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

Gergely Buda
Barcelona School of Economics

Vasco M. Carvalho
University of Cambridge and CEPR

Stephen Hansen
Imperial College and CEPR

Álvaro Ortiz
BBVA Research

Tomasa Rodrigo
BBVA Research

José V. Rodríguez Mora
University of Edinburgh and CEPR

July 20, 2022

Abstract

This paper provides the first proof of concept that naturally occurring transaction data, arising from the decentralized activity of millions of economic agents, can be harnessed to produce national accounts-like objects. We deploy comprehensive transaction-level data and its associated metadata arising from the universe of Spanish retail accounts of Banco Bilbao Vizcaya Argentaria (BBVA). We organize the resulting 3 billion individual transactions by 1.8 million bank customers in a large and highly detailed representative consumption panel to show (i) that the aggregation of such data, once organized according to national accounting principles, 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. Finally, exploiting the panel nature of our data, we (iii) offer a non-parametric analysis of individual consumption dynamics across the consumption distribution.

*We thank seminar participants at the Bank of Canada, CEPR Workshop on New Consumption Data, OECD, the UK Office for National Statistics, the Spanish National Statistical Office (INE) and the Bank of Spain. 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 metadata, 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 policymakers 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, Ehrlich et al. 2022).

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 traditionally structured survey data and slow-moving censuses. In turn, the latter are commonly perceived to be facing both increased funding and, on occasion, political pressures (European Commission 2010, Vinik 2017, AEA Committees on Economic Statistics and Government Relations 2020). Perhaps not surprisingly these difficulties are compounded in a developing country context and result in an underprovision of timely, high-quality data. Strikingly, by 2020, a third of countries worldwide still did not produce quarterly national accounts, with this number rising to 50% in Africa (Silungwe et al. 2022).[1]

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 organized via national accounting rules and then harnessed to produce a large-scale, high-quality and highly-detailed consumption survey. This, in turn, can then be deployed to produce national account objects via simple aggregation. In particular, we show how comprehensive transaction-level from Banco Bilbao Vizcaya Argentaria (BBVA), one of the largest banks in the world, can be organized to (i) reproduce current official statistics on aggregate consumption in the national accounts with a high degree of precision and, (ii) as a result of the richness of the underlying transaction data produce novel, highly detailed distributional accounts for consumption. Additionally, we show that (iii) the panel nature of our data can offer new insights on the nature of individual-level consumption risk and consumption dynamics, overcoming challenges associated with purely cross-sectional or short-panel consumption surveys. 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 of useful coincident indicators or proxies. More generally, our proof-of-concept results imply that transaction data, rather than providing a way to improve measurement within traditional national accounts and consumption surveys, provides a viable alternative.

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 2015Q2 to 2021Q4. Based on this data, the current paper makes four 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

[1] Further, Silungwe et al. (2022) document that (i) half of all countries do not produce quarterly GDP from an expenditure approach; (ii) despite the increasing familiarity of high frequency 'flash estimates' of GDP, only "four economies in Europe and five economies in Asia disseminate quarterly GDP within 30 days after the end of the reference period" and finally (iii) a quarter of countries in the world do not have a household budget survey with which to source data on consumption, with this figure rising to over 50% for African countries.

all cash withdrawals, which we assume are spent as consumption. Leveraging this comprehensive data, we further detail how to exploit metadata 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 spatial 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 largely based on quarterly 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 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, the fact that the aggregates implied by our data closely track both levels and dynamics of published national accounts consumption, immediately implies that our underlying micro-data can be additionally deployed to build distributional national accounts for consumption, characterizing both the distribution of aggregate consumption across Spanish adults - and hence consumption inequality - and its evolution over time - and therefore consumption inequality growth.

Following the work of Piketty et al. (2018), distributional national accounts for *income* already exist for a large number of countries and, arguably, this macro-consistent accounting methodology has had a significant impact in both academic and public discussions surrounding income inequality and its time evolution. Yet, to the best of our knowledge, distributional national accounts for *consumption* are virtually non-existent. To the extent that individual consumption, inequality and its evolution are more welfare-relevant objects than income per se, this is an important gap. [2]

Our third contribution is therefore to construct a detailed distributional accounting exercise for consumption. In particular, we first provide a description of macro-consistent consumption inequality in Spain in 2017 and across a variety of measures. For example, we find that in 2017, 22.4% of aggregate consumption in Spain accrued to the top 10% of the consumption distribution. Further, we benchmark our analysis in two ways. First, we benchmark our results against existing distributional accounts for post-tax income in Spain, concluding that macro-consistent consumption inequality is substantially smaller than its income counterpart (where for example, 31% of total national post-tax income accrues to the top 10%). Second, we benchmark our analysis against consumption inequality as given by the

[2] Arguably, this gap exists precisely because traditional consumption surveys - the typical data source deployed to analyze the extent and evolution of consumption inequality - *are not* consistent with national accounts, as extensively discussed in the literature reviewed below. In the main text we also discuss work by the joint OECD-Eurostat "Expert Group on Disparities in National Accounts Framework". Concurrently to our own work, this joint OECD-Eurostat effort has produced a first set of *experimental* distributional accounts for income and consumption (see Coli et al. 2022), which is only possible under arguably very strong assumptions on how to impute missing consumption in household surveys, both across consumption categories and households.

Spanish Household Budget Survey (Spain's equivalent to the US CEX). We show that the latter not only undershoots aggregate consumption figures (and hence averages) but also displays different properties at the upper tail of the consumption distribution, consistent with undersampling (or under-reporting) of high-income/high-consumption households. Finally, given the rich metadata available to us, we show that it is also possible to break down this distributional analysis of aggregate consumption further, across consumption categories, demographics (age and gender) and time frequencies. In doing so, we show that it is possible to reproduce - and then go beyond - analyses typically pursued in the consumption inequality literature, but this time in a way that is both consistent with the level and evolution of aggregate consumption and with data that is arguably less encumbered by the documented sampling biases of traditional consumption surveys.

Additionally, we present the distributional accounts for consumption growth. Our data spans only the period 2015Q2-2021Q4 and is therefore unable to resolve long term trends in consumption inequality. On the other hand, it does include both the onset of the COVID pandemic and associated lockdowns as well as the subsequent recovery period. We thus provide a macro-consistent account of the evolution of consumption inequality in Spain, in the years before the pandemic, during the large recession in 2020 and over the period of strong recovery in 2021. We find that consumption inequality was relatively stable in the three years before the pandemic, decreased markedly during the first year of pandemic and then increased strongly during 2021. Further, we show that this pattern is consistent with the decline and subsequent recovery in luxury and Veblen goods consumption, which affected disproportionately more those at the top of the consumption distribution.

Finally, our fourth contribution is to use the same data to analyze the characteristics of the micro-structure of consumption dynamics. We use the panel dimension of individual data to analyze its stochastic structure, documenting not only strong mean reversion but also the lumpy nature of consumption growth at the individual level. We find that micro-level consumption growth is difficult to approximate with a Gaussian distribution: for Spanish adults at both the top and the bottom of the consumption distribution (and particularly for older adults) consumption growth presents a high degree of skewness (positive for those at the bottom, and negative for those at the top) and excess kurtosis, indicating thick tails. Thus, the unprecedented size and detail of our dataset allow us to show not only that the distribution of consumption has a thick tail, but also that consumption *growth* has a thick tails in both the right and the left sides of the distribution, a surprising result difficult to reconcile with consumption smoothing at an individual level.

Our paper relates to five distinct literatures. First, our work is related to a small literature reviewing current 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, Ehrlich et al. 2022). 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 with increasing 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 national accounts grade 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), and 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), or Chetty et al. (2020), and Baker (2018) and 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 and retail point of sales data alone, we expand the scope of consumption expenditure significantly, by additionally considering direct debits and regular transfers, 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 or retail sales. Rather, differently from this literature, it is to offer evidence that national accounts can be based on such data.

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), and 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 frequency. 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.

Fourth, our paper builds on the literature on distributional accounts which, albeit stretching back to the pioneering work of Kuznets and others, it was not practically developed until the recent work of Piketty et al. (2018) and Alvaredo et al. (2021). This, in turn, has generated interest within national statistical agencies and international organizations, and provided the impetus for routine production and dissemination of such accounts (see for instance Statistics and Data Directorate of the OECD (2020)). The procedure for creating distributional accounts typically consists in generating synthetic data out of multiple sources of information (tax returns and tabulations, surveys to consumers and producers, and the underlying information in the generation of traditional national accounts to which the synthetic data is forced to aggregate). Our contribution to this literature is to build an encompassing rich and very dense set of distributional accounts, from a single source of naturally occurring data; without the need of generating synthetic data and adding, we believe, clarity and simplicity to the procedure.

Finally, our paper is also related to the large literature analyzing the extent of inequality in consumption (see, for example, Attanasio et al. (2014), Attanasio and Pistaferri (2016), Aguiar and Bils (2015), Coibion et al. (2021), Krueger et al. (2010)) and consumption dynamics and its relation to income dynamics (as in Madera (2019)). In particular, we adapt part of the analysis of the dynamics of individual income recently developed by Guvenen et al. (2021), using it to analyze the year-to-year changes in consumption at an individual level.

The paper is organized as follows. In section 2 we present our data and explain in detail the procedure that we use to move from raw transaction data to consumption spending. In section 3 we demonstrate that the aggregation of our data reproduces the measures of Consumption in National Accounts, and its distribution across consumption categories moves on par with the surveys performed by the Spanish National Statistics. In section 4 we present our distributional national accounts for consumption, and look at how they behave under different cuts of the data (by categories of consumption, gender, age, frequency of aggregation). We also study how this distribution evolves over time. Finally, in section 5 we use the microstructure of our data to offer a non-parametric characterization of consumption dynamics.

2 Building a Naturally Occurring Consumption Survey

The fundamental object we build using BBVA data is a consumption survey. We do so by accessing the universe of BBVA retail financial accounts in Spain beginning in 2015Q2. Unlike previous papers that have 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.[3]

It is first instructive to review how INE builds its own annual Household Budget Survey (HBS) for measuring individual and household consumption across different demographic groups and product and service types. In the HBS, households form the basic units of analysis and are chosen for participation according to a well-defined sampling procedure. INE first defines a set of 2,275 census tracts based on municipality size, employment, age, education and other socioeconomic characteristics. Within these tracts, ten dwellings are randomly selected and all households within them are invited to participate.[4] Sampled households take part in the HBS for two years, with a staggered rotation in which half the sample is replaced every year.

Households record their spending during a two-week period in standardized notebooks. Each purchase is assigned a classification based on the five-digit *Classification of Individual Consumption by Purpose* (COICOP) system. Quantities and prices are also recorded. Following this, households are interviewed by INE about items purchased at lower than two-week frequency. Recurring payments are estimated by the amount of the most recently issued bill. For households who own their homes, INE imputes the consumption value of housing services using information on house size and local rental prices in addition to subjective estimates of the respondents.

At a broad level, then, there are two problems one must address in converting naturally occurring spending data into a consumption survey. The first is sample definition: the total client pool of any single bank may not be representative of the national population. The second is that consumer spending does not equate to consumption. Spending on items outside the COICOP classification system should not be included in consumption measures, yet raw spending data from a bank typically has no such classification. Furthermore, housing services are an important part of total consumption but are not directly observed in spending data by those who own and occupy their homes.

In the remainder of this section, we describe how we approach these issues. We first detail the underlying sample of BBVA clients for whom we compute consumption measures, then how we map spending into consumption categories. Finally, we explain approaches to aggregating individual consumption to form representative national statistics.

[3]The data is held in a secure internal cloud environment that only BBVA employees and a limited number of non-BBVA individuals can access. All our database operations are GDPR compliant and have received additional legal approval from BBVA prior to execution.

[4]A dwelling is a single housing unit. For the purposes of the survey, a household is a group of people who at least partially share expenses. Household members need not have a family relationship. Multiple households can live within the same dwelling, for example renters who share no bills.

2.1 Sample frame

In total there are 10,270,041 unique BBVA retail customers who conduct at least one consumption-related transaction (explained in next subsection) between 2015Q2 and 2021Q4, when we end our sample.[5] We observe irregular volatility in consumption in some of the months between 2015Q2 and 2016Q4 while from 2017 on the data appears to have high quality. Information available about customers includes age, sex, and street address. 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. We define a balanced panel of *active* customers who make at least ten consumption-related transactions in each quarter. There are 1,827,866 such customers, after removing 181,918 self-employed customers whose transactions might reflect production inputs instead of consumption.[6]

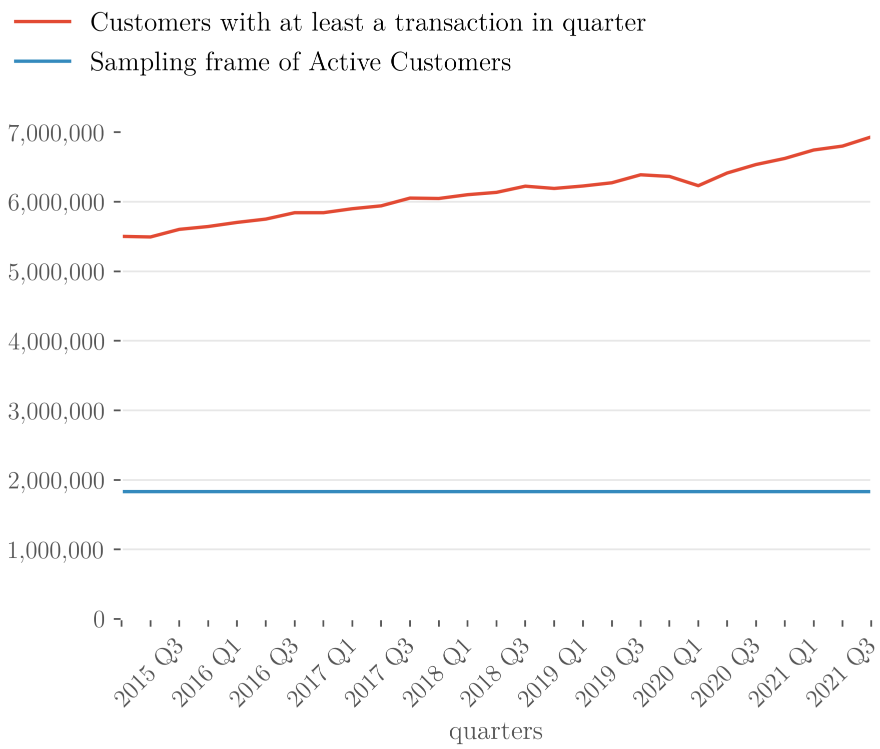


Figure 1: Frame of Active Customers

The red curve is a quarterly time series of the number of unique retail banking clients of BBVA in Spain who make at least one consumption-related transaction from 2015Q2 through 2021Q4. Active customers are those who make at least ten consumption-related transactions in each quarter. There are 1,827,866 in our final sample.

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. The 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.[7]

Figure 2 compares the distribution of geographic location of residence, age, gender, and neighborhood income for the population of active clients against those for all Spaniards as recorded by the census.[8]

[5]BBVA transaction data is updated daily, so in principle all our constructions can be conducted in near-real time.

[6]The 1,827,866 total is also after dropping a small number of outliers the procedure for which we describe below.

[7]In line with the HBS, 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. Moreover, the panel structure allows one to analyze within-client consumption changes over longer time horizons.

[8]We use the census year 2018 for the comparison. The geographic distribution is reported at the level of the 19 primary

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. They are also more likely to live in higher-income neighborhoods. When we come to form aggregate consumption measures, we address these imbalances appropriately.

Active customers' household structure is important for determining their consumption, but is not directly observed in the data. Instead we infer it by linking each active customer to the set of other 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.[9] 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 1,589,280 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 record the number of additional household members as equal to the number of dependent adults.

Figure 3 compares the resulting distribution of household sizes according to our grouping procedure against official data. 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.

	HBS 2016	**HBS 2017**	**HBS 2018**	**HBS 2019**	**HBS 2020**	**BBVA Sample**
Households	22,011	22,043	21,395	20,817	19,170	**1,589,280**
Adults	47,420	47,055	45,328	43,988	40,285	**1,827,866**

Table 1: Number of Households and Adults in HBS by Year

Our sampling frame consists of 1,827,866 active BBVA customers (who together form 1,589,280 households) whose consumption spending we observe from 2015-2021. For comparison, this table records the number of households and adults in INE's official Household Budget Survey spending by year.

In summary, then, our sample frame consists of a balanced sample of individuals along with their household structure. Table 1 tabulates the number of Spanish adults and households who participated in the Household Budget Survey during the years we track active customers' consumption spending. Naturally occurring data permits an enormous expansion of the number of individuals whose consumption can be recorded, which in turn allows for much finer cuts of data. Moreover, since no household participates in the HBS for more than two years, one cannot track spending over time within individuals. This is another advantage of our data that we illustrate below.[10]

2.2 From spending to consumption

The next challenge is to convert individual spending data into individual consumption data. For non-housing consumption, our overall strategy is to use transaction metadata to classify individual purchases as either consumption- or non-consumption-related and, if the former, to assign a COICOP classification. For housing consumption, we estimate a simple regression model that predicts observed rental

regions in Spain known as Comunidades Autónomas. For age, we plot the conditional census distribution for those at least 18 years old since the BBVA sample contains no minors. Details of neighborhood income are provided in figure 2 notes.

[9] Postal codes are coarser than census tracts. We do not match on census tracts to avoid privacy violations.

[10] Unfortunately, the publicly accessible data from HBS does not allow to use the panel component of the survey to study consumption dynamics, as the household identification code are redrawn every year.

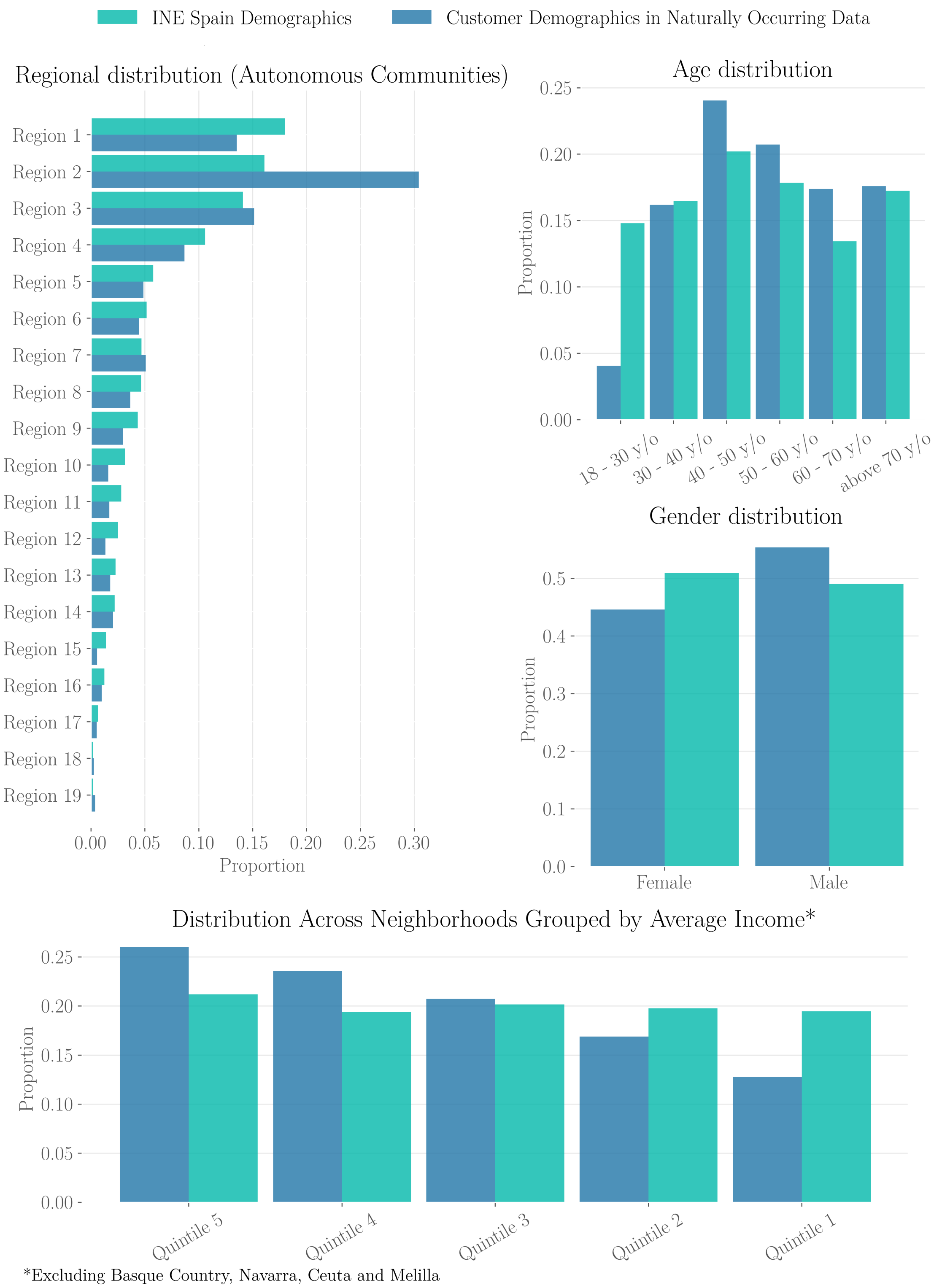


Figure 2: Demographics of Active Customers

For each BBVA customer, we observe data on age, gender, and address. The top three bar charts compare the distribution of Active Customers' characteristics against Spanish census data in 2018. 'Region' refers to a *Comunidad Autónoma* in Spanish terminology. To construct the neighborhood income distribution, we use information that INE provides about the average income of residents of each census tract in Spain (36,581 in total). This information exists for all Regions except for the Basque Country and Navarra. Ceuta and Melilla are small enclaves that are not subdivided into census tracts. On the basis of average income, we group census tracts into quintiles within each of 50 Spanish provinces and plot the distribution of the Spanish population across them. Because census tracts are approximately uniform in population, this distribution is close to uniform. We then assign Active Customers to census tracts based on street address and plot the distribution across quintiles.

