tracks official statistics, and can thus be used in conjunction with it to understand macroeconomic movements. We

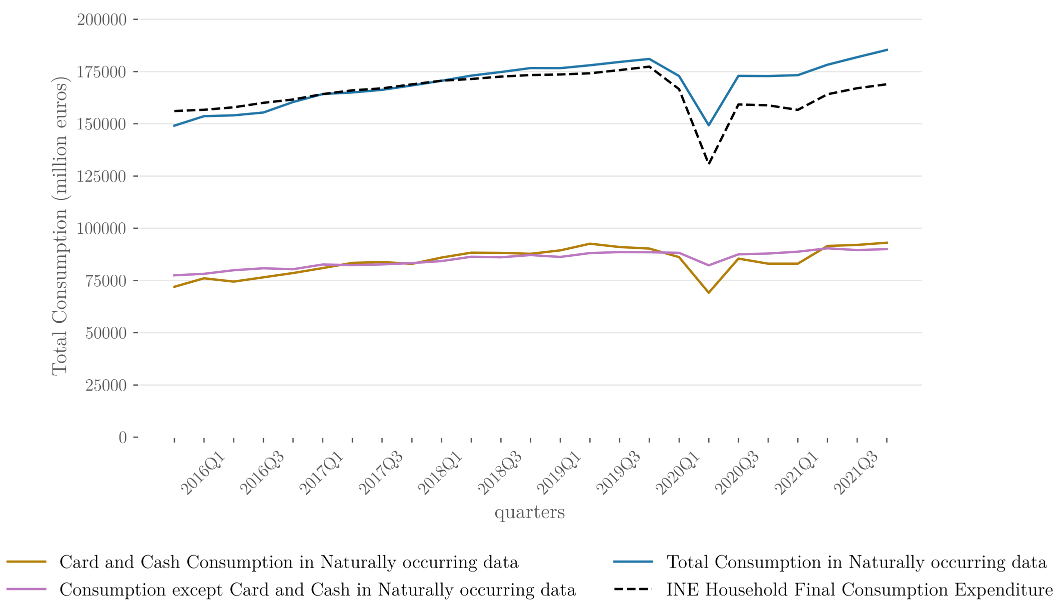

Figure 6: Final Consumption Levels. Quarterly

have here the double aim of validating our data and remark the need of including all type of transactions from financial institutions (and not only those generated by certain means of payments) when considering aggregate consumption.

Figure 6 plots four time series of quarterly data. The series labelled "INE" is the latest release of Final Household Consumption Expenditure[9] by INE, the natural aggregated benchmark of our data.

The series labelled "Total Consumption in Naturally Occurring Data" plots the final aggregation of our series at quarterly frequency. It consist of a simple summation of the expenditures that we deem as consumption assigning proper weights to regional, age and gender considerations, as explained in the previous section, plus the housing imputation.[10] Notice that the *levels* track each other exceedingly well, and there was no ex-ante guarantee that this would be the case. Our series and INE's are computed using vastly different technologies. INE's series is collected by summing up sales of firms and filtering the data via a statistical model. Ours can be viewed as the aggregation and escalation of a vast consumption survey.

The other two series present the evolution of all expenditure according to the means of payment used. The series "Cash and Card Consumption" is the aggregation of expenditure paid via card (both online and offline) plus extractions of cash by final consumers (be it via card at ATM or via direct extraction from a bank office). The remaining series ("Consumption except Cash and Card") aggregates the remaining of our consumption imputations: transfers, direct debits and housing services (both rentals of residences and imputations of owner occupied dwellings). Notice that both series are of about equal magnitude, suggesting that extraction information from only some of this methods of payments could bias results. In particular, the bulk of the literature using transaction data uses card data to infer aggregate consumption behaviour, which could be problematic in several counts.

First, because there are secular trends that can be observed even with the relatively sort time series at our disposal (2016-2022). In figure 7 we present a portfolio plot with the evolution of the monthly share of the six means of payments on total consumption. It is notorious that both online and offline

[9]Accessible from INE's webpage

[10]Following Eurostat procedure both the Naturally Occurring data and INE series are seasonally adjusted using the Jdemetra+ application and apply X-13ARIMA-SEATS