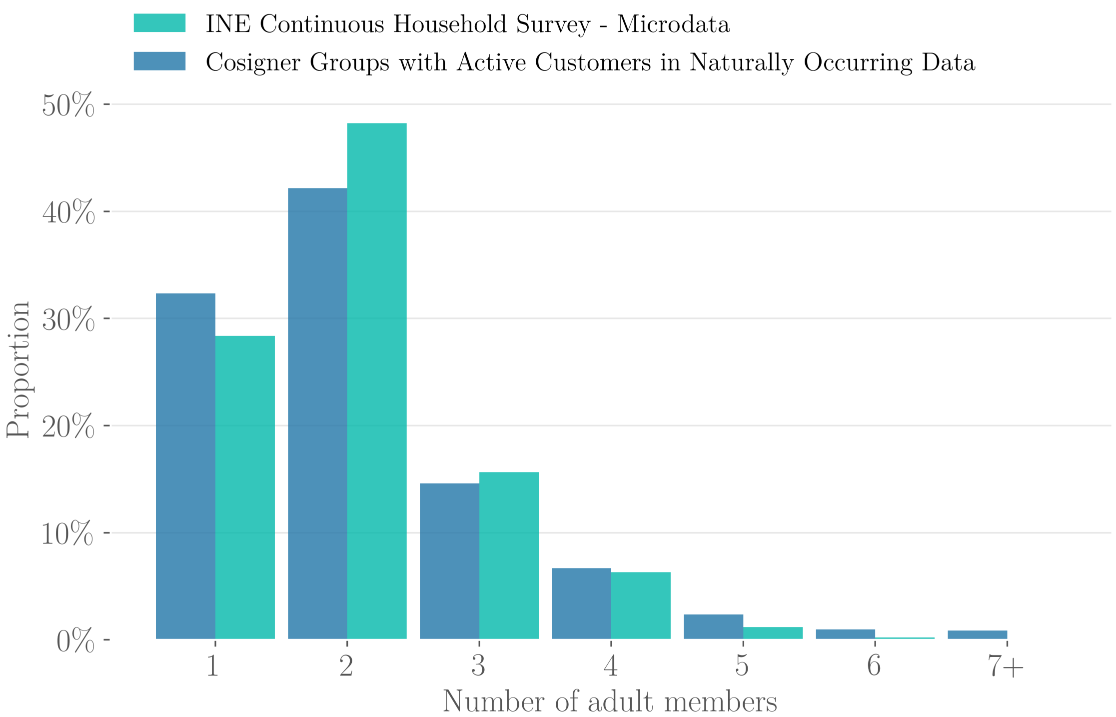


Figure 3: Household Proxy vs Official Data

To form the official distribution of household sizes, we use INE's *Continuous Household Survey* (`https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736176952&menu=resultados&idp=1254735572981`) and extract from each surveyed household the number of adults. We focus on adults since no minors are BBVA customers.

payments from household characteristics, and then use it to impute monthly rental payments for all active customers. Appendix A contains full details of the construction of consumption, with the main text providing a broad summary.

Throughout, we closely follow national accounting principles from the European System of Accounts (2010) as detailed in appendix B. The goal is to implement as faithfully as possible the official definition of consumption on naturally occurring data. We also seek to do so in a transparent way, so that the choices can be replicated (and further explored) using similar data from other banks and other countries.

2.2.1 Non-housing consumption

The national accounts concept we seek to construct is *final consumption expenditure* which according to section 3.94 of the European System of National and Regional Accounts (ESA 2010):

> ...consists of expenditure incurred by resident institutional units on goods or services that are used for the direct satisfaction of individual needs or wants or the collective needs of members of the community.

Many bank account outflows—for example, transfers to investment institutions, tax payments, or major building work—do not meet this definition. Appendix B lists all items that should and should not be included in consumption according to the ESA, and how we use transaction metadata to design appropriate filters. While in some cases we cannot exactly replicate national accounts principles, in general we are able to do so. After determining which transactions belong in consumption, we attribute wherever possible a COICOP classification at the two-digit level, which table 2 displays.

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

Table 2: COICOP Consumption Categories (Two-Digit)

This table displays the 12 COICOP categories we use for classifying consumption transactions. In line with INE, We use the European COICOP system in place of the international COICOP system. 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.

There are three separate transaction modes in the data—card spending, direct debits, and irregular bank transfers—and each has a distinct form of metadata we use for categorization. For cards, the relevant information is the Merchant Client Code (MCC) of the counterparty firm, which is a standardized system for classifying business activities. We manually categorize MCCs and make this resource publicly available at `https://www.dropbox.com/s/hroh7azjemtdh5x/mcc_to_coicop.csv`. There are 835 MCCs in total which allows for a fine-grained separation of spending categories. For example, there are separate MCCs for charity donations and tax payments, which do not form part of consumption, as well as an array of other non-consumption spending codes. Most other MCCs are specific enough to allow a clear mapping to a COICOP with two exceptions. First, there are MCCs that relate to generic consumption. The most prominent example is cash withdrawals at ATMs. Other examples include the MCCs for *shopping clubs* and *onboard sales (sea)*. Second, a limited number of MCCs refer to the sales of multi-product retailers such as supermarkets. In these cases, we use published statistics on the distribution of sales across COICOP categories by sector to allocate shares of a transaction's value.

Direct debit transactions are assigned one of approximately 100 labels by an internal BBVA classification system. These include labels like *utility bill payment*; *council tax payment*; and also more generic categories. Like for MCCs, we manually classify the labels but do not provide a public file since they are proprietary. If a label is not clearly categorizable, we instead attempt to link the counterparty firm's tax ID to the card transaction table and use the MCC mapping. If this fails to produce a classification, we instead use the firm's four-digit NACE sector code, which we again manually map to categories. We make available the NACE-to-COICOP mapping at `https://www.dropbox.com/s/9lcab2zajijxltn/nace_to_coicop.csv`.

Transfers contain the least relevant metadata, and even determining whether the counterparty is a firm requires care. In many cases, the only option is to string-match the counterparty name as recorded in the transaction metadata with firm names from commercial registry data external to BBVA. Conditional on identifying the counterparty as a firms, we categorize transfer-based transactions using our manual mapping above (full details in appendix A).

Table 3 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

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

Table 3: Consumption data volume of Active Customers (whole period)

This table displays the total number and value of consumption-related transactions made by the sample of 1,827,866 BBVA active customers from 2015-2021. These are broken down by transaction mode, where cash withdrawals—which we treat as consumption—are separated out.

from other transactions,[11] 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.

Figure 4 shows the distribution across consumption categories by payment mode. 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.

Finally, we remove active customers from the sample whose non-housing consumption is high relative to their census tract average income. Appendix A.2 details the procedure. This ensures that the properties of the consumption distributions analyzed below are not driven by outliers.

2.2.2 Housing consumption

Building a predictive model for rent begins with the extraction of rental payments, which we identify using a free-text field that payees can populate to describe both direct debits and generic transfers. 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.

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 then sum up all rental payments at the household-month level to form units of observation. 437,307 households have at least one rental payment. 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 (which proxies housing quality), utility payments (which proxies house size), 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 4 provides summary statistics for household-level observables

[11] 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, which appear as transfers. The former is more prevalent in our data, especially in later years.

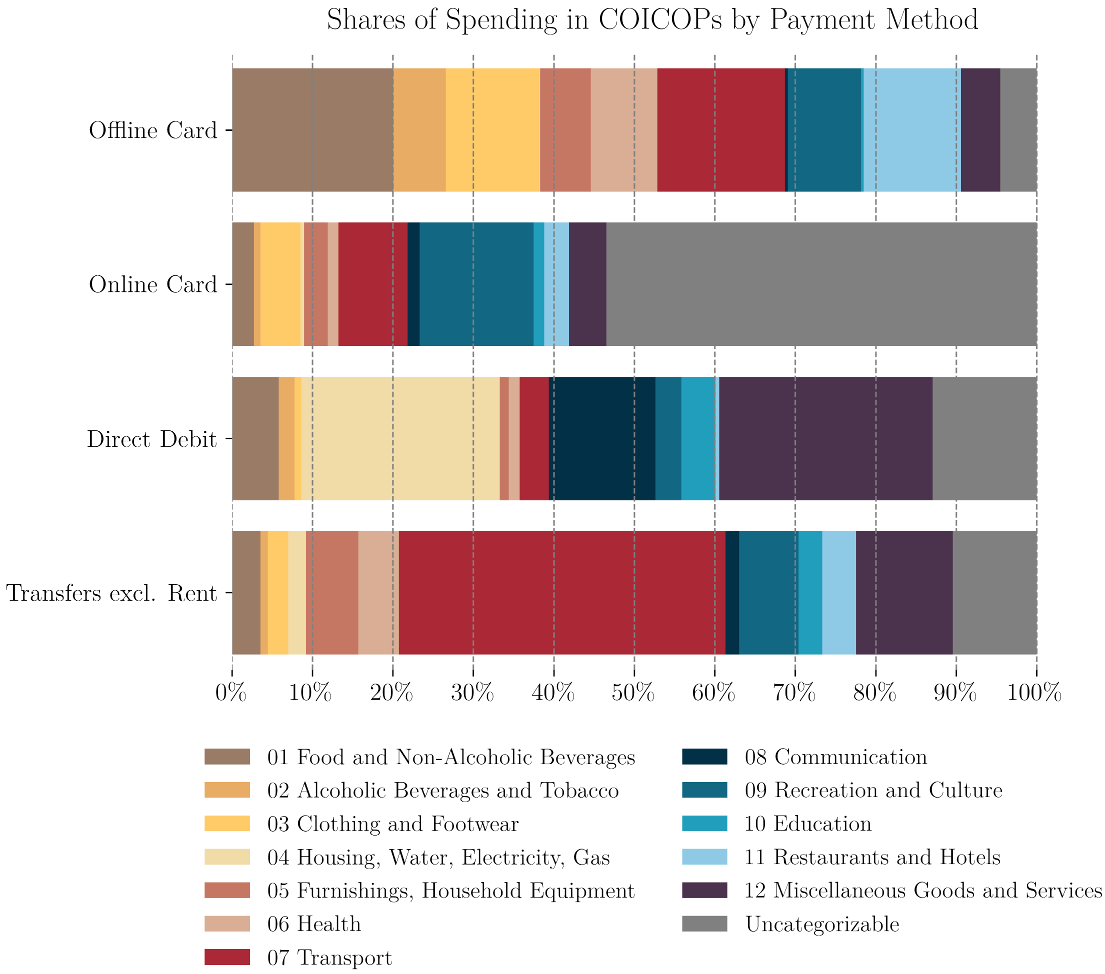


Figure 4: COICOP Shares by Payment Method

This figure shows the percentage of the total value of consumption spending across transaction mode that we manually allocated to COICOP categories.

in this set.

	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 €

Table 4: Summary Statistics of Training Sample for Rent Regression

We estimate a rental regression model using a subsample of 16,977 households that 1) have observed rental payments in 70 out of the 81 months in the sample; 2) have at least one month of observed utility payments; 3) have at least one month of labor, benefit, or pension income recorded internally by BBVA. This table provides summary statistics for these three variables for this sample.

For geographic location, we seek to define spatial units that are sufficiently well populated with households that fixed effects can be reliably estimated. Appendix A.3 details the specific approach, which consolidates postal codes together until a minimum of 30 households has been reached. The procedure produces 327 spatial units out of 2,687 unique postal codes in the observation sample. The average number of households and postal codes in each unit is 52.0 and 8.2, respectively.

Finally we regress monthly rental payments by household on spatial unit fixed effects, income, and utility payments via ordinary least squares. Where no income or utility information is available to form a given month's record, we use the household average over all months.

Table 5 displays the results. Although simple, the 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 shifts rental payments by 70 euros a month, or 0.28 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.

Variable	Model	Test set
Spending on House Utilities	0.0884	
	(0.0008)	
Income	0.0362	
	(0.0011)	
N of Households	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

Table 5: Regression for Rent

We regress household-level rental payments each month on spatial unit fixed effects, three-month total utility spending, and six-month average income. The 'Model' column contains point estimates obtained from OLS for the latter two, with standard errors in parentheses. The 'Test set' column provides goodness-of-fit information for the etimated model's performance out-of-sample. The test set is formed by drawing a random month from the rental payments of households with between 50 and 70 monthly rental payments.

We then use our estimated rental regression to impute monthly rents to all households. Where a household lies outside the spatial units defined for the estimation sample, we assign it to the closest unit based on centroid distance. Where no income or utility information is available to form a given month's record, we use the household average over all months. If a household has no utility or income records at all, we assign the spatial unit 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 compared to the estimation sample, which suggests 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.

2.2.3 Spending vs consumption

Finally, we compare how raw BBVA spending compares to consumption defined by our filters motivated by national accounting principles. Table 6 tabulates various quantities of interest for 2019. The value of total active customer account outflows across the three payment modes is over 50 billion euros. 20 billion euros flow to counterparties we remove from the data due to their not being relevant to consumption (mainly private individuals). Of the 31 billion euros of remaining value, another 7 billion is removed via our manually built filters applied to transaction metadata. Finally, the total value of imputed rent—which is not observed in spending data—is 8 billion.

	Volume of Transactions (2019)
(1) Cash Withdrawal	10,142 million €
Outflows under concepts of Card / Transfer / Direct Debit	51,342 million €
Out of which: Outflows to organizations	31,232 million €
(2) Out of which: Consumption-related transactions	23,959 million €
(3) Imputed Rent	8,158 million €
Total unweighted consumption (1) + (2) + (3)	42,259 million €

Table 6: Impact of filtering spending to consumption (2019)

This table tabulates total spending across transation modes in 2019, the spending allocated to consumption, and the value of imputed housing services.

In short, there is a large distance between raw spending and consumption. This suggests that using the former as a proxy for the latter in the absence of appropriate metadata is likely to produce a poor approximation.

2.3 Household and demographic weighting

The above operations yield a non-housing consumption measure c_i^{NH} for each active customer i, and a housing services consumption measure $c_{h(i)}^{\mathrm{H}}$ defined at the level of the household h. The final step in producing aggregate consumption measures is to adjust these totals for housing 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. Let $A(i)$ ($O(i)$) be the set of active (other) customers in i's household including himself. Household-weighted consumption is:

$$c_i = \frac{\sum_{j \in A(i)} c_j^{\mathrm{NH}} + c_{h(i)}^{\mathrm{H}}}{|A(i)| + 0.5|O(i)|}. \tag{1}$$

Suppose first that a household is made up just of active customers. (1) then aggregates all members' spending and divides it equally. If the household also contains non-active members, we apply an additional down-weighting that treats each non-active customer as 0.5 of an active customer. Non-active customers share the consumption of active customers, but also potentially generate consumption spending outside the BBVA universe. The down-weighting by $0.5|O(i)|$ accounts for these competing forces.

In much of analysis below, we aggregate individual spending into larger units and produce time series. To do this, we define cells at the gender (g), age group (a) and neighborhood income quintile (q) levels.[12] Since the neighbourhood quintiles are formed separately for each region (Comunidad Autónoma), the latter variable ensures regional representativity as well. Let $c_{t;g,a,q}$ be the sum over (1) for all active customers in cell (g, a, q) computed at time t. Depending on the setting, t might be yearly, quarterly, monthly, etc.

To aggregate across cells in each time period t, we account for demographic imbalances between the active customer sample and Spanish census data from 2018. Let $x^{\text{INE}}_{g,a,q}$ be the total count of Spanish adults according to census data in cell (g, a, q) in 2018. Also, let $x^{\text{BBVA}}_{\tau(t);g,a,q}$ be the total count of active customers in cell (g, a, q) in year τ which depends on the time period of interest t. Total consumption in each cell at time t is[13]

$$c^{W}_{t;g,a,q} \equiv c_{t;g,a,q} \times \left(\frac{x^{\text{INE}}_{g,a,q}}{x^{\text{BBVA}}_{\tau(t);g,a,q}} \right) \tag{2}$$

From here one can form arbitrary data aggregates by summing over the cells. Aggregate consumption is the sum over all cells; regional neighbourhood quartile consumption for q is the sum over all gender and age categories holding q fixed; and so on. Category-specific consumption is derived by only considering the subset of $c_{t;g,a,q}$ that pertains to the COICOP of interest.

While this weighting procedure accounts for demographic imbalances in forming aggregate consumption measures, it does not produce a nationally representative sample of individuals' consumption. In the analysis below for distributional national accounts and consumption dynamics, this is what we need. To obtain it, for each demographic cell (g, a, q) in year τ we draw with replacement from the population of active customers $x^{\text{INE}}_{g,a,q}$ times. This produces a national sample of size equal to the Spanish adult population for which one can perform distributional analysis. A full bootstrap procedure would compute consumption distributions across multiple national samples, but in practice we find little variance across draws. To avoid the computational cost of the full procedure, we proceed with a single national sample.

3 Measuring Aggregate Consumption

The remainder of the paper explores properties and develops applications of the naturally occurring consumption survey. This section focuses on aggregate consumption measures derived from $c^{W}_{t;g,a,q}$. It begins by summing across all cells at quarterly frequency to form a national aggregate consumption measure and compares this to the official Household Final Consumption published by INE. It also explores how the breakdown of consumption across COICOP categories compares across naturally occurring and official data. There is a surprisingly tight correspondence between our measures and official figures in both cases, in spite of markedly different approaches to arriving at aggregate numbers.

Next, the section presents novel aggregate objects not present in official data: a daily consumption series; aggregate household consumption at provincial level; a high-frequency measure of the national

[12] Recall that neighborhood income quintiles are formed within all regions using data on average income within census tracts. Two exceptions are the small enclaves of Ceuta and Melilla, which have too few census tracts to make the division into quintiles reasonable. Accordingly, we do not form income groups within them. The other two exceptions are the Basque Country and Navarra, where average income data is sparsely reported by INE. Instead, we use census data that maps each census tract in provinces within these regions into urban, semi-urban, and rural categories. We use these categorizations in place of neighborhood income quantiles for these provinces.

[13] In principle one could use year τ census data, but $x^{\text{INE}}_{g,a,q}$ is quite stable over time so we avoid the computational cost by using a single reference year 2018.

household consumption basket; and a breakdown of consumption by mode-of-payment. These illustrate the various ways that naturally occurring data can aid in generating new, policy-relevant measures.