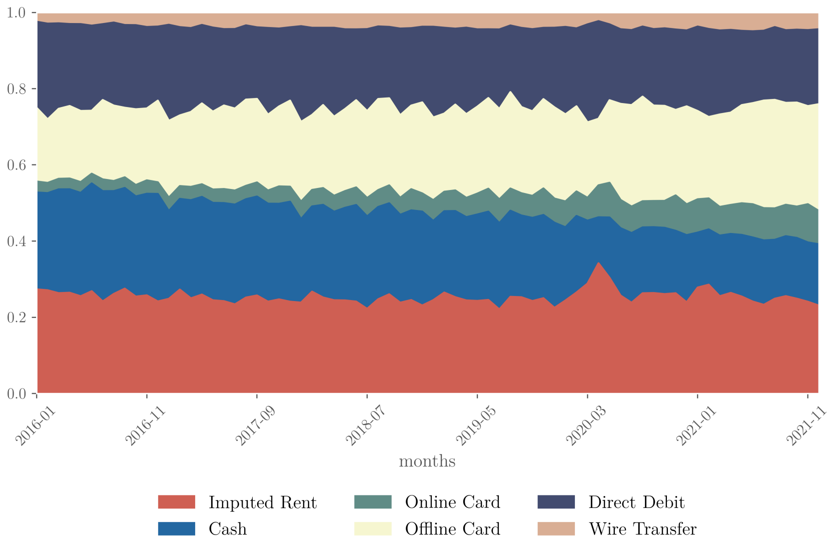

Figure 7: Shares of different payment methods

card transactions have a secular increase in share at the expense of cash. Transfers, direct debits and housing are approximately stable overtime, with the only notorious shock due to COVID-19 and the consequent Lockdown. We turn to that next, but before let us emphasize that if we were to rely only on card data not only we would have obviously missed more than fifty percent of consumption, but we would have over estimated its long rum growth.

Perhaps more important is that the volatility is substantially different across means of payments. Figure 8 presents a time series of the quarter on quarter rates for the four series, including the massive shock on consumption due to COVID-19 and the consequent Lockdown. It is apparent that the average levels of growth are not dissimilar in the four series, but also that and that the reaction to the COVID shock is much more dramatic when counting only cards and cash transactions, and much more muted for transfers, direct debits and housing. Notice also that the reaction of the official series to COVID follows closer the behavior of card and cash transactions. At this stage we do not know if this difference between our series and INE is a consequence of the official data not having included yet the re-evaluations that routinely take place as new surveys on retrospective data (typically annual in nature) inform the statistical office. Given the unique and massive effect of that event, it is certainly possible that the statistical model relies does not still capture the fact that housing, transfers and direct debits had (not surprisingly) a much smoother behavior than cards and cash purchases.

3.2 Consumption categories and some comparisons to Household Survey

Perhaps the best way of thinking on our data is as a massive survey on consumption with a panel structure in which the time dimension is exceedingly frequent. Without using yet the micro-structure of the data, in this section we aim to show that our data conveys the same rough structure of consumption patterns that the survey of reference in Spain. The main survey of consumption performed in Spain is the "Encuesta de presupuestos familiares" (Household Budget Survey) which interviews a large number of households (23920 in 2019) at an annual frequency obtaining detail information on their patterns of consumption.[11]

As explained in section 2.2 we develop an algorithm ascribing all expenditures in our data into one

[11] Available here

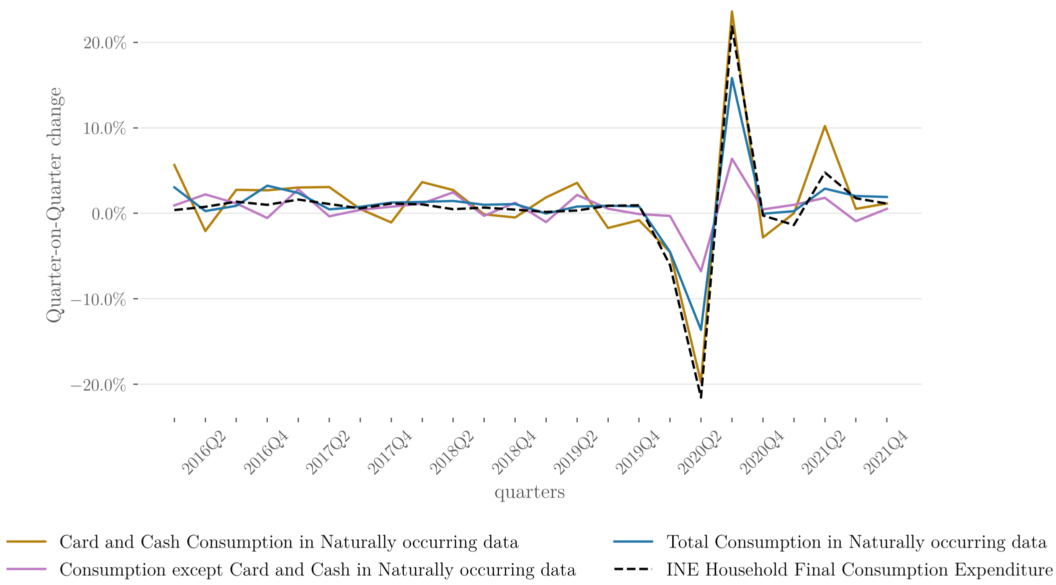

Figure 8: Quarter on Quarter growth rates (deseasonalized)