3.1 Relationship to National Accounts

Many datasets inform the estimation of household consumption in national accounts, including firm sales measured by survey instruments and obligatory reporting requirements; administrative data on, for example, car purchases; and household surveys for specific purchases, for example the Food Consumption Panel. These various data sources are then combined in a statistical model.[14] INE publishes household final consumption at quarterly frequency, and reports COICOP breakdowns at annual frequency. The Household Budget Survey is a raw input into the construction of published figures, but does not mechanically equate to it at annual frequency.

A natural question of interest is how national accounts consumption compares to our measure. Figure 5 plots INE's quarterly aggregate series against ours both in levels and in quarter-on-quarter growth rates. Seasonal effects in both time series are removed in line with Eurostat practice.

(a) Levels

(b) Quarter-on-Quarter Growth Rates

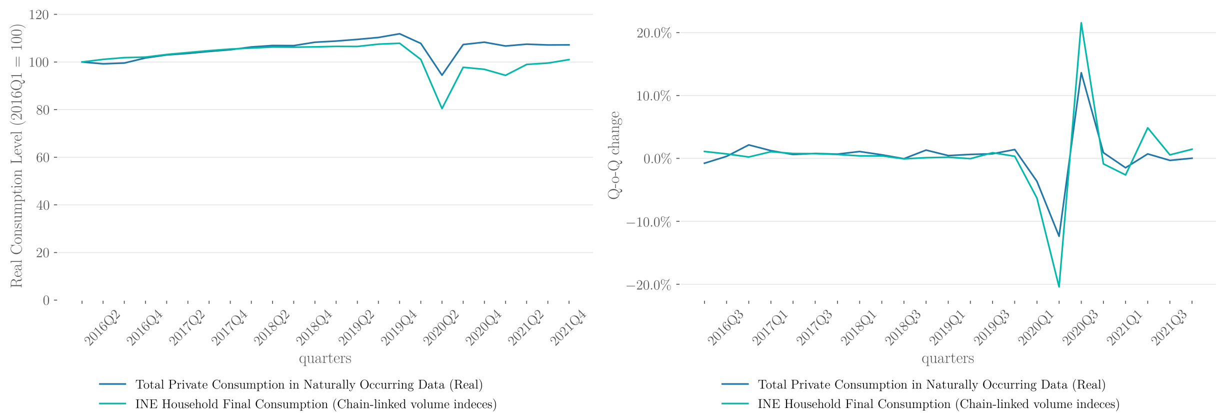
Figure 5: Aggregate Naturally Occurring Consumption vs. National Accounts

These figures compare quarterly aggregate household consumption according to official INE data and to naturally occurring data. To seasonally adjust both series, we use the Jdemetra+ application and apply X-13ARIMA-SEATS. The plot on the left shows the total level of consumption. The plot on the right dispalys the growth rate in aggregate consumption from quarter $t-1$ to quarter t.

The striking result is that naturally occurring and official data align exceedingly well in both levels and growth rates at quarterly frequency. This is despite their quite different constructions. While we carefully follow national accounting definitions in designing which transactions to filter, we obtain aggregate consumption by a simple weighted summation over individual consumption measures. On the other hand, the ESA (which INE follows to produce national accounts) defines a complex set of procedures for combining multiple data sources into national accounts aggregates. Indeed, one virtue of our approach is its simplicity: our philosophy is to design an accurate individual consumption survey which can then aggregate directly to the national level. The results suggest this transparent and straightforward approach is a close approximation to the output of INE's statistical model applied to noisier individual datasets.

One notable discrepancy in the series occurs during the COVID-19 crisis during which both the fall and recovery in consumption is more stable under our measure than under official data. Without

[14] A full description of the quarterly national accounts system is available at `https://ec.europa.eu/eurostat/documents/3859598/5936013/KS-GQ-13-004-EN.PDF/3544793c-0bde-4381-a7ad-a5cfe5d8c8d0`.



(a) Levels **(b)** Quarter-on-Quarter Growth Rates

Figure 6: Aggregate Naturally Occurring Consumption vs. National Accounts (Real)

Compared to the nominal consumption series plotted in Figure 5, this figure shows consumption in real terms. We deflated the nominal series by the official Consumer Price Index defined at the month-region-COICOP level. We plot the resulting series in terms of level with the base quarter being 2016Q1 (LHS) and in terms of quarter-on-quarter growth rates (RHS).

a ground truth to compare against, it is difficult to determine which of these series is more accurate. One hypothesis is that official data uses as an input firm sales but that during and after the pandemic consumers began to purchase inputs from different firms than those included in INE's model. This would imply that we would capture more relevant consumption and so generate a higher level.

Finally, it is very easy to create series of real consumption, controlling for price levels. Our data is by nature nominal as, obviously, in the transaction ledger there is no field indicating the price involved, but INE publishes a monthly price index series disaggregated by cells of autonomous community and COICOP category.[15] Thus, it is straight forward to create monthly series of real consumption at COICOP and autonomous community level.

In Figure 6, we plot the real analogues of the curves in figure 5. Unsurprisingly, we observe a tight correspondence between official real consumption and that derived from naturally occurring data.[16]

It is also instructive to compare how aggregate consumption as measured by the Household Budget Survey tracks national accounts. Figure 7 plots these at a common annual frequency. Consistent with the existing literature (Attanasio et al. 2014, Barrett et al. 2014, Passero et al. 2014), the HBS understates national accounts consumption, presumably due to household underreporting. In contrast, naturally occurring data captures all consumption spending within a given payments system, which ameliorates such mismeasurement and leads to more accurate aggregates.

We next compare the distribution of aggregate consumption across COICOP categories according to national accounts, the HBS, and naturally occurring data in 2019. In the latter, we distribute cash across COICOP categories using the same shares as we observe for offline card spending. The assumption is that cash and offline card spending are substitutes and so should be spent on related items. Figure 8 compares the levels of consumption in log space of the different measures.

In general, there is a strong relationship between national accounts COICOP-specific consumption levels and those of the HBS and naturally occurring data. Overall, though, naturally occurring data

[15] Available at https://www.ine.es/dynt3/inebase/es/index.htm?padre=8423capsel=8428

[16] We prefer to keep the paper in the realm of nominal data, as it is what it is directly observed in the transaction record, but we want to remark that using the price data from INE it is also straight-forward to evaluate the cost of inflation to different groups of individuals. Given that different income groups have different distributions of consumption across COICOPS, which we can measure as it will be explained below, it is immediate to evaluate the cost for different groups of the changes in prices that take place across all COICOPS and regions, for any given period of time.

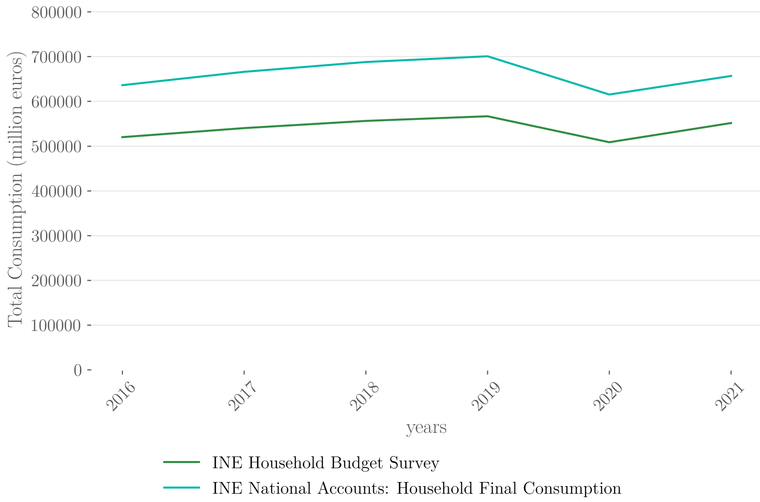


Figure 7: Aggregate Consumption: National Accounts and Household Budget Survey

This figure plots aggregate consumption at annual frequency as measured by the Household Budget Survey and national accounts.

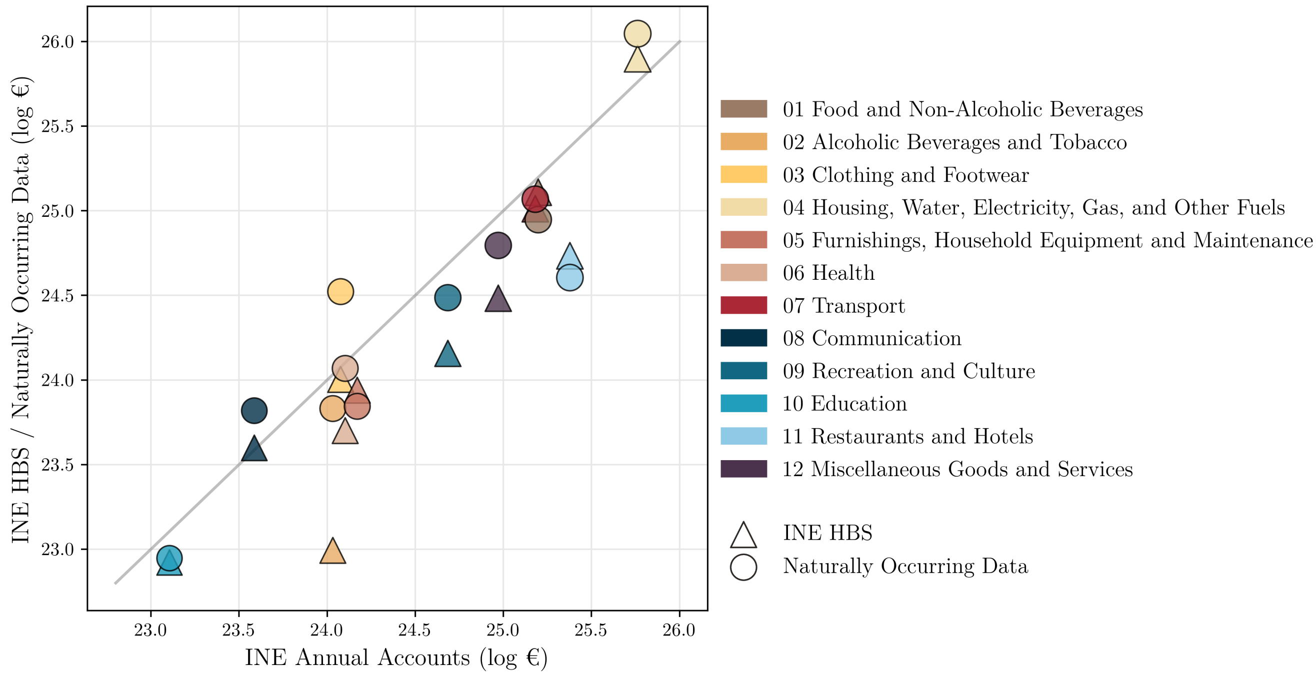


Figure 8: Distribution of Spending across COICOP Categories

This figure compares the (log) levels of consumption across COICOP categories from Spanish national accounts (x-axis) and from the Household Budget Survey and naturally occurring data, respectively (y-axis). For the naturally occurring data, we distribute cash across COICOP categories using offline card spending shares. In national accounts, COICOP spending is reported for sales of all goods in Spain. This includes sales of goods to foreigners and excludes non-domestic spending of Spanish residents, while final household consumption excludes the former and includes the latter. For this reason, we subtract sales of goods to foreigners and add non-domestic spending of Spanish residents in the same proportion as reported COICOP levels.

achieves better coverage of national accounts: the average absolute error with respect to national accounts of naturally occurring data (HBS) is 0.266 (0.333) log points across all COICOP categories. This is despite a portion of naturally occurring consumption having no COICOP assignment, which creates a downward bias in the level of any particular COICOP category. Interestingly, the largest difference appears for COICOP 2 'Alcoholic Beverages and Tobacco'. Households appear especially likely to underreport spending on this category due to social stigma, while naturally occurring data lines up with national accounts. The largest divergence between naturally occurring data and national accounts is for COICOP 3 'Clothing and Footwear'. This is likely due to the allocation of cash in proportion to offline card spending, where this COICOP is over-represented (see figure 4). Finally, it is reassuring that housing services—a category that is largely imputed not directly observed—line up in all three measures even though each has a distinct approach to imputation.

(a) Levels

(b) Quarter-on-Quarter Growth Rates

Figure 9: Aggregate Naturally Occurring Consumption vs. National Accounts (Card Only)

This figure replicates figure 5 but only using the portion of naturally occurring consumption spending arising from card transactions, including cash withdrawal with card.

Finally, we consider an aggregate consumption measure derived from card spending alone. Card spending is one of the most widely available forms of financial transaction data, and has recently been used to track the effects of COVID-19 in several papers (Andersen et al. 2020, Carvalho et al. 2021, Vavra 2021). A natural question is then: how far can one go in accounting for patterns in aggregate consumption using only card spending filtered through an appropriate MCC mapping? Figure 9 plots the same curves as in figure 5 but only using the part of consumption arising from card spending, including cash withdrawal with card.

The card series has poor aggregate coverage of national accounts consumption. Moreover, its growth rate has a notable upward bias not present in the full naturally occurring consumption measure. The average quarter-on-quarter growth rate of official household consumption from 2016-2021 is 0.55%; the growth rate of the full naturally occurring measure is 0.83%; and the growth rate of the card-based naturally occurring measure is 2.01%. Even as a coincident indicator, then, card spending has notable limitations at least in this setting.

3.2 Novel aggregate objects

The results above show that one can go far in recovering official statistics from the aggregated, naturally occurring consumption survey. One of the main advantages of this survey, though, is its ability to produce novel national accounting measures that go beyond what is already available from statistics agencies. We conclude this section by providing illustrations of this idea across several dimensions.

3.2.1 National accounts at high frequency

Particularly in the wake of large macroeconomic shocks, it is important to understand how the economy evolves at high frequency. The definition of t in $c^W_{t;g,a,q}$ depends on the user, and can be adjusted to whatever frequency one needs. To illustrate a high-frequency version of aggregate consumption, figure 10 takes t to be a single day. Housing is imputed at the monthly level, so we divide it equally across all days in a given month. Moreover, to avoid large payments on a single day driving the measure (e.g. regular dates on which bills are paid), we compute a 28-day, non-centered moving average. Finally, we account for daily seasonality by plotting year-on-year growth rates with respect to comparable days (e.g. the growth of consumption from the first Sunday of year $t-1$ to the first Sunday of year t). As one might expect, the daily series is more volatile – partly because of data quality issues in 2015Q2-2016Q4 – than the more traditional quarterly series, but is also able to capture the economic impact of large shocks. The drop in aggregate consumption due to COVID-19 lockdowns is stark and immediate.

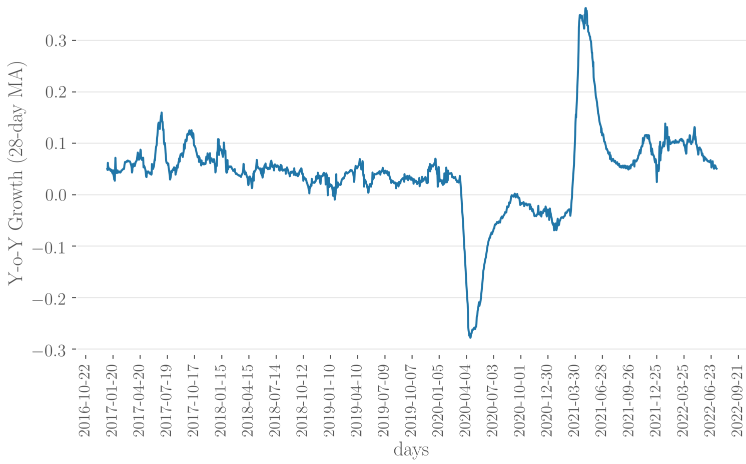


Figure 10: Daily Consumption Growth (Year-on-Year, 28-Day Moving Average)

This figure plots aggregate consumption growth at daily frequency. Monthly imputed housing services are divided equally across days within months. Aggregate consumption is smoothed using a 28-day, non-centered moving average. Year-on-year growth rates are then computed using comparable days-of-the-week in years $t-1$ and t.

Another high-frequency measure of interest is the consumption basket of consumers, for example to compute inflation rates in environments with unstable shopping patterns. Figure 11 shows the evolution of consumption shares of COICOP categories at monthly frequency, where again one observes a dramatic shift due to COVID-19. Spending on restaurants and hotels collapses in 2020, while other categories such as communication remain relatively stable throughout the sample.

3.2.2 Geographically detailed national accounts

In Spain, as in other countries, an important policy challenge is regional economic imbalances generated by the growth of a handful of economically dynamic cities and the relative decline of more marginal areas. Rural depopulation has been a major concern for the last decade. At the same time, the rigorous assessment of subnational geographic inequality is hampered by the lack of regional national statistics. In Spain INE produces estimates of GDP at the regional level but without separating out consumption

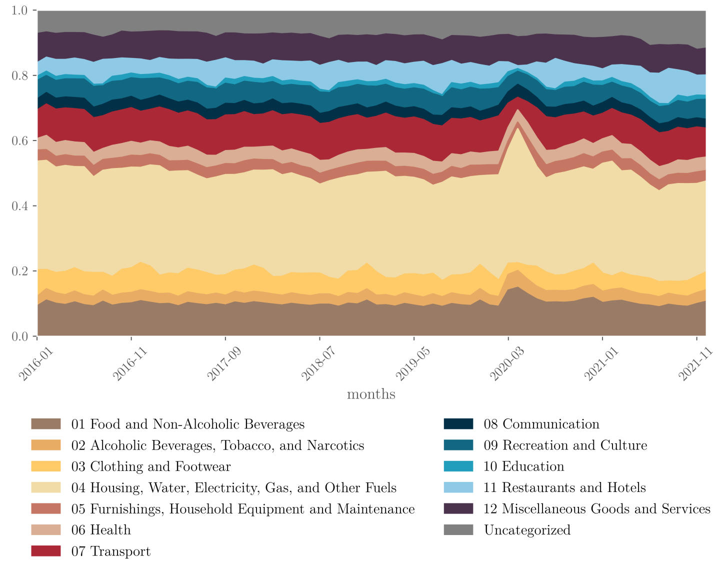


Figure 11: Proportions of Consumption by COICOP at Monthly Frequency

(or other components). INE produces no official statistics at the substantially more detailed provincial level.

Figure 12 illustrates provincial inequality as measured by the naturally occurring consumption survey. For each separate province (50 in total), we compute monthly, seasonally adjusted aggregate consumption and then divide by the number of adults in the province. Figure 12 plots the distributions of monthly per-person consumption for each month. In general, one observes a mass of provinces in the center of the distribution and a long right tail of higher-consumption provinces. We believe such statistics are likely to form an important evidence base for documenting and designing policy responses to spatial inequality.

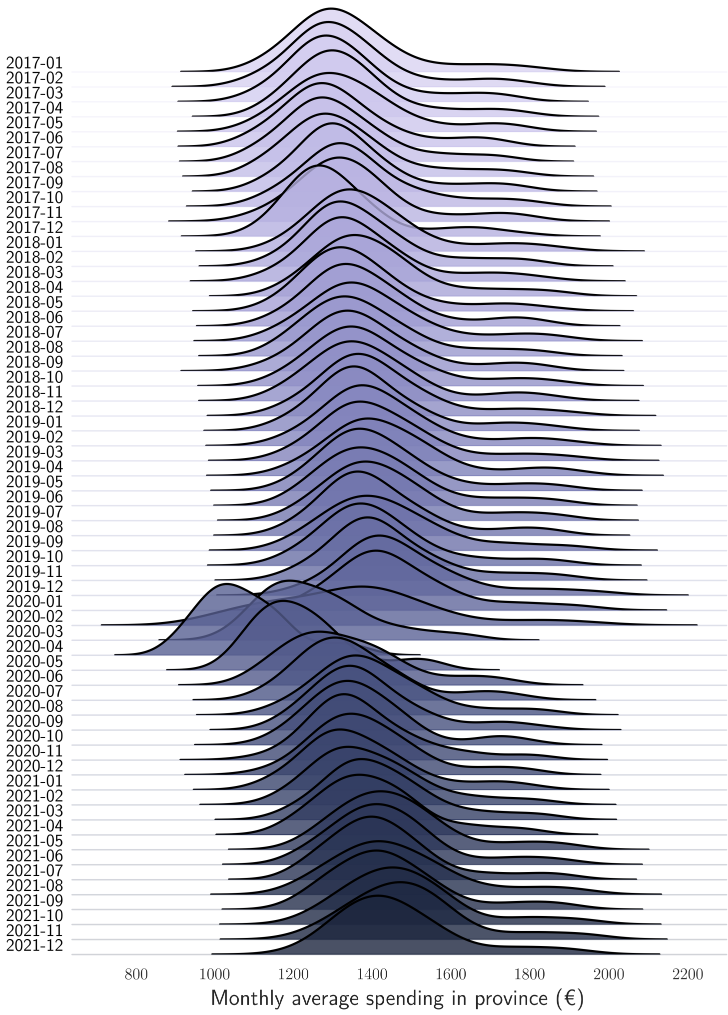
3.2.3 Means of payment

The naturally occurring consumption survey also contains useful information on mode of payment. While sporadic survey evidence exists documenting how often individuals use different payment modes, these typically do not condition on payments related specifically to consumption. Figure 13 shows the evolution of transaction mode over the sample. One observes a steady increase in the use of online card transactions to satisfy consumption needs, and a steady decline in the use of cash. Such information is potentially important for determining the importance of alternative payment technologies for aggregate welfare.