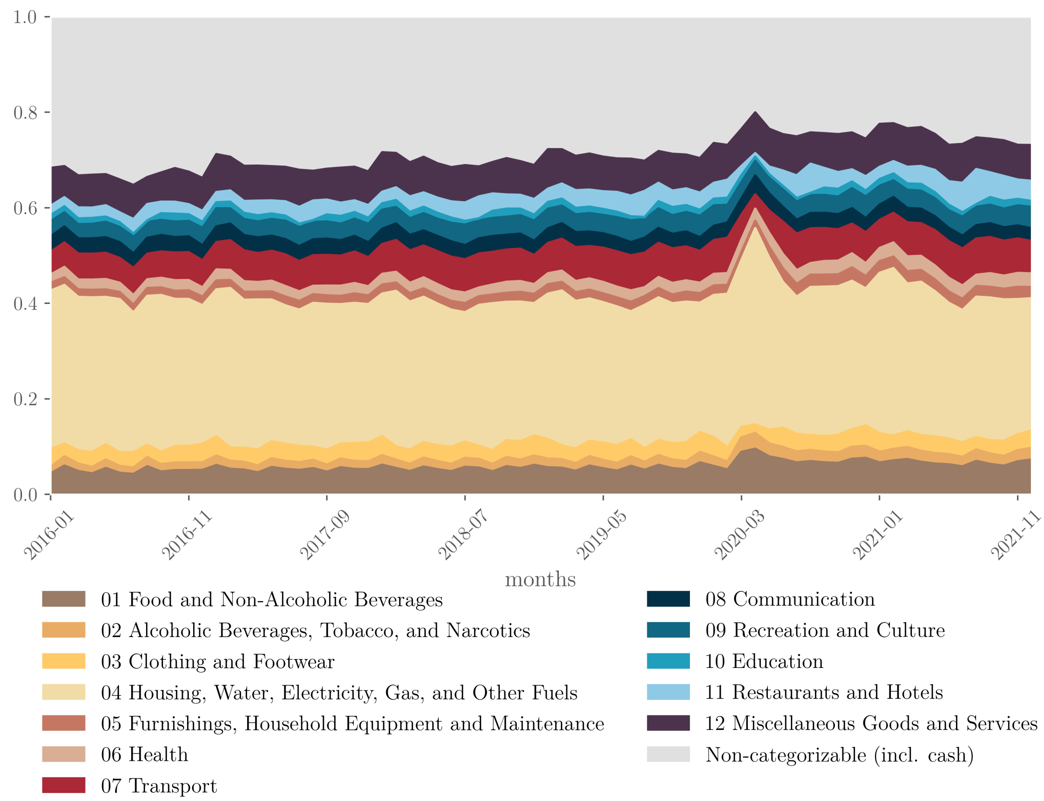

Figure 9: Proportions of Consumption by COICOP

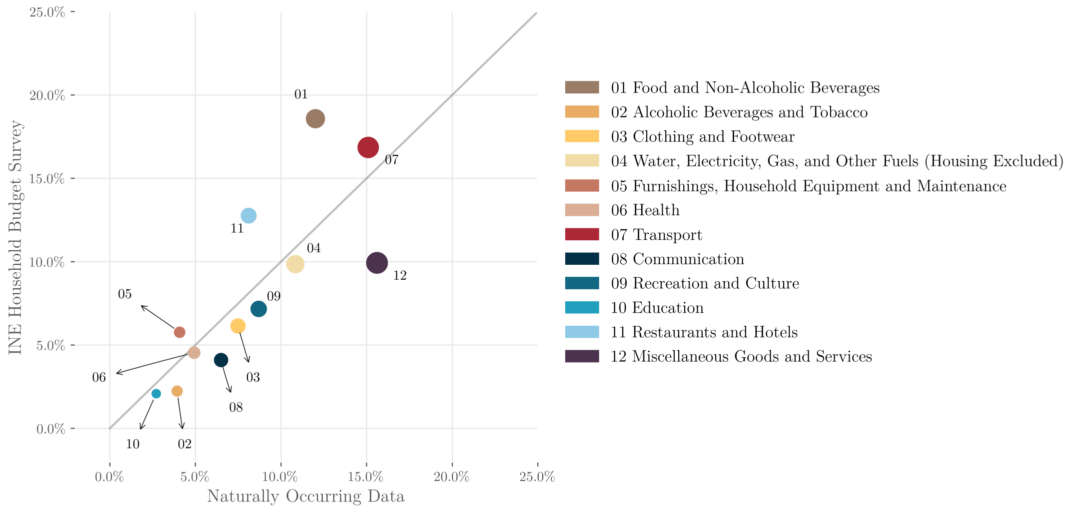

Figure 10: EPF versus Naturally Occurring distribution of consumption per categories (2019).

of 12 COICOP categories (plus a set of non-classifiable expenditures which includes cash).

In figure 9 we plot the evolution of the shares of these thirteen categories at a monthly frequency. Housing services represent the bulk of the consumption expenditures, while non-categorized (mostly cash expenditure) follows in magnitude, other than that the only noticeable fact is that with the exception of the COVID shock, the distribution of spending is remarkable stable.

Figure 10 compares the shares of consumption per categories in the EPF of wave of 2019 with the averages for the same year in our data. They track each other reasonably well, with a raw correlation of 0.798. In the figure we do not include into COICOP 4 (housing and utilities) the imputation of housing services as its magnitude (around 40% of spending) increases the correlation artificially to a level of 0.94.

Notice that this correlations are remarkable, particularly if one realizes than, obviously, we do not include the set of non-categorizad purchases, which correspond to more than 25% of spending in our data and it is concentrated in a particular means of payment (cash) that most likely is not spread across categories in the same manner than other means of payments. We thus conclude that our data and algorithm categorize consumption in a manner very much compatible with the one observed in the surveys. Moreover, the nature of our data makes it available at a much higher frequency and it is cheaper and easier to get than a comparable survey (as the EPF). These two advantages are patent in the development of the COVID crisis.

3.3 Durables versus non durables

TBW

4 Some examples of use of our data

4.1 Restrictions imposed by COVID

Our data, even without using its micro-structure, allows to make explorations which would be almost impossible otherwise. It allows to cut the data in arbitrary manners. Thus, as an example, figure 11 plots the time series of per capita levels of consumption per age groups at quarterly frequency (deseasonalized). It is apparent an increase of the spread of average consumption per age. The increase in the differences that was apparent before COVID, and while it was (of course) dramatically reduced during the lockdown,

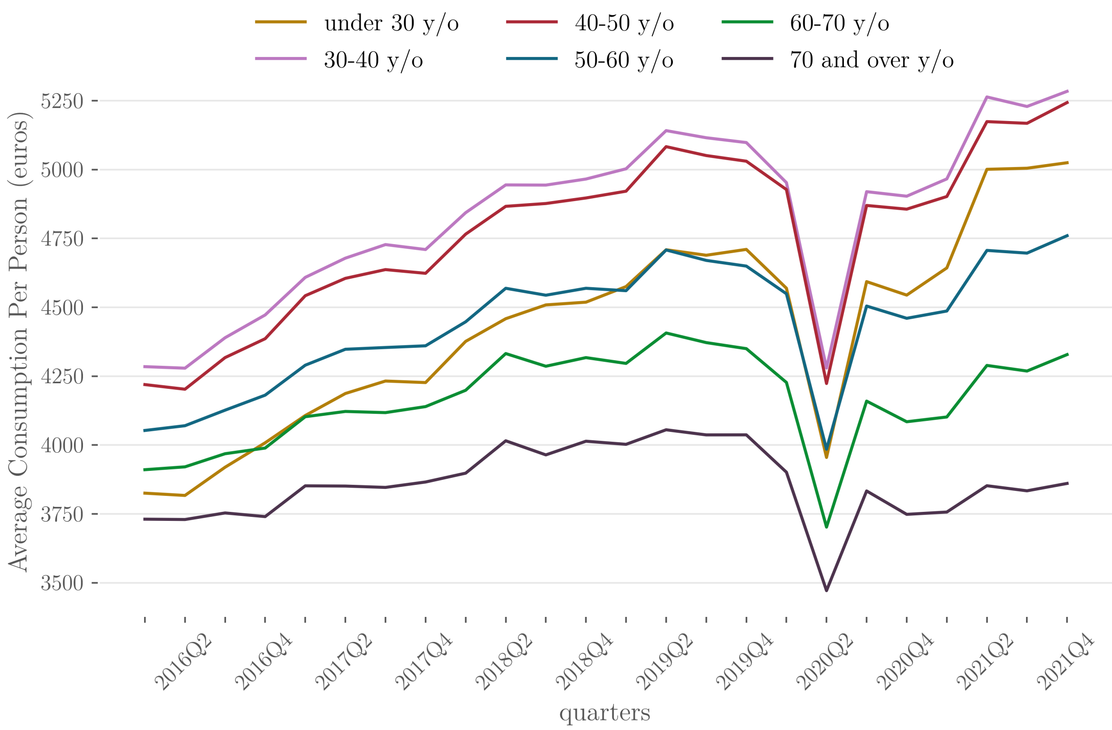

Figure 11: Levels of consumption per age group

it has amplified notoriously since. Notice that for people over 70 the path is essentially flat (and the level remains lower than pre-covid), while it has surpassed pre-covid levels for all age groups below 60. Notice as well the surprising path of consumption of young individuals: their per capita consumption was increasing exceedingly fast before COVID, and while it plummeted more than for any other age group in the lockdown, it has since surged. Apparently, Spain is Not a Country for Old Men.

This path for young individuals can be visualized in the context of the restrictions placed by COVID. Madrid and Barcelona are the two largest cities in Spain, and are comparable in many dimensions, but the governments of the regions to which they belong have had very different attitudes towards imposing restrictions to activity, particularly in what respect to hospitality industry. During the first weeks of lockdown activity was regulated by the Spanish central government, initially equally in all the country (during the very strict lockdown), and later following province to province decisions based on the perceived level of infection. But the Spanish State is extremely devolved, and once the state of emergency was canceled (summer 2020) all decisions on restrictions were devolved to the regional governments. On this respect Madrid's government has had a much more relaxed attitude, emphasizing testing, while the government of Catalonia has been much more prone to restricting activity. Thus, in December 2020 Madrid essentially did impose only minor restrictions to the activity of the hospitality industry, while in Barcelona the restrictions were stiff (in schedule, minimum distances and the mere possibility of opening).

The upper panel of Figure 12 plots average consumption in Restaurants and Hotels in December 2019 and 2020 of individuals living in each postal code of both Madrid and Barcelona. At first glance it does not seem very informative, it indicates that in 2019 consumption was higher in both cities in richer neighborhoods. In 2020 while there was a decrease in consumption of Restaurants and Hotels in both cities. Albeit it is clear that the decrease is mostly concentrated among the richer neighborhoods (because Restaurant and Hotels are the kind of things than richer people consume), the fall seems ballpark similar