4 Distributional National Accounts for Consumption

As we have just seen in the previous Section, our transaction data, when organized and classified via national accounting principles, provides a high-quality match with published national accounts by Spain's National Statistical Institute. Importantly, this result immediately implies that our underlying microdata can be additionally deployed to build distributional national accounts for consumption, characterizing both the distribution of consumption levels - and hence consumption inequality - and the evolution of this distribution over time - and therefore consumption inequality growth.

Following the seminal work of Piketty et al. (2018), distributional national accounts for *income* already exist for a large number of countries. Combining existing national accounts aggregates, censuses, household surveys, and micro income tax data, this macro-consistent accounting methodology has ar-



a

Figure 12: Distribution of Spending per Person across Provinces

[a]For each province in Spain (50 in total), we compute aggregate consumption at monthly frequency. We then seasonally adjust each monthly series using the Jdemetra+ application and applying X-13ARIMA-SEATS, and divide each series by provincial population. For each month, we then produce a kernel density to describe the cross-sectional distribution of per-person consumption. Since the aggregate consumption is representative at the regional (Comunidad Autónoma) level, there may be some demographic imbalance present at province level.

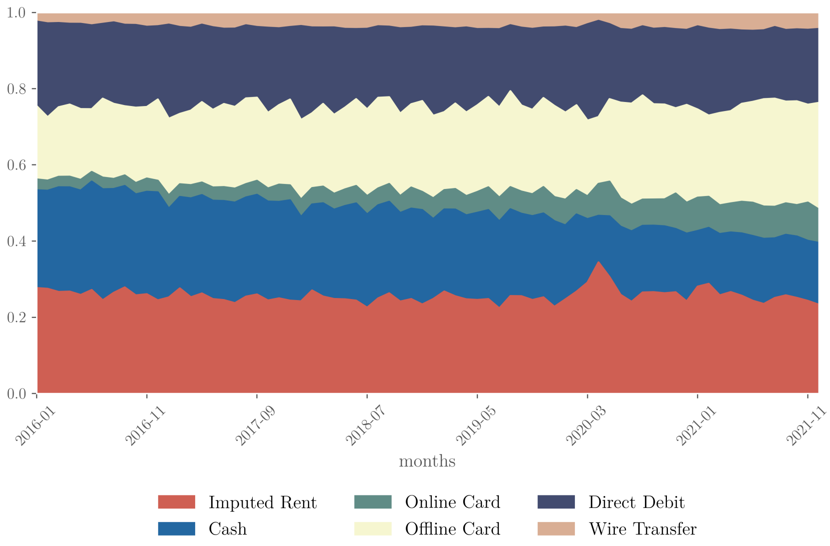


Figure 13: Shares of different payment methods

guably had a large impact in both academic and public discussions surrounding income inequality and its time evolution. Yet, to the best of our knowledge, distributional national accounts for *consumption* are virtually non-existent.[17] To the extent that individual consumption, consumption inequality and its evolution more closely reflect welfare-relevant objects than income per se, this is an important gap. Arguably, this gap exists precisely because traditional consumption surveys - the typical data source deployed to analyze the extent and evolution of consumption inequality - *are not* consistent with national accounts, as extensively discussed in the literature reviewed in the Introduction and also in the previous Section.

Thus, in this Section, we present a first distributional accounting exercise for consumption based on BBVA transaction data. We do this distributional analysis both for consumption levels and its growth. Specifically, Section 5.1. below provides a detailed discussion of the macro-consistent distribution of consumption across adults in Spain in 2017. This allows us to provide a description of consumption inequality in Spain across a variety of measures. Further, we benchmark our analysis in two ways. First, we compare our results against existing distributional accounts for post-tax income in Spain, concluding that macro-consistent consumption inequality is substantially smaller than its income counterpart. Second, we benchmark our analysis against consumption inequality as implied by the Spanish Household Budget Survey (Spain's equivalent to the US CEX). We show that the latter not only underestimates aggregate consumption figures (and hence averages) but also displays different properties at the upper tail of the consumption distribution, consistent with undersampling (or under-reporting) of high-income/high-consumption households. Finally, given the rich metadata available in our setting, we show that it is also possible to break down this distributional analysis further, across consumption categories, demographics (age and gender) and time frequencies. In doing so, we show that it is possible to reproduce - and then go beyond - analyses typically pursued in the consumption inequality literature but, importantly, (i) in a way that is both consistent with the level and evolution of aggregate consumption and (ii) with data that is arguably less encumbered by the documented sampling biases of traditional consumption surveys.

In Section 5.2 we turn our attention to the distributional accounts of consumption growth. Our data

[17]The exception to this is the recent work by the joint OECD-Eurostat 'Expert Group on Disparities in National Accounts Framework.' Concurrently to our own work, this joint OECD-Eurostat effort has produced a first set of *experimental* distributional accounts for income and consumption (see Coli et al, 2022), which is only possible under arguably very strong and counterfactual assumptions. We review this experimental work below and compare to our own results.

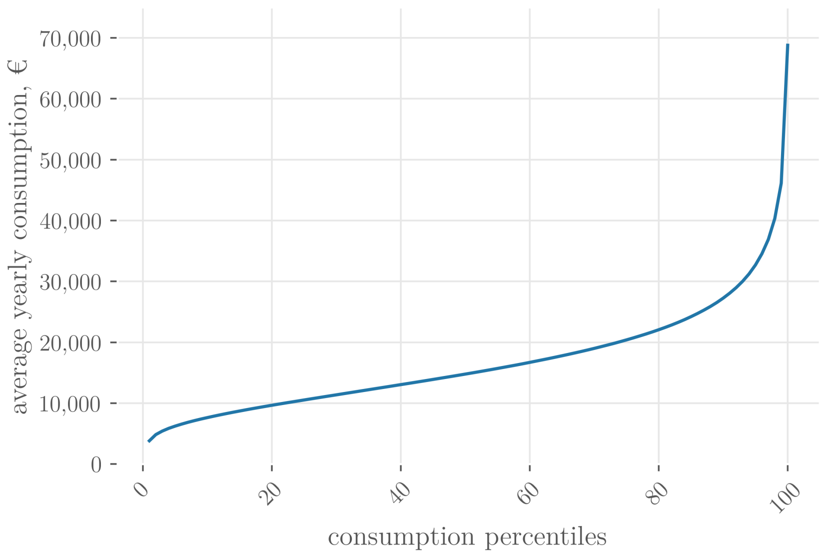


Figure 14: Yearly adult consumption across the distribution of levels of consumption in 2017; Weighted sampling procedure, 2017 Euros. Y-axis gives annual average consumption per adult in the corresponding X-axis percentile of consumption.

spans only the period 2015Q2-2021Q4 and is therefore unable to resolve long term trends in consumption inequality. On the other hand, it does include both the onset of the COVID pandemic and associated lockdowns as well as the subsequent recovery period. We thus provide a macro-consistent account of the evolution of consumption inequality in Spain, in the years before the pandemic, during the large recession in 2020 and over the period of strong recovery in 2021.

Before turning to our analysis, a word on methodology. In order to ensure that the distribution of consumption does aggregate to national accounts consumption, we need to ensure that our sample is representative of the adult population in Spain. In the previous section this was achieved by properly weighting the micro-data as we aggregate. Here, however, we are interested in analysing the micro-data itself. In order to do this while preserving representativity with respect to the population, throughout we follow the sampling (with replacement) scheme described at the end of Section 2.3, where sampling weights reflect the corresponding population weights for a given cell.[18]

4.1 The Distribution of Aggregate Consumption: Levels of Consumption Inequality in 2017

We start by analysing the extent of inequality in consumption in Spain. As stressed above, our distributional national accounts for consumption capture 100% of aggregate consumption, allowing us to compute consumption for each quantile of the consumption distribution, in a manner consistent with macroeconomic aggregates.

Figure 14 plots the cross-sectional distribution of consumption across consumption percentiles in 2017, with the Y-axis giving the annual average consumption of a Spanish adult in a given percentile in the consumption distribution.

Given our distributional national accounts framework notice that, mechanically, average adult consumption arising from this micro-distribution necessarily coincides with the per-adult aggregate Spanish

[18] We currently present results based on a single representative draw from the underlying micro-data. As discussed above, these samples are large enough to minimize sampling noise and obviate the need for a full bootstrapping procedure averaging over multiple samples.

consumption figure presented in the previous Section. For 2017, this number stands at 16,907 Euros or – combining this with official Spanish GDP figures – 56% of per-adult GDP in Spain. A first indication of the extent of inequality in consumption is then given by the fact that, instead, the median Spanish adult in 2017, consumed 14,971 Euros. Thus, indicative of inequality in the upper tail of the consumption distribution, the median adult in 2017 consumed 12% less than the average consumer in Spain.

This first indication of consumption inequality is confirmed by looking directly at the tails of the consumption distribution. Thus, the typical adult consumer at the 90th percentile of the 2017 consumption distribution consumed roughly 2 times more than the median consumer, at 28115 Euros. In other words, the familiar p90/p50 ratio in 2017 is 1.87. Going further towards the tail of the consumption distribution, the average consumer at the top 1% consumed about 68893 Euros, implying a p99/p50 ratio of 4.6. Finally, at the very top of the consumption distribution, the 0.1% consumption-richest adults consumed 128907 Euros in 2017 while the top 0.01% consumed roughly double that, at 242490 Euros. In other words, the typical adult at the top 0.1% (0.01%) of the consumption distribution, consumed 8.6 times more (respectively, 16.2 times) than the median consumer in Spain for the year 2017. [19] Finally, and switching our attention to the consumption-poor end of the distribution, a typical adult at the bottom 10th percentile of the 2017 consumption distribution, consumed only 7869 Euros, roughly half of the median adult consumer and 3.6 times less than the top 10% adult.

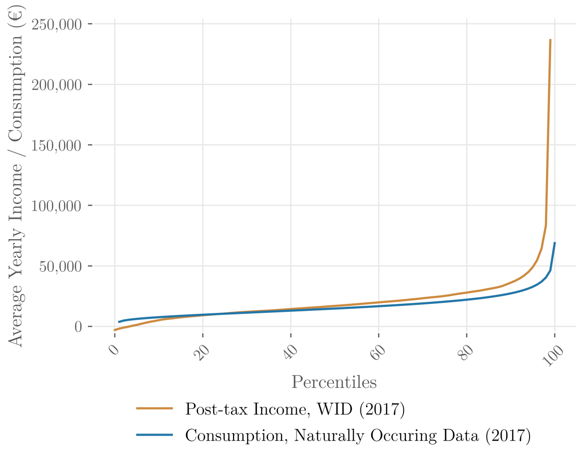
An related metric of interest is given by the associated cumulative distribution function of micro-level adult consumption. In particular, one can ask, for example, how much of total 2017 aggregate consumption in Spain accrued to the top 10% consumption-richest adults? The answer implied by the empirical consumption distribution is that 22.4% of total consumption in Spain accrues to the top 10%. Furthermore, there is again evidence for concentration of consumption at the top, with the top 1% (0.1%) accounting for 4.1% (respectively, 0.8%) of total consumption in Spain. In contrast, the bottom 50%, account 31% of total aggregate consumption in Spain, whereas only 4% of accrues to the 10% consumption-poorest adults in Spain.

The above distributional analysis of consumption can be readily comparable with available distributional accounts for *income.* To do this, we source tabulated data made available by the World Inequality Database (https://wid.world/country/spain/). The latter follows the methodology pioneered in Piketty et al. (2018), combining existing national accounts aggregates, censuses, household surveys, and micro income tax data. The specific methods and concepts used in the WID are reviewed in detail in Alvaredo et al. (2021) and the resulting long-run analysis for Spain is presented in Alvaredo et al. (2019). Throughout, given substantial redistribution and progressivity in the Spanish tax system, we focus our attention on measures of post-tax income inequality.

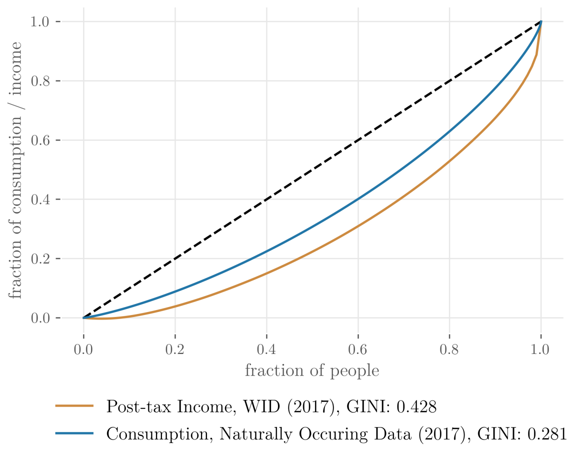
The top panels of Figure 15 summarize the comparison between these two sets of macro-consistent distributional national accounts, for consumption and income respectively. Thus, the left top panel overlays our own consumption distribution with the post-tax national income distribution for Spain (by income percentiles in 2017), made available by the World Inequality Database. The top right panel plots the implied Lorenz curves for these two distributions.

Clearly, by either metric, and however unequally distributed consumption in Spain is, inequality in consumption is substantially smaller than income inequality. For example, the p90/p50 ratio for post-tax income is 2.13 vs 1.87 for consumption as introduced above. As is clear from Panel (a) in Figure 15 these differences become increasingly more pronounced as we move to the top of the income and consumption distributions, with a p99/p50 (p99.9/p50) ratio for income of 14.82 (respectively 57.6), three

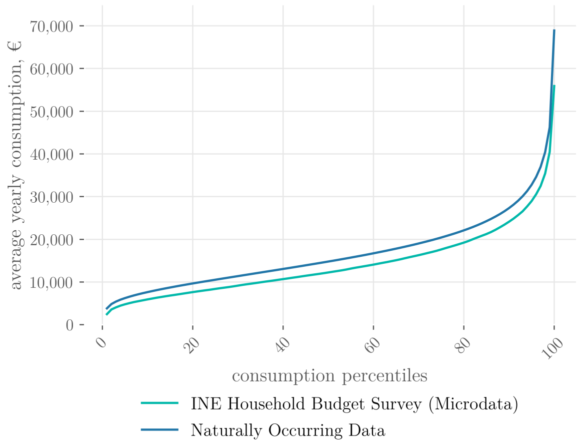
[19] Note that Figure 14 plots the average adult consumption per percentiles of the 2017 consumption distribution. Therefore the maximum Y-axis value corresponds to that of the top 1% of the consumption distribution rendering inequality within the top 1% invisible in this figure.



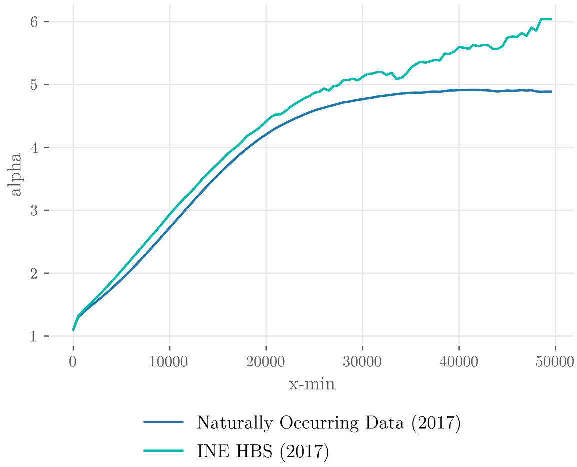
(a) Distribution of 2017 consumption per Spanish adult in BBVA vs. WID 2017 post-tax income distribution.



(b) Lorenz Curves of BBVA 2017 consumption and WID 2017 post-tax income in Spain.



(c) Distribution of 2017 consumption per Spanish adult in BBVA vs. 2017 Spanish Household Budget Survey consumption per adult distribution.



(d) Estimated shape (α) parameter for best fit power law distribution as a function of scale parameter ($x - min$), BBVA vs Spanish Household Survey.

Figure 15: BBVA Consumption Inequality vs Income Inequality (Top Panels: a,b) and HBS Consumption Inequality (Bottom Panels: c,d)

(respectively, 8.6) times the value observed for the equivalent moment(s) in the consumption distribution. Alternatively, focusing on CDF measures underlying the Lorenz curves in Panel (b), note that 31% of total national post-tax income accrues to the top 10%, a larger proportion than that of consumption (at 22%), reviewed above. Consistent with our discussion, the gap between the concentration of income and consumption increases as we look into the very top quantiles of the distribution, where the top 1% (0.1%) post-tax share of income is 11% (respectively, 4.2%), nearly three times (respectively, 5.2 times) the concentration of consumption observed at the top. Conversely, looking at the bottom of these distributions, we find that the 10% income-poorest account for only 0.4% of aggregate income, ten times smaller than the corresponding number for consumption. Taken together, the measurements for income and consumption imply that overall inequality (as measured by the Gini index) is 50% larger for post-tax income relative to consumption.[20]

An alternative comparison is possible with the Spanish household budget survey (HBS henceforth), discussed above. When properly weighted, the Spanish HBS is designed to be representative of the Spanish population. However, unlike our data, consumption is reported (mostly) at the household level while making note of the number of adults in the household. In order to render the implied consumption distribution comparable to that obtained with our data, in the remainder of this section, we split the total household consumption reported in HBS equally across all adults in the household.

Recall further that, as discussed above and unlike our data, the Spanish HBS is not consistent with the aggregate level of consumption reported in the national accounts. This is a not an idiosyncratic problem of the Spanish HBS but a rather more general problem for consumption surveys across the world, as reviewed in the Introduction. Further possible shortcomings of consumption surveys, beyond recall failures by the households interviewed and relatively small sample sizes, include (i) under-reporting in certain consumption categories (for example, the consumption of tobacco, alcohol or gambling services) and (ii) non-response, under-reporting or under-sampling of high-income/high-consumption households at the very top of the distribution. These concerns suggest that it is important to assess whether conclusions regarding inequality in the distribution of aggregate consumption change as we move from a traditional consumption survey to a macro-consistent naturally-occurring consumption survey, enabled by large scale transaction-level data.

Panel (c) in Figure 15 summarizes our results by plotting the implied consumption distributions in our data vs. that in the HBS. Consistently with our discussion in the context of aggregate national accounts, note that the consumption distribution implied by the BBVA data is uniformly above that of the Spanish household budget survey. This confirms a lower average (and hence also lower aggregate) consumption per Spanish adult in the HBS. Comparing the two distributions, we additionally document that this discrepancy in average adult consumption worsens as we move to the very top of the consumption distribution. To see this, note that the median adult consumer in our data consumes 21% more than the median adult in the Spanish HBS. This discrepancy is still stable at the 99th percentile of the consumption distribution, where we observe a 23% increase in average consumption as we move from the HBS to our data. However, at the top 0.1% of the distribution this discrepancy increases to 54% and then further at the very top 0.01% to 90%. In other words, the consumption-richest consume almost double the amount of goods and services as we move from HBS to our data.