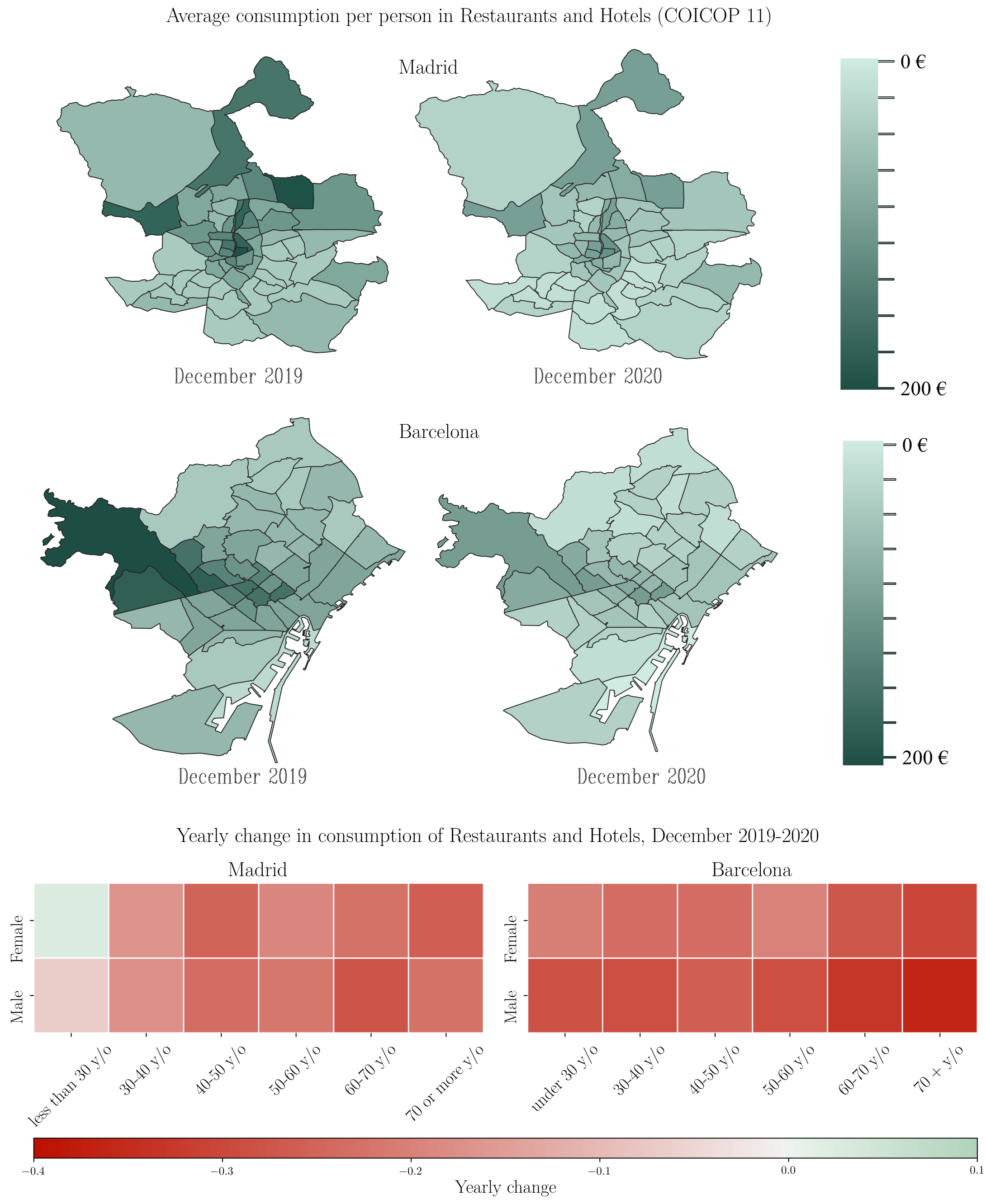

Figure 12: Private Consumption in Restaurants and Hotels, December 2019 and 2020

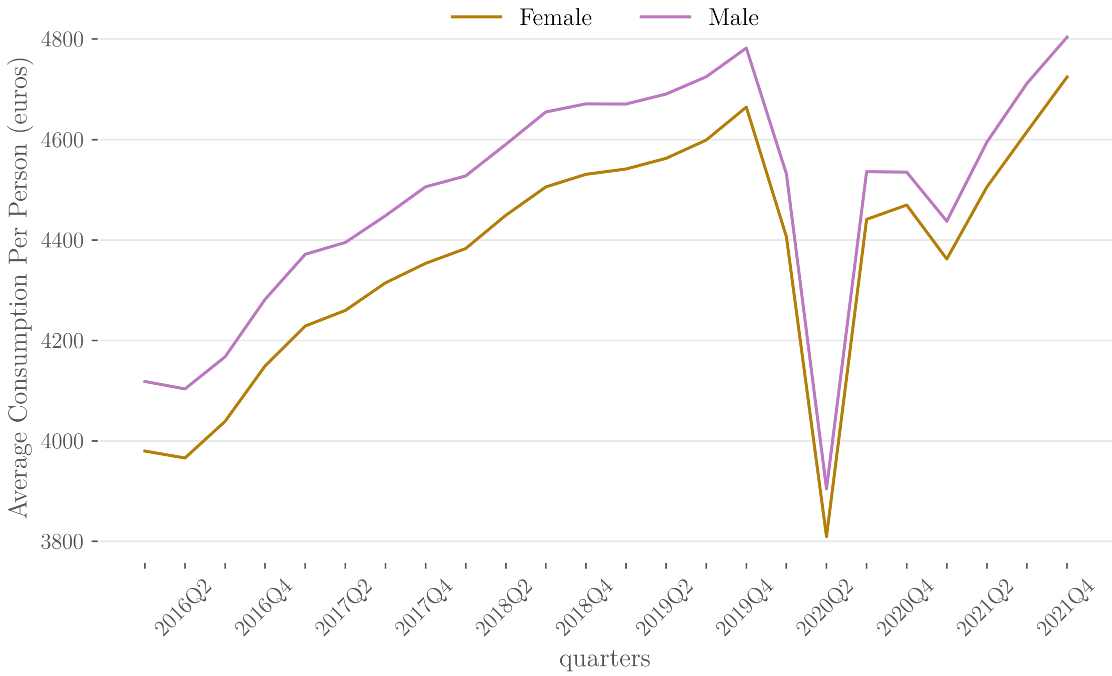

Figure 13: Levels of consumption per gender

in both cities. It would seem that the difference in policy did not have distinctive effects.