This finding is consistent with under-sampling or under/non-response at the very top of the consumption distribution and, in turn, implies that some familiar inequality ratios (such as p99.9/p50) would be

[20] The finding that consumption inequality is smaller than income inequality also suggests that the post-tax savings distribution is more unequal than income distribution. Note, however, that for this to be a direct implication of the findings Panel (a), the ranking of Spanish adults in the income distribution would need to correlate highly with their respective consumption ranking. While this does not seem an overly strong assumption, we cannot at this stage verify it, as we (currently) do not deploy BBVA measures of income in this analysis.

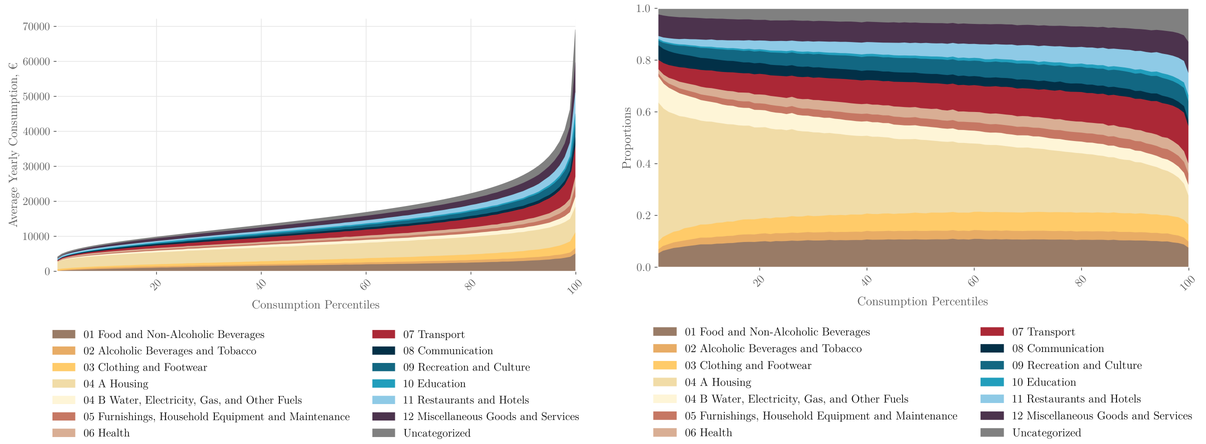
particularly underestimated in the HBS relative to our data. More generally, this also has implications for the characterization of tail behavior in the consumption distribution. In particular, as is well-known, Pareto distributions typically offer a particularly simple parametric way to encode fat tail distributions. Here we follow Clauset et al. (2009) maximum-likelihood methods to estimate the implied power-law behavior (estimating both the scale, $x - min$, and shape parameters, α) in both our data and in the HBS. We also subsequently follow their likelihood ratio methodology to assess the respective fit against alternative parametric distributions.

For our BBVA transaction data the maximum-likelihood estimates of the power law shape parameter is 3.91, holding at the tail of the distribution, for observations of above a minimum scale of 38000EUR. Consistently with the discussion above, comparing income and consumption distributions, this behavior is considerably less fat tailed than the typical benchmark of a Zipf distribution for income (implying a shape parameter of 1). This value of 3.91 is also similar to the only comparable published estimate we are aware of, that of Toda and Walsh (2015) who obtain a value of 3.65, based on the implied cross-sectional consumption distribution by the CEX consumption survey for the US.

Redoing the same exercise on the Spanish HBS, we instead obtain an estimated shape parameter of 4.07. Thus, conditional on imposing a power law fit, our results imply somewhat less mass at the top tail of the consumption distribution in the HBS relative to our transaction data. More importantly, formally testing the power law assumption against other parametric alternatives reveals a different behavior across the two empirical distributions. Using likelihood ratio tests, we find that for BBVA data, the power law parameterization provides a statistically significant better fit when compared to lognormal or exponential alternatives. However, for the HBS data, these findings are reversed, with lognormal providing a statistically significant better fit. Panel (d) of Figure 15 provides an intuitive visualization of these differences. Specifically, we plot the implied Pareto shape parameter estimate as a function of the scale parameter. For the BBVA transaction data, we see that the implied tail behavior - as encoded by the shape parameter - is stable after the minimum scale has been reached. For the HBS data instead, the estimated shape parameter diverges (towards ever thinner parameterizations of the tail behavior) as we move towards the top of the HBS consumption distribution. This different behavior of the tails of the consumption distribution across the two datasets again suggests that - as hypothesized in the literature on consumption surveys - the Spanish HBS is undersampling the top of the consumption distribution. In turn, this rationalizes differences both in the levels of implied aggregates but also in the analysis of inequality of consumption at the upper tail. Our distributional accounts seem to improve on this outcome, being both consistent with macro-aggregates and providing further resolution at the upper tail.[21]

Finally, note that concurrent with this paper, a joint OECD-Eurostat initiative has produced a first set of *experimental* distributional accounts for income and consumption for OECD countries, including Spain, for the year 2015 (see Coli et al. 2022). These experimental accounts recognize the first-order coverage limitations of household consumption surveys which prevent the compilation of distributional accounts for consumption. In order to make progress, they propose to heavily impute by scaling the household survey data up, at the level of COICOP category subsaggregates, and then applying the same correction factor (given by macro total/micro total) to all households in the sample. By construction, this forces household surveys aggregates to match national account aggregates, thus circumventing the failure of aggregation problem. As Coli et al. (2022) discuss "the assumption behind this approach is that the distribution found in the sample survey is close to the real distribution of the household

[21] Note that additionally, this contrasting behavior at the tail has important implications for the consistency of GMM Euler equation estimates, as stressed by Toda and Walsh (2015). This is because, as is well known, higher order moments do not exist for sufficiently fat tailed power law distributions but do exist under log-normality.



(a) Consumption inequality disaggregated by COICOP consumption categories (Levels).

(b) Consumption inequality disaggregated by COICOP consumption categories (Shares).

Figure 16: Consumption distribution disaggregated by COICOP consumption categories, levels (LHS) and shares (RHS).

population, meaning that potential under-reporting or sampling errors are evenly distributed among the population."[22]

Our analysis provides a first benchmark with which to evaluate these imputation concerns. As discussed above, we find that undersampling and underreporting at the very top of the consumption distribution is indeed a first order concern for the HBS relative to our data. Thus, by dimputing missing consumption in the HBS under the assumption of evenly distributed sampling errors, the proposed OECD-Eurostat experimental accounts will necessarily impart bias in inequality calculations.[23]

4.1.1 The Distribution of Aggregate Consumption Across Consumption Categories

Thus far we have focused our attention on the distribution of total consumption per adult. However, it is possible to gain further understanding of key drivers of this distribution by disaggregating further total consumption into specific consumption categories. Again, recall that this disaggregated distributional accounts exercise is made possible by leveraging the abundant metadata associated with each transaction and classifying it, as discussed above in Section 2.

Thus, in Figure 16 we present results on the decomposition of consumption across major COICOPS consumption categories. Panel (a), on the left of Figure 16 plots the distribution of consumption in levels, allowing us to inspect how spending (in 2017 Euros) in specific categories of consumption varies across the consumption distribution, the latter still given by percentiles of total consumption in 2017. Note that, as before, integrating consumption over the distribution - for either total consumption of COICOP subaggregates - produces macro-consistent aggregate figures. We complement this information with Panel (b), Figure 16, on the right, where we plot shares of total consumption across the distribution of total consumption.

Perhaps not surprisingly, a first order implication of this disaggregated analysis of the distributional

[22]Further, as a recognition of this very strong assumption, OECD-Eurostat warns that the accuracy of the implied results maybe "imperfect" or "insufficient" and unable provide fine-grained analysis of consumption inequality beyond quintiles of the population due to "a higher uncertainty at both the very top and the very bottom of the distribution, which is inherent to survey data."

[23]Note further that, as we shall discuss in the following subsection, the consumption bundle of households is not stable across the consumption distribution, with high consumption households consuming proportionately more luxury items. This, in turn, suggests that consumption inequality of certain types will be relatively more distorted than others by the proposed imputation.

accounts for consumption, is that inequality in consumption is largely the result of highly unequal discretionary, or luxury-type, consumption.

To see this, consider constructing two subsaggregates, necessities vs luxuries. We include in consumption necessities, Food and Non-Alcoholic Beverages (COICOPS category 01), Alcohol and Tobacco (02) Clothing and Footwear (03), Housing and Utilities spending (4A and 4B) and Health (06). Conversely, we include in luxuries spending, Furnishings and Household Equipment (05), Transport (07), Communication (08), Recreation and Culture (09), Education (10), Restaurants and Hotels (11), Miscellaneous Goods and Services (12) and the unclassified residual category, Uncategorized expenditures.[24] Clearly, given the coarse COICOPS classification we work with, any such disaggregation will be fraught with some amount of miss-classification. For example, the COICOPS category Transport, includes both necessary commuting and public transportation expenses as well as discretionary type consumption such as vehicle purchases or tourism. By the same token, Clothing and Footwear includes both low quality/low price apparel and high-end luxury-brand consumption. With this proviso in mind, we proceed with our analysis based on these groupings.

Considering first the expenditure share on necessities, these constitute 57.4% of total consumption for the median adult in the consumption distribution. Consistent with the concept of necessities, this share declines strongly over the consumption distribution, accounting for 67% of total consumption of adults at the bottom 10%, 49% of total consumption of the top 10% and only 29% of the top 0.1%. The upshot of this is that, though total consumption of necessities does rise with total consumption (as clear from Panel (a) for levels of spending), the implied consumption inequality arising from consumption of a necessities is somewhat smaller than that for total consumption. For example, the p90-p50 ratio is 1.60 (relative to 1.87 for total consumption) and the top 10% share of aggregate consumption of necessities is 0.19 (relative to 0.22 for total aggregate consumption).

The flip-side of this argument is that the distribution of luxury consumption is highly unequal. Indeed, the bottom 50% of the consumption distribution only accounts for 24% of aggregate luxury spending, while the top 10% accounts for a disproportionately large 30%. As expected, luxury consumption is concentrated at the very top and, for example, accounts for 71% of consumption of the average adult at the top 0.1% of the consumption distribution. Alternatively, using the Gini Index as a univariate measure to summarize inequality in the distribution, the distributional facts above imply that luxury consumption is 38% more unequally distributed than total consumption.

4.1.2 The Distribution of Aggregate Consumption by Age and Gender

Given that our raw transaction includes information on each consumer, we can also present distributional national accounts disaggregated by demographic characteristics. In particular, in this section, we analyse consumption inequality by age and gender.

Thus, the left Panel Figure 17 depicts the 2017 consumption distribution by age group, where the y-axis gives both the average adult total consumption in a given age category and its breakdown across COICOP consumption categories. Again note that given our distributional accounts setting all figures aggregate up to macro-consistent totals.

The familiar hump-shaped pattern of consumption previously documented in the literature is also evident in our data. Thus, consistently with Aguiar and Hurst (2013) and Fernandez-Villaverde and Krueger (2007), we see that adult consumption grows throughout the 20s and 30s, peaks in middle age, and declines smoothly thereafter. Quantitatively, the average Spanish adult between 35-40 years old,

[24]Note that pre-college education is largely publicly provided and free at the point of use in Spain and undergraduate education, while not free, has low yearly tuition fees between 750EUR and 2500EUR a year.

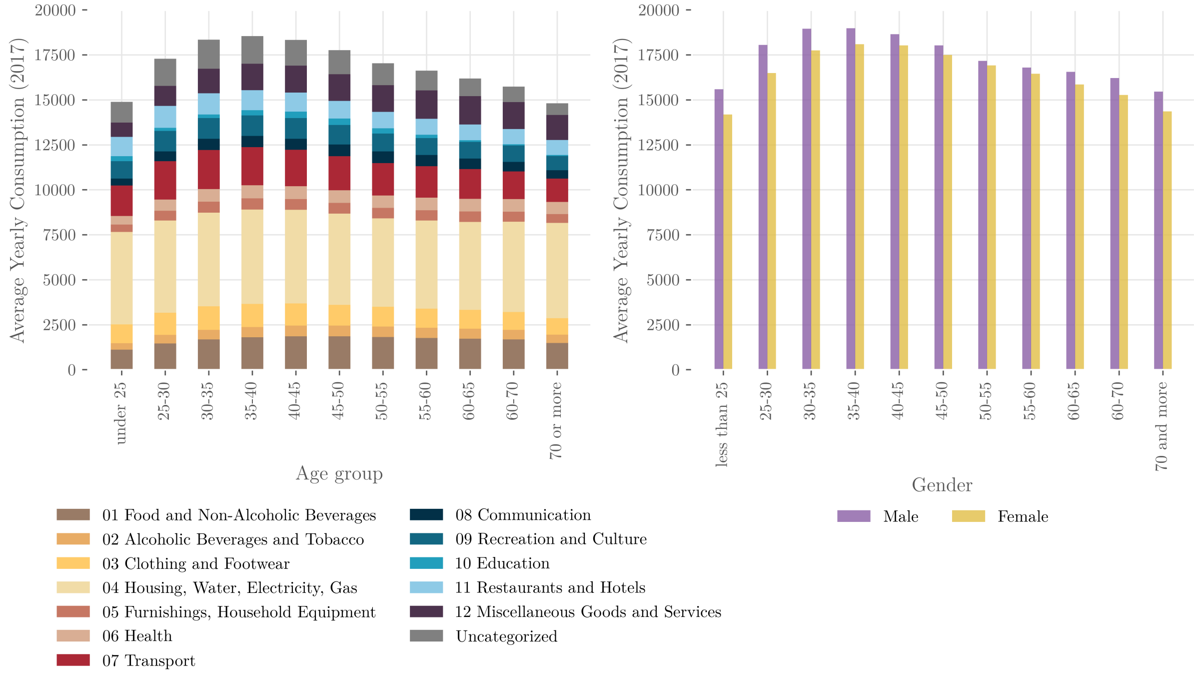


Figure 17: The Distribution of Consumption by Age and Gender.

consumed around 18500 Euros during 2017, 10% more than the average adult consumer in Spain during the same year (and almost a quarter more than the median adult consumer). Conversely, both those under 25 and those over 70, consume 9% less that the average adult in Spain. That is to say, over the life cycle we observe a 20% increase in consumption from young adulthood till middle age, followed by a similarly sized decline in consumption into old age. These quantitative findings on consumption are broadly consistent with Aguiar and Hurst (2013), albeit somewhat smaller (the latter documents 25% declines from peak to through but focuses only on non-durable expenses).

As before, it is also possible to explore the rich metadata associated with each transaction to obtain a distributional accounting of this hump-shape across age *and* consumption categories. In particular, like Aguiar and Hurst (2013) we confirm that the post-middle age decline in consumption is partly the result of a decline in consumption of Restaurants and Hotels, Transport, and, to a lesser extent, Clothing and Footwear. Unlike Aguiar and Hurst (2013) findings for the US, we also find an important role for the decline in Education expenses and Recreation and Culture.

Again exploiting the demographic data at our disposal, the right panel in Figure 17 depicts another aspect of consumption over the life cycle, focusing on the heterogeneity across females and males. First note that, despite splitting consumption equally for all those within collective consumption units (e.g. married couples), our distributional accounting of aggregate consumption, still exhibits a 6% gender gap in consumption. Thus, in 2017, the average adult male in Spain consumed 17390EUR whereas the average adult female consumed roughly 1000EUR less, at 16399EUR.

Interestingly, as Figure 17 displays, this gap is not constant over the life cycle of males and females. While both mean and women exhibit a clear life cycle profile peaking in middle age, the consumption gender gap is largest for those in their 20s and early 30s, then declines slowly attaining a near-parity minimum for those aged between 50 and 55, while opening up again from the 60s onwards.

This evidence is consistent with a broadly documented gender income gap penalty due to career interruptions during typical childbearing ages; see for example Guvenen et al. (2020) for recent evidence.

However, it can also be mechanically driven by our assumption of equally splitting consumption within household units. In particular the decline in gender gap observed from young age till the 50s, could be the result of a gradual selection into co-habitation or marriage over the life cycle coupled with the assumption of equal-split consumption. To address this possibility, we redo the computations above for a subsample of singles (i.e. those unassigned to a collective unit), where results are unaffected by assumptions on household-level consumption and its distribution across members and, in particular, gender.[25] In this singles subsample we find a slightly larger consumption gender gap, with the average single Spanish female consuming 8.6% less in 2017 than its average male counterpart. This larger gap for singles is consistent with consumption redistribution within the household playing a non-negligible role in the level of gender consumption inequality. Importantly, the life-cycle patterns observed for singles are qualitatively similar to the ones reported for our full-sample baseline. Again, we find a U-shape pattern, with the consumption gender gap gradually declining as move from young adulthood till late middle age - again attaining a minimum at age 50-55 - followed by a (more marked) worsening of the gap from the late 50s onwards. [26]

4.1.3 The Distribution of Aggregate Consumption Across Time Frequencies

While we have been focusing our attention on distributional accounts of consumption over the course of a year, one additional advantage of high-resolution transaction data is that it allows us to conduct distributional analysis at varying time frequencies, from daily to multi-year time windows. This flexibility is likely important to policy-makers and analysts when considering the real-time/high-frequency implications of major shocks, such as the COVID-19 crisis and subsequent policy responses. It is also important more generally, for understanding how the frequency of measurement of consumption may interact with conclusions on the level and dynamics of consumption inequality.

For example, as Coibion et al. (2021) conclude, "a decline in shopping frequency as households stock up on storable goods will lead to a rise in expenditure inequality when the latter is measured at high frequency, even when underlying consumption inequality is unchanged." That is, the level of consumption inequality may be spuriously affected by the conjunction of two facts. First, at high frequencies, individual consumption is lumpy, due to infrequent purchases of durable (e.g. cars or household equipment) and non-durable yet storable goods kept in inventory by households over weeks or months. Second, many consumption surveys (the CEX in the US being a prime example, but also the Spanish HBS), include a high-frequency "diary" component, where households are asked to provide an account of the level and distribution of consumption over a limited - 2 weeks in the case of the CEX - time window. This is aimed at both improving measurement - by reducing recall errors in household reporting - and more generally reducing the burden (and hence attrition) imposed by consumption surveys.

Thus, the upshot of such a survey design is that, given lumpiness in consumption, whenever the frequency of consumption purchases is lower than the survey recall period, measured consumption inequality may mechanically be biased upwards. This argument holds at both high frequencies for storable goods (as in Coibion et al. (2021) and at lower frequencies for durable goods.[27]

[25] Note that for computations in this subsample we do not weight the observations according to their distribution in the population. Our objective here is not to provide a distributional analysis of the subaggregate given by Spanish singles' consumption; rather this serves simply as a robustness check to our main distributional exercise along gender.

[26] Quantitatively, and consistent with the larger average consumption gap in this singles subsample, the consumption gender gap we observe is larger at all points on the life cycle of singles, attaining a minimum of 5.4% at age 50-55

[27] To see the first case, suppose that two households consume the same amount of, say, Food and Non-Alcoholic beverages at a monthly frequency, based on a single Superstore trip every month. If the two household trips are not synchronized in the same week, a survey conducted at the weekly frequency will conclude for inequality in consumption across the two households when there is none at the biweekly frequency. The second case is a simple extension of this argument at longer frequencies. Take for example two households purchasing the same car every five years but at different position in their buying cycle over time. Then, a yearly consumption survey will again conclude for consumption inequality when there is

Frequency	Gini index	Variance of Log
Daily avg. (2017)*	0.629	1.019
Weekly avg. (2017)	0.439	0.572
Monthly avg. (2017)	0.338	0.347
Quarterly avg. (2017)	0.307	0.296
Yearly (2017)	0.281	0.257
Pre-Covid 3 Years (2017-2019)	0.273	0.244
All 5 Years (2017-2021)	0.265	0.231

*30 days sampled randomly

Table 7: Inequality of consumption in different time windows

We now exploit the flexibility that our transaction data allows to quantify the extent to which the level of inequality in the distribution of aggregate consumption depends on the frequency of sampling. In particular, given that we observe real-time expenditure at every frequency, we can simulate what a hypothetical survey would conclude, depending on the frequency design of such survey. Further, as the result of the distributional accounts framework, this measurement is consistent with aggregate consumption, at every frequency.

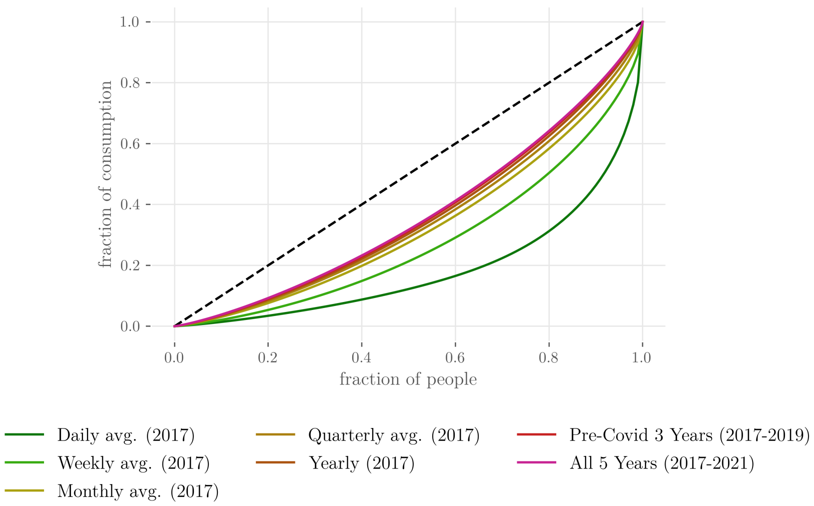


Figure 18: Lorenz Curves of the distribution of consumption across different time frequencies.

To do this, Table 7 presents two traditional univariate measures - the Gini index and variance of log consumption - giving the extent of inequality in the distribution of aggregate consumption, where the latter is measured at different frequencies, from daily, to week, to month and quarterly, to yearly, to lower frequencies, over a single three years (pre-covid) and the five years spanning the 2017-2021 sample. To avoid being distorted by high-frequency outliers, for frequencies below one year, we take the average of either inequality measure, over the available 2017 observations at that frequency. For example, the extent of inequality, when measured at the monthly level, is given by the average of the relevant inequality statistic over the 12 months of the year.[28] At yearly and lower frequencies, we simply report the observed value of total consumption during the relevant period, be it one, three, or five years. Finally, Figure 18 complements these results by showing the associated Lorenz curves over the entire distribution of aggregate consumption, implied by the different sampling frequencies.

none at a five year horizon.

[28]The exception to this is the daily frequency due to computational constraints on the size of data. We opt to sample 30 days uniformly at random during 2017.

The first-order results are clear from both tabulations and Lorenz curves. Inequality in the distribution of total consumption declines strongly with the sampling frequency. Thus, a hypothetical consumption survey of all Spanish adults tabulating their consumption the previous day would find inequality, as measured by the Gini Index (log variance of consumption), to be 2.4 times (respectively, 4.4 times) larger than another survey reporting inequality in total consumption over all five years of our sample.

Notice also that the bulk of this measured decline in inequality happens as we move from very high frequency to the year level. Though the infrequent purchases of consumption durables does still drive the level of inequality down as we move from a one-year to five-year window, the effect on measured inequality is strongest when additionally incorporating the higher frequency – yet relatively infrequent purchases of storable and semi-durable goods.

To further understand this point, we additionally perform the same analysis at the level of specific consumption categories. Thus we compare the behavior of measured consumption inequality across time frequencies, for consumption of Food and Non-Alcoholic Beverages (COICOPS 1) vs that of Furnishings and Household Equipment (COICOPS 5). Intuitively, the first category should correspond to non-durable - but storable at high frequencies - consumption, while the second provides an example of household durables, displaying more low frequency purchasing behavior. Consistently with our analysis above, measured inequality in Food consumption declines very rapidly across high frequencies. Inequality in food consumption at the monthly frequency is roughly half (55%) of that measured at the daily frequency. Conversely, for Furniture and Household equipment this decline in measured inequality at high frequencies is considerably slower, decaying by 32% as we move from the daily to the monthly level. The corresponding Lorenz curves and further results for lower frequencies are given in the Appendix to the paper.

Taken together, these strong high frequency effects, in turn, suggest that - consistently with Coibion et al. (2021) - whenever (*i*) consumption is to be measured via a diary or other high-frequency consumption survey methodologies and (*ii*) consumption purchasing habits are shifting across frequencies, the effects on measured inequality growth will be counterfactually high and, further, that this bias is heterogeneous across consumption categories.

4.2 The Distribution of Consumption Growth: 2017-2021

Our data not only allows us to analyze the differences in consumption patterns across Spanish adults but also allows us to look at the evolution of these differences, while still consistent with the level and dynamics of national accounts macro-aggregates. That is, the distributional properties of the data allow us to go beyond Figure 5 (where we saw that our aggregates track remarkably accurately the growth rate of consumption in aggregate national statistics), and analyze the micro-level *distribution* of growth rates of consumption. Thus, we not only determine the evolution of inequality of consumption, but we also use the microstructure of our data to determine who benefits from this growth, and by how much. We can also perform this exercise by finer cuts of the data. In particular, we focus on the evolution of the distribution of consumption in different categories, partly as an example of use of our data, and partly because it helps in order to make sense of the dynamics of consumption during the COVID pandemic.

In the remainder of this section, we will follow Piketty et al. (2018) and treat our data as a succession of cross-sections. In the following section, we will make use of the fact that our data is in fact much richer than the synthetic data typically used to construct distributional accounts, as individuals can be followed over time, showing that we can determine individual (not only aggregate) consumption dynamics.

Arguably the most influential result that has arisen from the literature on distributional accounts, is

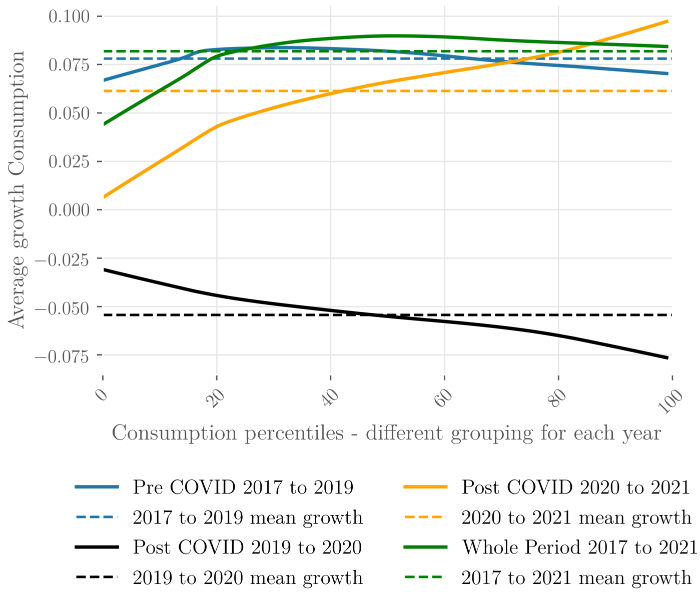


Figure 19: Growth Rate of Consumption per Percentiles (cross-sections)

the attribution of aggregate income growth to different percentiles of the income distribution. Attributing the growth rate of income to each percentile of the income distribution generates a decomposition of aggregate growth, determining where in the distribution the gains or losses are located, as well as their magnitude. Nevertheless, to the best of our knowledge, these distributional dynamics exercises have so far been limited to income. Here we provide the first analysis of such distributional exercises for consumption. Figure 19 gives the main result plotting the distribution consumption growth across the consumption distribution.

Our data spans only five years and is thus unable to resolve secular trends. This being said, it includes a particularly interesting period, the COVID-19 pandemic, as this period experienced what is one the largest consumption shocks ever recorded, with massive changes that were unevenly distributed across agents and consumption categories.

Specifically, in Figure 19, following Piketty et al. (2018), we index the percentiles of consumption on the horizontal axis, and on the vertical axis we plot the total growth in consumption for each consumption percentile. For instance, the green solid line in the graph plots the distribution of growth between all percentiles in the whole period (2017-2021).[29] It maps, to each percentile in the horizontal axis, the growth rate in the total consumption for Spanish adults who were in that percentile in 2021 and the consumption of the individuals who were in that percentile four years earlier That is, we sum the total consumption of all the individuals in the sample (appropriately weighted) during the year 2017. We then order individuals by total consumption levels and assign them to percentiles. We do the same in the year 2021 and report the change in consumption in each percentile between the two periods. Importantly, note that individuals within the percentile do not need to be the same at the beginning as at the end of the period (actually, we will see in the next section that, in general, they are not the same).

Given that we follow our weighted sampling procedure at all points in time, our growth rates aggregate mechanically into the growth rate of private consumption in national accounts. In particular, the horizontal dotted green line is the average (and hence aggregate) growth rate of nominal consumption during this 5-year period. Thus, if for a given percentile the solid green line lies below the dotted line

[29] As before we present a smoothed version of the underlying scatter plot.

(as we observed for the lowest percentiles), the share of total consumption for this percentile decreased; if it lies above it, it implies that the corresponding share increased.

In the Figure, we represent four sets of lines informing on the distribution of consumption growth in the years before the pandemic, the year when the pandemic hit (2020), the subsequent recovery (2021) and also the distribution of growth rates over the whole 5-year period that our data covers.

In blue we show the evolution of the distribution of consumption in the "normal years", before the pandemic shock. Growth is larger in the top 10% of the distribution than in the bottom 10%, indicating that the p90/p10 ratio did increase. Nevertheless, the curve is rather flat and is increasing only at the very bottom of the distribution. From the 20th percentile onwards the growth rate actually decreases. Thus, it is difficult to argue that overall consumption inequality increased during these "normal" years. Indeed, we find that the Gini Index decreased from 0.281 in 2017 to 0.279 in 2019. Thus, while the share of the lowest percentiles indeed increased less than at the median[30] or the average (the horizontal line) the share of consumption in the highest percentiles did decrease with respect to the median, and simply kept pace with average consumption growth in Spain. In fact, the share of the top percentiles decreased with respect to all other individuals above the 20th percentile. Thus, even if some measures of inequality (P90/P10 ratio for instance) can result in an apparent worsening of consumption inequality, the overall picture is rather different from the one that appears when considering the evolution of the income distribution: it is not that the top percentiles of consumption were decoupling in relation to the rest of society. The evolution of the distribution of consumption over these years seems strikingly different from the one that we are used to observe when thinking of income inequality, as in the equivalent plot in Piketty et al. (2018).[31]

As we saw above, we can also generate disaggregated distributional accounts at the consumption category (COICOP) level, and we can equally measure its evolution. We do so in Figure 20, where we present the same growth allocation per percentile of total consumption for each of the COICOP categories separately. Looking only at the blue (pre-COVID) lines, it is clear that inequality decreased in all categories except "Communication" (COICOP 08) and, to a lesser extent, "Education" (10)[32].

4.2.1 Distributional Effects on Consumption of the Pandemic

As mentioned in section 1 a large literature has used card transaction data to study the effects of the pandemic on consumption,[33] showing that the combination of the pandemic itself and the consequent restrictions on activity resulted in (i) a large decrease in consumption, that (ii) was skewed towards the relatively better-off.

Our data has the advantage of including *all expenditures*, not only those paid via cards. As we saw in section 3, this means that we can account for roughly twice the amount of consumption. Given that this extra spending (paid via transfers, cash, or debits) does not react equally to card spending to the COVID shock and, further, that it is not distributed equally across the population, it is worth using our data to take a look at the distributional changes on total consumption during the pandemic.

The yellow and black lines in Figure 19 are a graphical representation of those distributional effects of the COVID pandemic on total consumption. Unlike previous measurements, they do aggregate into changes in consumption in National Accounts.

The black lines graph the distributional change of aggregate consumption between 2019 and 2020, the

[30]The p50/p10 ratio increased from 1.902 to 1.909

[31]We have not found a paper using the time dimension of distributional accounts to generate the same plot for income in Spain, but Alvaredo et al. (2019) finds that the income share of the top 1% grew in Spain since 2010 at a higher speed than in France, and close to US speeds.

[32]Remember that this does not include public education provision

[33]For Spain, see Carvalho et al. (2021)

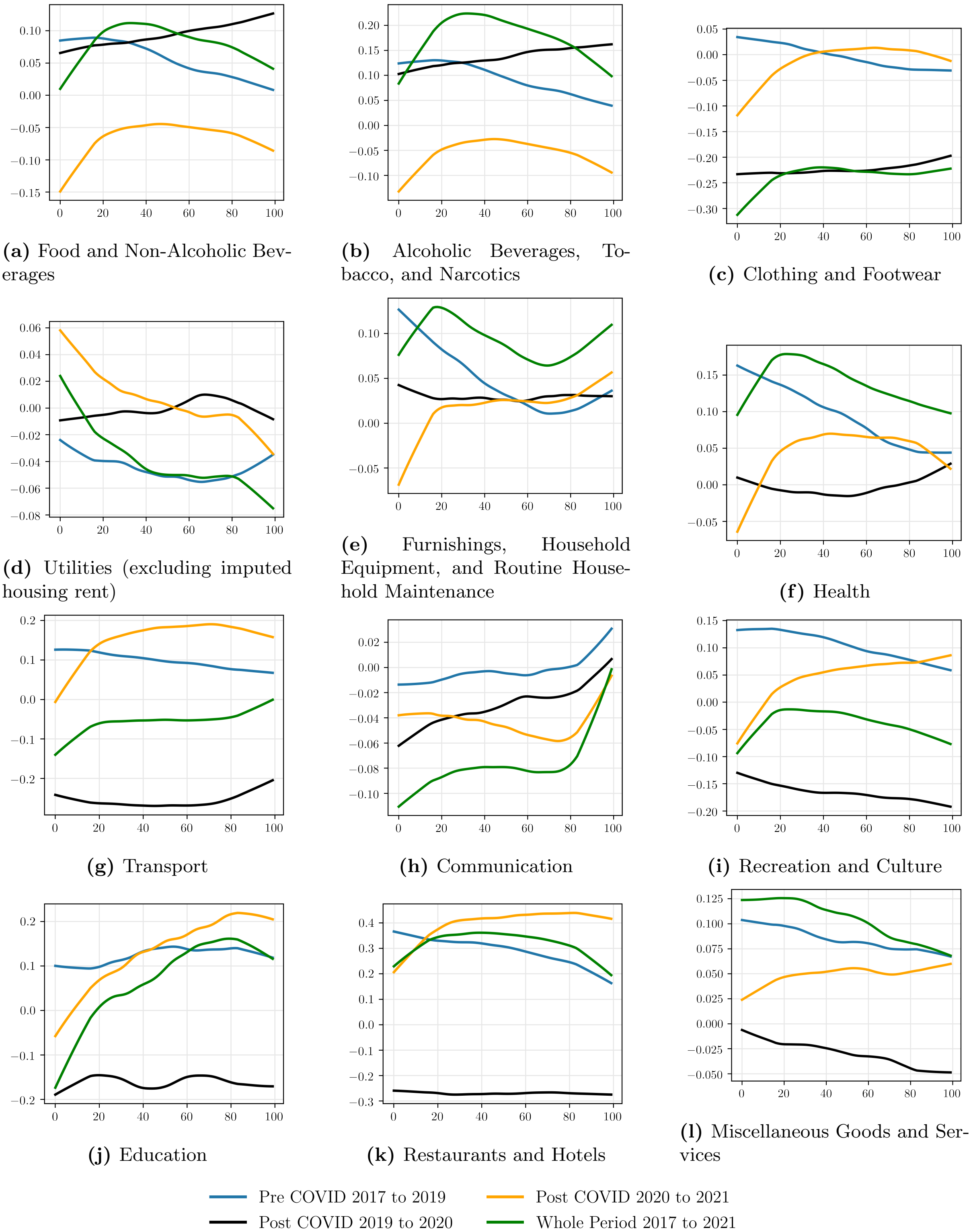


Figure 20: Growth rates per consumption percentile and COICOP. Cash allocated per percentile as offline card purchases.

year of the pandemic. It shows (i) a generalized decrease in consumption, and importantly (ii) that this decrease is larger at the right tail of the distribution. Therefore, in this very peculiar year consumption inequality decreased unambiguously, as the shares of all percentiles above the median did monotonically decrease and those of all percentiles above it did monotonically increase.

This finding is consistent with previous analysis in Carvalho et al. (2021), showing that the restrictions imposed by lockdowns, had a larger impact on the consumption of relatively well-off agents relative to poorer ones, as they applied primarily to luxury items (such as traveling, restaurants, etc.) whose consumption tends to be concentrated among the better-off. This is apparent in Figure 20, where the very large decrease in inequality took place in "Recreation and Culture" and "Miscellaneous" (within our definition of "luxuries"). Instead, for categories like "Food" and "Alcoholic beverages and Tobacco" (all within our definition of "necessities") not only did consumption *increase* in all groups during this year, but its consumption got more concentrated at the top of the distribution.

Going back to aggregate consumption (Figure 19), notice that the slope of the curve is less pronounced than one would expect looking only at card data. This is because of two reasons: (i) as we saw above goods and services not paid with cards were not as affected by the restrictions and the consequent recovery, and (ii) their share in total consumption is flatter across all percentiles. In other words, luxuries are more likely to be paid for with cards and were more affected by the pandemic. Thus, the decrease in inequality in consumption is substantially smaller when looked at in our distributional accounts than when looked at only in the light of card spending.

The yellow lines show the distributional consumption dynamics during the recovery from the pandemic, for both aggregate consumption and per category in Figures 19 and 20, respectively.

Perhaps not surprisingly, we see that total consumption (Figure 19) increased across all percentiles. However, this recovery was more substantial at the top of the distribution. As a result, the cross-sectional inequality in consumption increased substantially and unambiguously during this period: all percentiles below the median decreased their shares while all above it monotonically increased them. Switching to our the COICOPS-disaggregation of these dynamics, this is largely due to an increase in the consumption of luxuries at the very top. In turn, this is also consistent with high-income consumers accumulating extra savings during 2020. Thus, the increase in consumption inequality during the COVID recovery of 2021, likely reflects pent-up demand for luxuries at the top of the distribution and excess savings by consumption-rich Spanish adults. Consistent with this, further note - in Figure 20, that the increasing profile of the yellow curves is also present in the "Recreation" and "Miscellaneous" categories (and, incidentally, in "Education", suggesting a movement toward private education in the academic year following the pandemic), again suggesting strong growth in luxury consumption at the top of the distribution. Conversely, for necessities like "Food" and "Alcoholic Beverages" there was a decrease in consumption across all percentiles during this recovery period, suggesting a substitution from food at home towards expenditure in restaurants, which grow strongly in this period.

Finally, the green lines plot the overall change from 2017 to 2021. Remarkably, despite the massive shock in between, the overall evolution during the whole period is not that different from the evolution in the pre-covid period.

Taking the distribution of aggregate total consumption growth, there is a very small increase in the Gini Index (from 0.281 to 0.283) reflecting two different opposite movements in the distribution. First, we observe that the shares of the lowest percentiles (up to p20) decrease relative to the top (and the median). Second, this slight increase in inequality at the bottom is partially counteracted by an increasing consumption share of the median adult relative to the consumption-rich. These countervailing effects acting on different ends of the consumption growth distribution imply a weak overall effect on inequality which increases slightly.

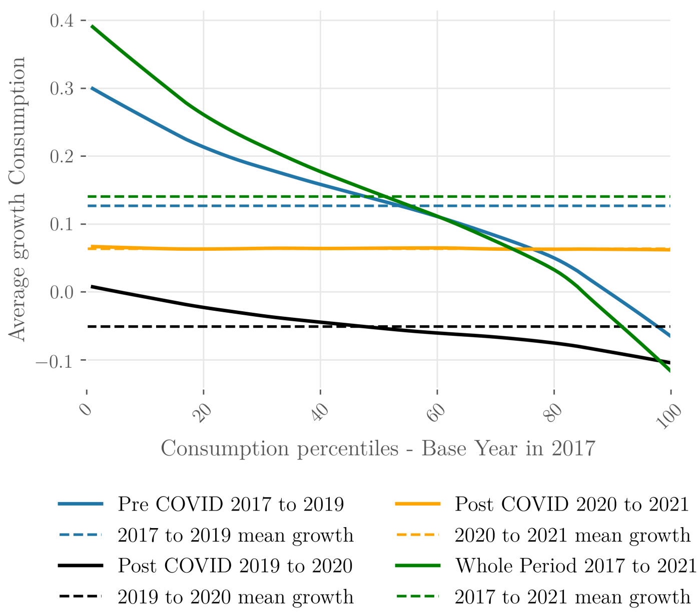


Figure 21: Growth rates conditional on initial consumption percentile

Turning to the distribution of consumption growth across COICOP categories the entire COVID decline and recovery episode has amounted to large increases in the inequality of consumption in "Education" and "Communications", while for the bulk of consumptions categories, including both necessities and luxury goods, the overall trends are similar to the pre-COVID ones, with a general tendency for a slight decrease in the dispersion of consumption in these categories.