Interestingly, when looked with attention some remarkable differences appear. In the lower panel we present the change in consumption per age group and gender. It is now manifest that the average decrease is slightly bigger in Barcelona, but much more interestingly, the effects on the age composition are very different. Young people in Madrid decrease their consumption much less than in Barcelona. Actually, young females did *increase* their consumption.

This gender difference may seem surprising, but perhaps figure 13 may help explain it. There we plot the nation-wide per capita levels of consumption per gender. Two facts are quite salient. First, that the level of consumption per gender is higher for males, but not by a large amount. Second, that howver large or small this difference was, it has *decreased* in the period posterior to COVID. Not only during the lockdown (when obviously could not be very different), but even when the aggregate level has returned to pre-covid values.

Granted that in the process of household formation we make assumptions that affect the imputation of certain consumption to one or another member of the household. Thus, albeit we have done our best to make assumptions that stand to reason, one could think that the equality in levels in the pre-covid period might be due to our imputations. Nevertheless, our imputations are the same before and after COVID, so the decrease of this difference can not be attributed to our methods and needs to reflect a change in pattern of consumption in all likelihood related to the COVID experience.

4.2 Inequalities of Consumption

TBW

References

Aguiar, M. and Bils, M. (2015). Has Consumption Inequality Mirrored Income Inequality? *American Economic Review*, 105(9):2725–2756.

Aguiar, M. and Hurst, E. (2013). Deconstructing Life Cycle Expenditure. *Journal of Political Economy*, 121(3):437–492.

Aladangady, A., Aron-Dine, S., Dunn, W., Feiveson, L., Lengermann, P., and Sahm, C. (2021). From Transactions Data to Economic Statistics: Constructing Real-Time, High-Frequency, Geographic Measures of Consumer Spending. In *Big Data for Twenty-First-Century Economic Statistics*. University of Chicago Press.

Andersen, A. L., Hansen, E. T., Johannesen, N., and Sheridan, A. (2020). Consumer Responses to the COVID-19 Crisis: Evidence from Bank Account Transaction Data. SSRN Scholarly Paper ID 3609814, Social Science Research Network, Rochester, NY.

Attanasio, O., Hurst, E., and Pistaferri, L. (2014). The Evolution of Income, Consumption, and Leisure Inequality in the United States, 1980–2010. In *Improving the Measurement of Consumer Expenditures*, pages 100–140. University of Chicago Press.

Baker, S. R. (2018). Debt and the Response to Household Income Shocks: Validation and Application of Linked Financial Account Data. *Journal of Political Economy*, 126(4):1504–1557.

Baker, S. R. and Kueng, L. (2021). Household Financial Transaction Data. Working Paper 29027, National Bureau of Economic Research.

Barrett, G., Levell, P., and Milligan, K. (2014). A Comparison of Micro and Macro Expenditure Measures across Countries Using Differing Survey Methods. In *Improving the Measurement of Consumer Expenditures*, pages 263–286. University of Chicago Press.

Bean, C. (2016). Independent review of UK economic statistics: Final report. Technical report.

Carvalho, V., Garcia, J., Hansen, S., Ortiz, A., Rodrigo, T., Rodriguez Mora, S., and Ruiz, P. (2021). Tracking the COVID-19 crisis with high-resolution transaction data. *Royal Society Open Science*, 8.

Chetty, R., Friedman, J. N., Hendren, N., and Stepner, M. (2020). How Did COVID-19 and Stabilization Policies Affect Spending and Employment? A New Real-Time Economic Tracker Based on Private Sector Data. Working Paper.

Coibion, O., Gorodnichenko, Y., and Koustas, D. (2021). Consumption Inequality and the Frequency of Purchases. *American Economic Journal: Macroeconomics*, 13(4):449–482.

García, J. R., Pacce, M., Rodrigo, T., Ruiz de Aguirre, P., and Ulloa, C. A. (2021). Measuring and forecasting retail trade in real time using card transactional data. *International Journal of Forecasting*, 37(3):1235–1246.

Gelman, M., Kariv, S., Shapiro, M. D., Silverman, D., and Tadelis, S. (2014). Harnessing naturally occurring data to measure the response of spending to income. *Science*, 345(6193):212–215.