5 Individual Consumption Dynamics Across the Consumption Distribution

Thus far we have shown that our data is able to both reproduce aggregate national accounts and – by acting as a high quality, highly detailed consumption survey – renders feasible the production of rich distributional accounts. In this section, we show that, by leveraging its panel dimension, the data allows us not only to study national aggregate and distributional accounts but also to produce a detailed analysis of the underlying patterns of individual consumption growth processes.

Thus, in this section, we move away from standard distributional accounts and make use of the fact that, in our data, individuals can be tracked over time; our goal is to describe the dynamics of consumption at an individual level.

To be clear, we do not aim here to provide a rationalization of these dynamics, as doing so we would need to look also at the income and portfolio position of agents at any given point in time, and that remains beyond the limit of this paper. Nevertheless, the results we obtain are interesting in and by themselves. Indeed, detailed high quality consumption panels where representative populations can be tracked across multiple time periods are still rare. As such, we are able to establish non-parametric facts that are seldomly studied. In particular, we show that individual consumption changes are not only very large indeed, but they follow non-linear dynamics that, further, do not seem to be well approximated by Gaussian distributions.

In Figure 21 we start our analysis by plotting a graph apparently similar to Figure 19, but with the important difference that we group agents by their percentiles of total consumption for the year 2017. That is, we order individuals (after sampling according to the method explained in Section 2.3) by their total consumption percentile during the year 2017. In the vertical axis, we plot the average growth rate of consumption of *these* agents. We are thus following a panel where each individual is associated with her percentile in 2017, and we measure the average consumption growth of the members of each 2017 percentile.[34]

First notice how different the resulting plot is from Figure 19. The blue line, denoting the growth rate between 2017 and 2019, is decreasing in all percentiles, and negative for the highest ones. That is, adults that were in the lowest percentiles of consumption in 2017 did, on average, increase their nominal consumption strongly over this 3-year period (a 30% for those in the bottom percentile), while agents initially located in the top percentiles of the 2017 distribution did, on average, decrease consumption (almost five percent). This provides a first suggestion for strong mean reversion behavior in consumption, which we retake below. Before that, it is worth additionally noting the black and yellow lines portraying (as before) the pandemic and subsequent recovery respectively. The black line is only slightly decreasing, while the yellow line appears flat. In other words, consumption mobility is such that conditioning on the the fact that a particular adult was located in a certain percentile of consumption in 2017 provides very little information - and indeed does a very poor job at predicting - on the *growth* of its consumption just a few years later, from 2020 to 2021. The green line - depicting the overall change over the five-year period - is strongly decreasing because it contains information about 2017: those who consume a lot in this initial year will tend to consume less in the next ones while those who consumed little initially are likely to increase consumption substantially in the following ones.

In principle, this mean-reversion can be due to two effects acting simultaneously. First, to the extent that income is mean reverting and correlated with consumption this, by itself, must generate mean reversion in consumption.The second effect is related to the above discussion in Section 4.1.3: Consumption is lumpy, with infrequent purchases of storable/durable goods being prevalent. The upshot of these consumption spikes is that an individual's consumption level in a given base year may not be strongly related to her present or lifetime income but, instead, reflect the lumpy nature of purchases that year. That is, individuals may experience an infrequent high-consumption year (say due to the purchase of household equipment or vehicles) which, due to the durability of the underlying high-value goods, both significantly affects the ordering of individuals over total consumption *and* is likely to be followed by relatively low consumption years as individuals enjoy the service stream of the durable good with no need to replace it, thus inducing mean-reversion in individual consumption. While we recognize the importance of differentiating the role due to changes in income (or information on the income stream) from the role due to grouping and unusual purchases,[35] we do not, at this stage, aim to do so. Still, let us remark that we perform our exercise with one-year growth innovations, which as we have seen are vastly less lumpy than at daily or weekly frequency, and with similar lumpiness than five-year frequencies.

In a recent key paper, Guvenen et al. (2021) have used US social security data to perform a similar (much broader) exercise for the US income distribution. While we do not fully characterize the non-parametric distribution of consumption (as they do for income), we follow some of their methods and start by plotting in Figure 22 the log density of the individual growth rates of consumption. This is, we calculate all the annual growth rates at individual level in our data and plot the (log) kernel distribution. In this plot, we draw the distribution of 7.2 million individual annual growth rates of total consumption

[34] A consequence of this is that the data aggregates to national accounts for 2017, but not exactly for the following years. The reason is that the weights of different ages, regions, etc., change slightly from year to year. The data shown here should be thought of as the distribution growth rate of adults in 2017, which does not need to be the same as in the following years. The average growth rates shown are, thus, slightly different from the ones shown in Figure 19.

[35] There is a literature aiming to differentiate these effects, see for instance Madera (2019)

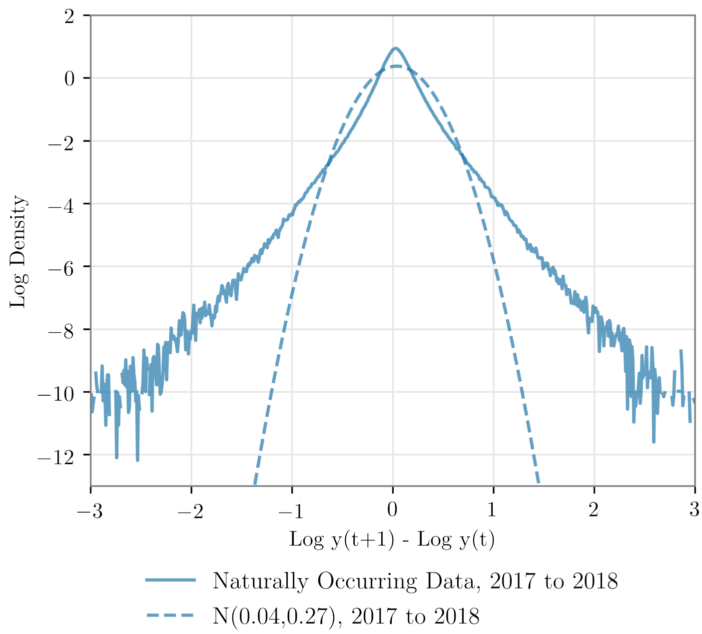


Figure 22: Log density of one-year consumption growth.

of a representative sample of society; being able to take a much more careful look at the distribution than what would be possible by following any previously existing panel of consumers. In the same plot, we draw the distribution of the best Gaussian fit of our data (by definition a parabola in log space). It is clear that - just as Guvenen et al. (2021) report for individual income growth in the US - consumption growth does not seem to be well approximated by a Gaussian distribution. Rather, the linear log-log relationship suggests a form of Pareto distribution for consumption growth in both the left and the right tail.

This is a surprising result, as traditional consumption smoothing thinking suggests that consumption is not supposed to change dramatically from one year to the other. Notice that we are not remarking here that the distribution of consumption has a fat tail (that we already did in section 4), but that the consumption *growth* in itself has fat tails on both sides, with a non-negligible mass of agents both decreasing and increasing their consumption by very large magnitudes from one year to the next. While it is interesting that consumption levels have fat tails, it is perhaps more surprising that the changes do, as consumers with concave preferences should strive to avoid those large oscillations.

It is worth, thus, to zoom-in and carefully examine who is suffering (or enjoying) these large consumption shocks. We devote the rest of the section to describing the distribution of these changes when conditioning on the level of consumption and age of the agents.

In Figure 23 we plot the average annual growth for each of the percentiles of consumption over the whole period. That is, for each year we group people according to percentiles of total consumption in that year, measure its growth in the following year, and report the average growth per percentile *over all of the years in our data*.[36] As in Guvenen et al. (2021), we plot this separately for different age groups, so that we can resolve any differences per consumption group and age.

There are two lessons to take from this figure. In relation to our discussion in Section 4.1.2, notice that the average growth rate of consumption decreases with age independently of the initial consump-

[36]For this and the following plots we performed the same exercise only with the years before the COVID pandemic and obtained the same qualitative results.

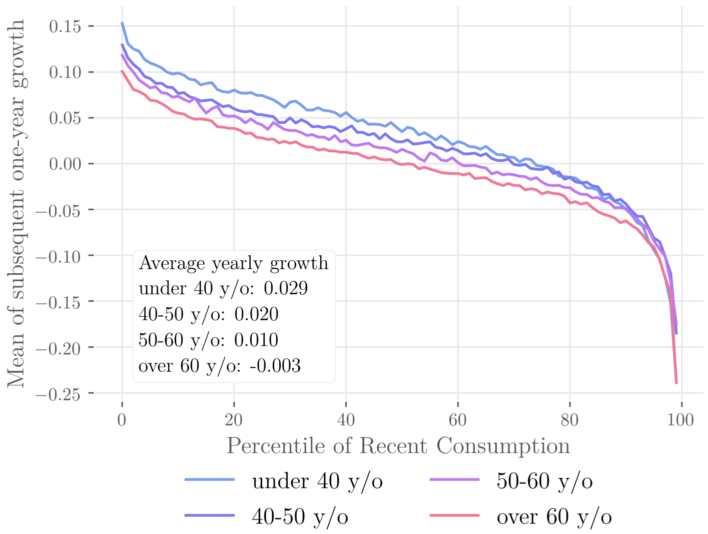


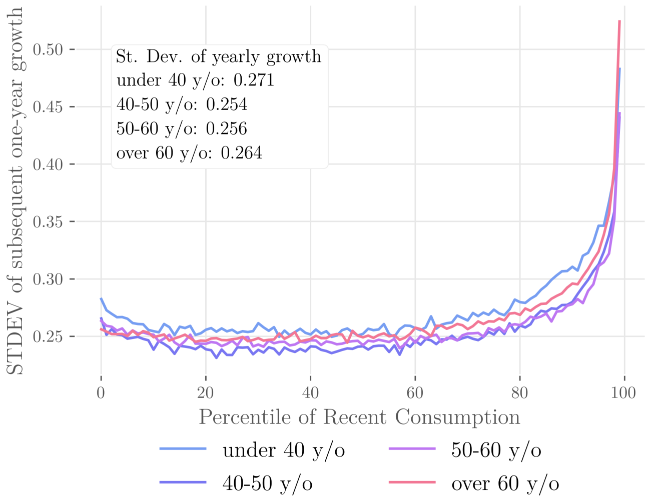
Figure 23: Average one year growth per consumption percentile and age group

tion percentile of the agent. More relevant to our current discussion, notice the strong and fast mean reversion. Agents in the lowest percentiles of the consumption distribution have large average increases in consumption over the following year (of about 10%), while agents that are in the highest percentiles sustain a severe decline.

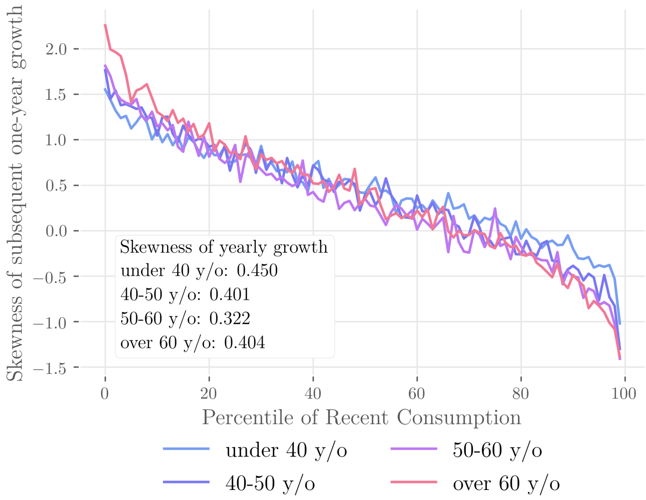
In Figure 24 (again, following the exercise that Guvenen et al. (2021) perform on income changes in the US) we plot the second, third, and fourth moments of the one-year change in consumption on the percentiles of the consumption level the previous year, as in the previous figure. To the best of our knowledge, there is no analysis of micro-level consumption processes to which we can compare or benchmark our results. Rather, in what follows, we compare the properties we observe for individual consumption with the moments reported by Guvenen et al. (2021) for US income processes.

For example, notice first (Figure 24a) that the standard deviation of individual adult growth in our data is much smaller than in the US income data, where the minimum is 0.6 and the maximum is 1.1. It is also slightly higher for younger relative to older adults, but these differences are small in comparison to the difference in volatility between agents with low and high consumption. Albeit the volatility of the growth rate is flat (at around 0.25) up to the 70th percentile, it then increases strongly, being twice that level for the highest percentiles. Thus, one-year ahead consumption for individuals with currently high consumption is highly volatile, but not so for those in the lowest percentiles. This contrasts with the patterns in US income data, where the standard deviation of growth is U shaped, being high for rich but also for poor agents, as exiting unemployment produces high volatility for low incomes. Thus, and recapping our findings, relative to consumption poor adults, those at the top of the consumption distribution are both expected to, on average, decrease their consumption subsequently (mean reversion) but also face substantially higher consumption risk.

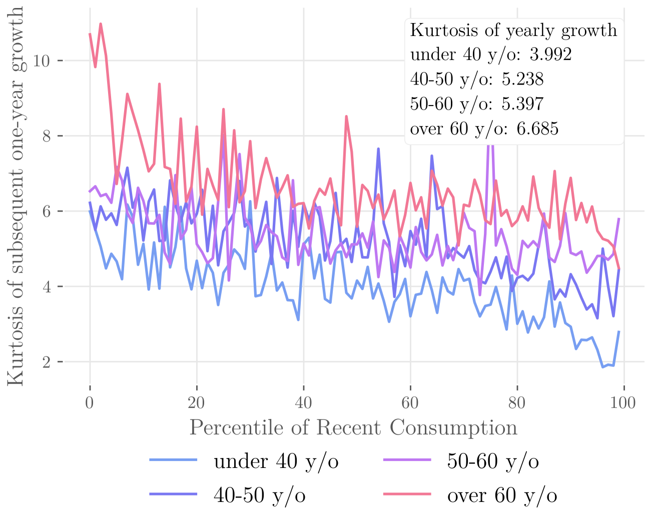
Skewness in our consumption data (Figure 24b) also behaves differently relative to the patterns reported on US income. Specifically, innovations in for US individual incomes display negative skewness for all percentiles of current income (they are left-tailed), implying that there are frequent small gains and few large losses. In contrast, in our data, consumption growth skewness is highly negatively correlated with the level of consumption in the current year. Individuals in the lower percentages of consumption present positive skewness, i.e. a right-tailed distribution of consumption growth. As we saw, for these individuals average growth is expected to be positive; thus, the positive skewness behavior suggests



(a) standard Deviation



(b) Skewness



(c) Kurtosis

Figure 24: Moments of one year growth per consumption percentile and age group

frequent small increases below this average and a few unusual but very large positive consumption growth episodes. Adults in the top percentiles of consumption present, on the other hand, negative skewness. Their average growth is negative, so that the left-tail behavior of the distribution implies frequent small declines and unusual but large declines for some individuals.

It is also interesting to note that this skewness at the tails is larger for older agents than younger ones. The large and infrequent shocks driving *consumption leaps* (both increases and decreases) are more prevalent among the old than among the young. For individuals close to the median, the distribution of consumption shocks is essentially symmetrical.

Finally, we find positive excess kurtosis (Figure 24c) in the distribution of consumption growth for almost all percentiles and age groups, suggesting fat tails. The degree of kurtosis is clearly decreasing in the consumption percentiles, and higher for older agents than for younger ones. Indicating the possibly non-expected finding that older agents who currently are at the bottom of the consumption distribution are more prone to have extreme changes in their consumption than other agents.

Moreover, there are interesting differences between our kurtosis curve and the one that Guvenen et al. (2021) report for US income growth. Specifically, the shape is the reverse. Kurtosis of income growth (at least in the US) is increasing with the level of income, taking a value of about 3 (similar to a normal distribution) for the lowest percentiles of income, and growing to about 15 for percentiles close to the top of the US income distribution.

The picture that arises from putting together these facts, and contextualizing them with the distribution of income changes in the US, is both interesting and challenging. Our data suggest a process of consumption at the individual level that is lumpy, non-linear, and not easily reproducible with Gaussian distributions. In that respect, it is very much like the recent characterization for income processes in the US; but of course, there are important qualitative differences across the two.

Summarizing, in our data, there is an extremely strong mean reversion in consumption but in a lumpy manner. Agents with low consumption most often have small positive consumption growth, but sometimes enjoy large positive changes. Particularly if they are old, there are thick tails on the right of the distribution.

The consumption growth of agents with high current consumption is particularly difficult to predict. They are also very likely to have small changes (declines in this case) in their consumption but face unusual large negative shocks moving them far towards the left, and generating high volatility.

Our naturally occurring data, thus, not only allows for the computation of detailed national accounts (both aggregate and distributional in nature) but also allows us to explore its microstructure, opening new avenues to improve our understanding of economic behavior. It suggests that if agents in our data are aiming to smooth their consumption, they seem to be failing at doing so, and not in a small manner.

Clearly, our results suggest the need to understand these consumption changes; and to incorporate income, its stochastic structure, and the expectations on its evolution that agents may have. Moreover, it demands to do so while looking at the lumpy character of purchases of durable and storable goods.

6 Conclusion

Plentiful, naturally occurring transaction data can be used at a relatively low cost to generate complex, careful, accurate, and encompassing information on economic activity. Our paper advocates the use of this unstructured, but readily available, data for both the construction of national aggregate and distributional accounts as well as the study of the microstructure of economic activity.

Our proof of concept results imply that simple and transparent procedures - resulting from organizing the data around national accounting principles and ensuring a representative sample of a country's population - followed by bottom-up aggregation tracks with remarkable accuracy not only the growth rate of consumption in national accounts, but also its level. Further, due to its granularity it allows immediate decomposition across goods, demographics, space or time frequencies. In particular, the good aggregation properties of the data allow for a distributional analysis of aggregate consumption, providing a rich, macro-consistent description of consumption inequality and its time-evolution. Finally, we have seen that this same data that aggregates into national accounts can be used to analyze the microstructure of the economy. Specifically, we have studied the dynamics of individual consumption, demonstrating that at one-year frequency they are characterized by lumpy movements from the extremes of the distribution of consumption growth. All this is made possible because transaction data, once properly organized, can be deployed of as a high-quality, large scale, real-time consumption survey containing both information on the consumption decisions of millions of individuals along with rich metadata which tags billions of transaction.

Clearly, there are a plethora of further covariates within the BBVA data that can be used in order to go beyond consumption and better track other areas of economic activity. For example, transaction data includes a vast amount of firm-level information that can be used to understand production activity. We believe it is also possible to further analyse the workings of the external sector and government activity, or generate regional and sectoral input-output tables, a key object for national accounts. Finally, returning to consumption, at the micro-level it is also possible to augment our consumption panel with further covariates (chiefly income and wealth), rendering it possible to better understand consumption-savings decisions and help the development of theory.

References

AEA Committees on Economic Statistics and Government Relations (2020). Statement on the Need for Accurate and Reliable Data from the 2020 Decennial Census. Technical report, American Economic Institution, Nashville.

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.

Alvaredo, F., Artola Blanco, M., and Martínez-Toledano, C. (2019). Long-Run Income Inequality in Spain, 1932-2015. Evidence from Distributional National Accounts.

Alvaredo, F., Atkinson, A. B., Bauluz, L., Fisher-Post, M., Garbinti, B., Goupille-Lebret, J., Martínez-Toledano, C., Neef, T., Piketty, T., Robilliard, A.-S., Saez, E., Yang, L., and Zucman, G. (2021). Distributional National Accounts Guidelines - Methods and Concepts Used in the World Inequality Database. Technical report, World Inequality Lab.