Jarmin, R. S. (2019). Evolving Measurement for an Evolving Economy: Thoughts on 21st Century US Economic Statistics. *Journal of Economic Perspectives*, 33(1):165–184.

Koijen, R., Nieuwerburgh, S. V., and Vestman, R. (2014). Judging the Quality of Survey Data by Comparison with "Truth" as Measured by Administrative Records: Evidence From Sweden. In *Improving the Measurement of Consumer Expenditures*, pages 308–346. University of Chicago Press.

Kreiner, C. T., Lassen, D. D., and Leth-Petersen, S. (2014). Measuring the Accuracy of Survey Responses Using Administrative Register Data: Evidence from Denmark. In *Improving the Measurement of Consumer Expenditures*, pages 289–307. University of Chicago Press.

National Academies of Sciences, Engineering, and Medicine (2018). *Reengineering the Census Bureau's Annual Economic Surveys*. The National Academies Press, Washington, DC.

Olafsson, A. and Pagel, M. (2018). The Liquid Hand-to-Mouth: Evidence from Personal Finance Management Software. *The Review of Financial Studies*, 31(11):4398–4446.

Passero, W., Garner, T. I., and McCully, C. (2014). Understanding the Relationship: CE Survey and PCE. In *Improving the Measurement of Consumer Expenditures*, pages 181–203. University of Chicago Press.

Pistaferri, L. (2015). Household consumption: Research questions, measurement issues, and data collection strategies. *Journal of Economic and Social Measurement*, 40(1-4):123–149.

Vavra, J. (2021). Tracking the Pandemic in Real Time: Administrative Micro Data in Business Cycles Enters the Spotlight. *Journal of Economic Perspectives*, 35(3):47–66.

A Transaction Categorization

In this section we provide full details on how we categorize the various transactions available in the BBVA database. Each transaction is given one of 14 labels that comprise twelve two-digit COICOP categories plus non-categorizable consumption plus non-consumption-related.

A.1 Card transactions

For card transactions we manually define a mapping from Merchant Client Codes into labels which is available at `https://www.dropbox.com/s/hroh7azjemtdh5x/mcc_to_coicop.csv`. In several cases, a particular MCC refers to a multi-product retailer with goods that comprise a variety of COICOP categories. For these situations, we first assign a label according to the type of the retailer: 'Supermarkets', 'Supercenters', 'Household Electronics', 'Building Material Supplier' or 'Sporting Goods'.

In order to determine the distribution of products sold by these multi-product establishments, we rely on official statistics. Whenever possible, we use INE's breakdown of turnover according to products sold by retailers [12]. For example, NACE category 4710 is made up of supermarkets and supercenters while household electronics appliances fall under NACE code 4750; thus we can match these retailer labels to their underlying product distribution. Nevertheless, available data on products sold is at a higher aggregation level than ECOICOP categories. For instance, we learn that 72.5% of supermarket and supercenter sales correspond to food, alcoholic beverages and tobacco, products that are broken down into two separate categories in the ECOICOP system. Also, other retailer labels are difficult to match with NACE codes on a one-to-one basis. To fill in the gaps, we resort to the U.S. statistics on retail trade by product lines [13]. This source provides a broader disaggregation of retailers and products, classified based on NAICS and Product/Services Codes, respectively. These statistics allow for a more precise matching between retailer types, e.g. the 'Sporting Goods' label is matched with the NAICS code for 'Specialty-line sporting goods stores'. We manually label the relevant Product/Services Codes with their corresponding ECOICOP categories.

Ultimately, we compile the corresponding product distribution for each retailer label by first relying on INE's breakdown. If no matching retailer category is identified here—such as in the case of 'Sporting Goods'—we fully rely on the 'NAICS to Products' distribution provided in US data. On the other hand, if a retailer is successfully matched to a NACE code in INE's data but a specific product is at a higher aggregation level than the ECOICOP categories—such as the food, alcohol and tobacco in supermarkets and supercenters—we take the corresponding percentage obtained from INE's data and allocate it between ECOICOPs in proportion to the distribution of the relevant categories in the U.S. data.

In the card database we also observe the tax ID (NIF) and sector membership (four digit NACE code) of the selling firm, which will allow us to link across tables as discussed below.

A.2 Direct debits

[12] `https://www.ine.es/jaxi/Tabla.htm?tpx=36388&L=0`

[13] `https://data.census.gov/cedsci/table?q=EC1244&g=0100000US&tid=ECNLINES2012.EC1244SLLS1&hidePreview=false`