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.

Attanasio, O. P. and Pistaferri, L. (2016). Consumption Inequality. *Journal of Economic Perspectives*, 30(2):3–28.

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.

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.

Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703.

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

Coli, A., Istatkov, R., Jayyousi, H., Oehler, F., and Tsigkas, O. (2022). Distributional national account estimates for household income and consumption: methodological issues and experimental results. Technical report, Eurostat.

Ehrlich, G., Haltiwanger, J. C., Jarmin, R. S., Johnson, D., and Shapiro, M. D. (2022). Reengineering Key National Economic Indicators. In *Big Data for Twenty-First-Century Economic Statistics*, pages 25–68. University of Chicago Press.

European Commission (2010). Report on Greek Government Deficit and Debt Statistics. Technical report, European Commission, Brussels.

Fernandez-Villaverde, J. and Krueger, D. (2007). Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data. *The Review of Economics and Statistics*, 89(3):552–565.

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.

Guvenen, F., Kaplan, G., and Song, J. (2020). The Glass Ceiling and the Paper Floor: Changing Gender Composition of Top Earners since the 1980s. In *NBER Macroeconomics Annual 2020, Volume 35*, pages 309–373. University of Chicago Press.

Guvenen, F., Karahan, F., Ozkan, S., and Song, J. (2021). What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics? *Econometrica*, 89(5):2303–2339.

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.

Krueger, D., Perri, F., Pistaferri, L., and Violante, G. L. (2010). Cross-sectional facts for macroeconomists. *Review of Economic Dynamics*, 13(1):1–14.

Madera, R. (2019). The Consumption Response to Tail Earnings Shocks. *SMU*.

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.

Piketty, T., Saez, E., and Zucman, G. (2018). Distributional National Accounts: Methods and Estimates for the United States. *The Quarterly Journal of Economics*, 133(2):553–609.

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

Silungwe, A., Baer, A., and Guerreiro, V. (2022). 2020 Global Stocktaking of National Accounts Statistics: Availability for Policy and Surveillance. *IMF Working Papers*.

Statistics and Data Directorate of the OECD (2020). Distributional Information in Household Income, Consumption and Saving In Line With National Accounts. Technical report.

Toda, A. A. and Walsh, K. (2015). The Double Power Law in Consumption and Implications for Testing Euler Equations. *Journal of Political Economy*, 123(5):1177–1200.

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.

Vinik, D. (2017). Is Washington bungling the Census? *Politico*.

A Further Details on Construction

In this section we provide further details on the construction of the consumption panel. First, we describe the categorization of transactions as related to consumption or not, and in case of the former how to allocate a COICOP. Second, we detail the procedure for identifying outliers in non-housing consumption that we remove from the sample. Third, we provide additional detail on the housing imputation model.

A.1 Transaction categorization

There are three main types of transaction data in the sample: card payments, direct debits, and irregular transfers. Each payment class has different associated metadata, which we use to classify transactions. We first describe how we recover information on counterparties before discussing how we allocate payments to consumption categories.

A.1.1 Extracting counterparty information

For *card payments*, we retain the full set of counterparties as potentially providing consumption services and use their Merchant Client Codes (MCCs) to categorize transactions as described below. MCCs are available for all card transactions, although in some cases they appear as '0000 - Non-categorizable' especially for online transactions. In the majority of cases, we observe the tax ID (Número de Identificación Fiscal—or NIF—in the Spanish tax system) of the counterparty. The same NIF can be associated with multiple MCCs.

For *direct debit payments*, we typically directly observe the NIF and NACE sector code associated to the counterparty for each payment. When we do not, we instead rely on a free-text 'Description' field that provides information on the counterparty. In most cases with a missing NIF, the Description indicates the counterparty is a homeowners' association.[37] We collectively assign such transactions to housing services when we categorize consumption categories below.

In the remaining cases, the Description field often contains the first 17 letters of the counterparty's name. We attempt to search for the name in two auxiliary datasets. The first is the *Sistema de Análisis de Balances Ibéricos* (SABI) database which contains financial information on the near-universe of Spanish firms. When we are able to uniquely match the first 17 letters of a counterparty name to the first 17 letters of a firm in SABI, we use the retrieved NIF and NACE codes. Failing this, the second database we attempt to link to via string match is a BBVA internal database of all corporate clients. If we obtain a unique match, we use the associated NIF/NACE recorded by BBVA.

In some cases, the NIF of the counterparty indicates it is a private individual and not a firm.[38] We drop such transactions from consideration, unless the Description suggests a housing-association-related payment.

Irregular transfers are substantially more heterogenous than card and direct debit payments. Our overarching goal is to identify the set of payments to firms that are not related to housing rental payments, a category we treat separately.[39] Irregular transfers also contain the least immediately relevant metadata. If the counterparty is a BBVA client, one can retrieve a NIF/NACE code by linking internal client files. Otherwise, the only available information is a free-text 'Beneficiary Name' field.

As with direct debits, we first attempt to match the beneficiary name with a firm name in SABI. Unlike with cards and direct debits, the counterparty of an irregular transfer can be a private individual

[37] The Spanish term is *Comunidad de Propietarios*. The Description field either contains this or its variants, e.g. *C.P* or *cmdad prop*.

[38] We identify these because the first character is a digit not a letter.

[39] The problem of how to filter housing rental payments is addressed in the section on housing services imputation.

which creates ambiguity for matches involving personal names. We therefore exclude from the matching process beneficiary names that contain a common Spanish personal name from a list we compile. If a positive match is not obtained from SABI, we next use the same BBVA internal database of corporate clients as for direct debits. Finally, we export the top 2,000 remaining beneficiary names according to the total value of account inflows in 2019 and 2020 and manually assign a NIF and NACE where possible.[40] This manual inspection revealed 17 NIFs that are providers of consumer credit. We treat these separately from other financial firms in the assignment of consumption categories below.

A.1.2 Assigning consumption categories

Each transaction is assigned exactly one of the following categories: non-consumption-related, non-categorizable consumption, the twelve two-digit COICOP categories from table 2, or a multiproduct retailer label comprising 'Supermarkets', 'Supercenters', 'Household Electronics', 'Building Material Supplier' or 'Sporting Goods'. We describe below how purchases made at multiproduct retailers are distributed across COICOP categories.

For *card payments*, we manually define a mapping from Merchant Client Codes into categories which is available at `https://www.dropbox.com/s/hroh7azjemtdh5x/mcc_to_coicop.csv`. In defining consumption vs. non-consumption we follow national accounting principles as closely as possible. We provide further details in appendix B. One MCC identifies withdrawals at cash machines, which we treat as non-categorizable consumption.

Each *direct debit payment* is assigned one of approximately 100 labels (concepts) by an internal BBVA classification system although some of these are generic and not useful for categorization, e.g. 'regular charge'. We again create a manual mapping between concepts and consumption categories. Since the concepts are proprietary, we have not made available the manual mapping. Before applying these, we assign certain transactions separately. Direct debits marked as relating to housing association payments (see above) are assigned COICOP 4. If the counterparty is one of the 17 providers of consumer credit identified in our manual search of large receivers of irregular transfers (see above), we categorize on that basis. Twelve of the firms are providers of generic credit, so direct debits received by them are considered non-categorizable consumption. Five of the firms are providers of car-related credit, so we assign direct debits received by them as to COICOP category 7.

For every other direct debit payment, we proceed through the following sequence of steps. If a step assigns the transaction to a COICOP category, multiproduct retailer category, or non-consumption, we stop and use that assignment. Otherwise, we proceed to the next step. If the final step still produces no assignment, we treat the transaction as non-categorizable consumption. Two additional sources of information are used in these steps. First, we build a mapping from NIF to a unique consumption category via the card table. Each NIF present in the card table is assigned to whichever MCC appears most frequently in the payments it receives. The NIF is then assigned a consumption category based on our manual mapping from MCC to categories. Second, as described above, we attempt to obtain information on the NACE code of counterparties. We construct a manual mapping from NACE codes to consumption categories which is available at (`https://www.dropbox.com/s/9lcab2zajijxltn/nace_to_coicop.csv`). The sequences of steps is:[41]

1. Apply the manual mapping from concepts to consumption categories.
2. Apply the mapping from NIF to MCC to consumption category.

[40] In most cases these are retrieved from SABI. A manual match is necessary due to differences in how company names are recorded in the BBVA payments table and how they are recorded in SABI.

[41] The final two steps relate to counterparties that are not standard corporations and so would not have entries in SABI nor the BBVA internal database of corporate clients.

3. Apply the manual mapping from counterparty NACE to consumption categories.

4. NIFs that begin with 'E' or 'H' refer to a housing association, so payments received by them are given COICOP 4. Those that begin with 'R' are related to religious organizations, so payments received by them are given COICOP 12.

5. If the Description field contains text related to education (e.g. 'COLEG' or 'CEIP') assign COICOP 10.

For each *irregular transfer*, we assign a consumption category based on the following sequence of steps.[42] If a step assigns the transaction to a COICOP category, multiproduct retailer category, or non-consumption, we stop and use that assignment. Otherwise, we proceed to the next step. If the final step still produces no assignment, we treat the transaction as non-categorizable consumption. One additional source of information is used in these steps. We build a mapping from NIF to a unique consumption category via the direct debit table. Each NIF present in the direct debit table is assigned to whichever concept appears most frequently in the payments it receives. The NIF is then assigned a consumption category based on our manual mapping from direct debit concepts to categories. The sequences of steps is:

1. Apply the mapping from NIF to concepts to consumption categories.

2. Apply the mapping from NIF to MCC to consumption category.

3. Apply the manual mapping from counterparty NACE to consumption categories.

In the irregular transfer table we also observe cash withdrawals from bank tellers. These are included in non-categorizable consumption.

A.1.3 Multiproduct retailers

Our categorization procedure in some cases terminates by assigning a transaction to a multiproduct retailer. In order to determine the distribution of products sold by these establishments, we rely on official statistics. Whenever possible, we use INE's breakdown of turnover according to products sold by retailers [43]. 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 [44]. 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

[42] As with direct debits, we separately assign payments to the 17 consumer-credit firms.
[43] `https://www.ine.es/jaxi/Tabla.htm?tpx=36388&L=0`
[44] `https://data.census.gov/cedsci/table?q=EC1244&g=0100000US&tid=ECNLINES2012.EC1244SLLS1&hidePreview=false`

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.

A.2 Outlier detection

After computing total non-housing, consumption-related spending for each active customer, we remove outliers from the sample. The overall strategy is to find instances in which active customers spend far more on consumption than would be predicted from the distribution of income in the census tracts they live in.

To this end, we first define c_i for each active customer as the total consumption spending during our sample period within the BBVA universe. We then geolocate each active customer to a Spanish census tract (*sección censal*) using customer metadata. For each census tract τ, INE provides distributional information on income each year including 'Average Income per Consumption Unit' where a Consumption Unit is a weighted version of population depending on economic status (e.g. children receive a consumption unit less than 1). For each census tract, we assign 2018 average income directly from the INE files. For other years in our sample, we compute average income per census tract by taking the 2018 average and scaling by the national inflation rate.[45] Finally, we sum these average incomes for all years to arrive at a tract-level estimate of total income received on average by residents over the 2015-2021 period. Let $y_{\tau(i)}$ be the resulting income estimate for active customer i residing in tract $\tau(i)$.[46] We then regress c_i on $y_{\tau(i)}$ for all active clients. The fitted regression line is $50138.89 + 0.534 * y_{\tau(i)}$.

Next we estimate the standard deviation of income within census tracts. The INE files provide a Gini coefficient G_τ for each tract τ. Under the assumption that income is lognormally distributed within tracts, an estimate of tract-level standard deviation of income σ_τ can be obtained by inverting the formula

$$G_\tau = 2\Phi\left(\frac{\sigma_\tau}{\sqrt{2}}\right) - 1$$

where Φ is the standard normal cdf. We then use $\hat{y}_\tau \equiv y_\tau + 3 * \sigma_\tau$ as an estimate of the the level of income that places a resident three standard deviations above average tract-level income.[47]

We consider an outlier to be any active customer i whose consumption c_i is greater than $50138.89 + 0.534 * \hat{y}_\tau(i)$. Out of the 1,842,010 original active clients, 1,827,866 remain after this outlier treatment. The average threshold consumption across census tracts for defining an outlier is 454,957 (64,994 consumption per year).

A.3 Rental imputation model details

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

[45] This saves the computation time of extracting multiple years' worth of census tract files.

[46] If we cannot precisely geolocate active customers, we use coarser province-level income information also provided by INE (there are 50 provinces in Spain). We also use province-level information for residents of Navarra and the Basque Country because INE does not supply tract-level income information for these regions.

[47] σ_τ is the standard deviation of 2018 income, not cumulative income over 2015-2021. σ_τ in general understates the standard deviation of the latter.

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 A.1 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.

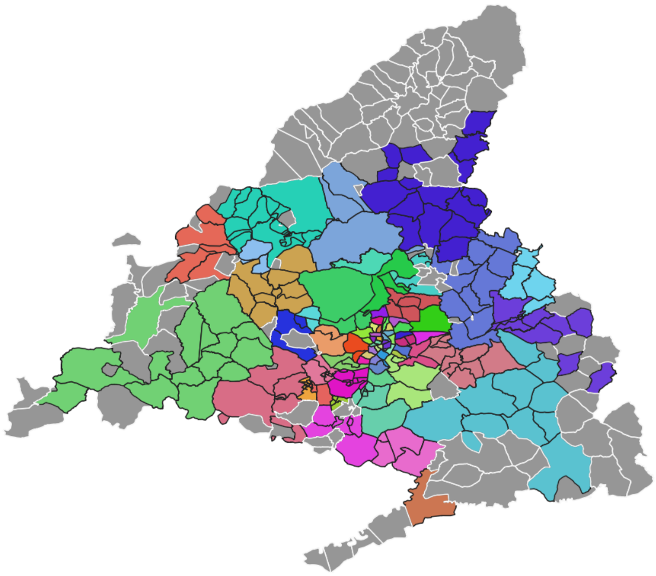


Figure A.1: Merged Regions in Madrid

B Conformity with National Statistics Definition

According to section 3.94 of the European System of National and Regional Accounts (ESA 2010):

> Final consumption expenditure consists of expenditure incurred by resident institutional units on goods or services that are used for the direct satisfaction of individual needs or wants or the collective needs of members of the community.

The choices we make in appendix A about the classification of transactions into consumption categories conform as much as possible with this definition. In this appendix, we go point-by-point through the official ESA documentation to describe how. We also acknowledge places where our data is unable to fully account for all principles.

B.1 Items included in final consumption expenditure

Section 3.95 of ESA 2010 outlines the following examples (in italics) of spending items included in final consumption expenditure. We then detail (in plain text) how each item influences our payment categorization choices.

Household final consumption expenditure includes the following examples:

a. *services of owner-occupied dwellings*;

   We impute this to all active customers. We exclude rental payments found in direct debit and irregular transfer payments to avoid double-counting.

b. *income in kind, such as*:

   1. *goods and services received as income in kind by employees*;

      These are not seen in our transaction data, and we do not account for them.

   2. *goods or services produced as outputs of unincorporated enterprises owned by households that are retained for consumption by members of the household. Examples are food and other agricultural goods, housing services by owner-occupiers and household services produced by employing paid staff (servants, cooks, gardeners, chauffeurs, etc.)*;

      We filter out self-employed clients from our sample frame. We also exclude transfers from business accounts owned by active customers in our sample frame in defining consumption-related payments. These choices mean that this type of income in kind should not be present in the sample frame. A further assumption is that the consumption patterns of the population which does not receive this income in kind does not diverge substantially from the population that does.

c. *items not treated as intermediate consumption, such as*:

   1. *materials for small repairs to and interior decoration of dwellings of a kind carried out by tenants as well as owners*;

      We define a multiproduct retailer category 'Building Material Supplier' to account for spending on this type of consumption (which is distributed across COICOPS 4 and 5 according to the procedure described in section A.1.3). Examples of retailers assigned this category are Bauhaus and Leroy Merlin.

   2. *materials for repairs and maintenance to consumer durables, including vehicles*;

      As above.

d. *items not treated as capital formation, in particular consumer durables, that continue to perform their function in several accounting periods; this includes the transfer of ownership of some durables from an enterprise to a household*;

   If the consumer durable is not bought with credit, we account for it as we would any other good. Consumption-related credit is accounted for in a variety of ways. First, we observe direct debit payments made to pay non-BBVA credit card bills. Second, certain concept labels are explicitly associated to providers of consumer credit, for example the consumer finance arms of major financial institutions (not necessarily BBVA). Third, other concept labels are associated with direct debits to large retailers. Finally, our manual search of receivers of large amounts of irregular transfers reveals 17 providers of consumer credit we account for in consumption.

e. *financial services directly charged and the part of FISIM used for final consumption purposes by households*;

   We observe payment concepts related to BBVA-provided financial services, such as card-issuance and account opening fees. We do not directly observe charges for service provision of other financial institutions.

f. *insurance services by the amount of the implicit service charge*;

   We include the payment of insurance premiums in consumption, but do not separate out the implicit service charge.

g. *pension funding services by the amount of the implicit service charge*;

   We do not observe contributions to the public pension system, and categorize transfers to private pension funds as 'non-consumption' since the bulk of these are investments rather than payments for service charges.

h. *payments by households for licences, permits, etc. which are regarded as purchases of services (see paragraphs 4.79 and 4.80)*;

   We appropriately define MCCs that denote such payments, for example port fees and parking fees.

i. *the purchase of output at not economically significant prices , e.g. entrance fees for a museum*

   We compute payments based on observed transaction values and so cannot account for divergences between posted and economically significant prices.

B.2 Items not included in final consumption expenditure

Section 3.96 of ESA 2010 outlines the following examples (in italics) of spending items not included in final consumption expenditure. We then detail (in plain text) how each item influences our payment categorization choices.

Household final consumption expenditure excludes the following:

a. *social transfers in kind, such as expenditures initially incurred by households but subsequently re-imbursed by social security, e.g. some medical expenses*;

   In Spain the government directly funds public goods rather than requiring individuals to claim back expenses.

b. *items treated as intermediate consumption or gross capital formation, such as*:

1. *expenditures by households owning unincorporated enterprises when incurred for business purposes — e.g. on durable goods such as vehicles, furniture or electrical equipment (gross fixed capital formation), and also on non-durables such as fuel (treated as intermediate consumption)*;

   We filter out self-employed clients from our sample frame. We also exclude transfers from business accounts owned by active customers in our sample frame in defining consumption-related payments.

2. *expenditure that an owner-occupier incurs on the decoration, maintenance and repair of the dwelling not typically carried out by tenants (treated as intermediate consumption in producing housing services)*;

   Unlike payments to shops that sell goods relating to basic repairs (see point c. in the previous subsection), we mark payments related to large home repair and improvement projects as non-consumption. Examples are the MCC for plumbing and heating equipment, and the NACE code 4322 for 'water, gas, heating, air conditioning installation'.

3. *the purchase of dwellings (treated as gross fixed capital formation)*;

   We exclude counterparty real estate firms (e.g. those with NACE 6831 'real estate agents'), construction firms (e.g. those with NACE 4121 'residential building construction'), and private individuals. Our outlier strategy would also remove individual who make large purchases relative to local income.

4. *expenditure on valuables (treated as gross capital formation)*;

   The distinction between valuables and jewelry (which is included in COICOP 12) is ambiguous, and we choose to include payments to jewelers in consumption.

c. *items treated as acquisitions of non-produced assets, in particular the purchase of land*;

   See item b.3 above.

d. *all those payments by households which are to be regarded as taxes (see paragraphs 4.79 and 4.80)*;

   We exclude payments associated with taxes. Examples include MCC 9311 for 'tax payment', as well as direct debit concepts for social security contributions and taxes.

e. *subscriptions, contributions and dues paid by households to NPISHs, such as trade unions, professional societies, consumers' associations, churches and social, cultural, recreational and sports clubs*;

   Separating these payments from those to associations included in COICOP 12 is challenging, so we opted to include all payments to associations as consumption.

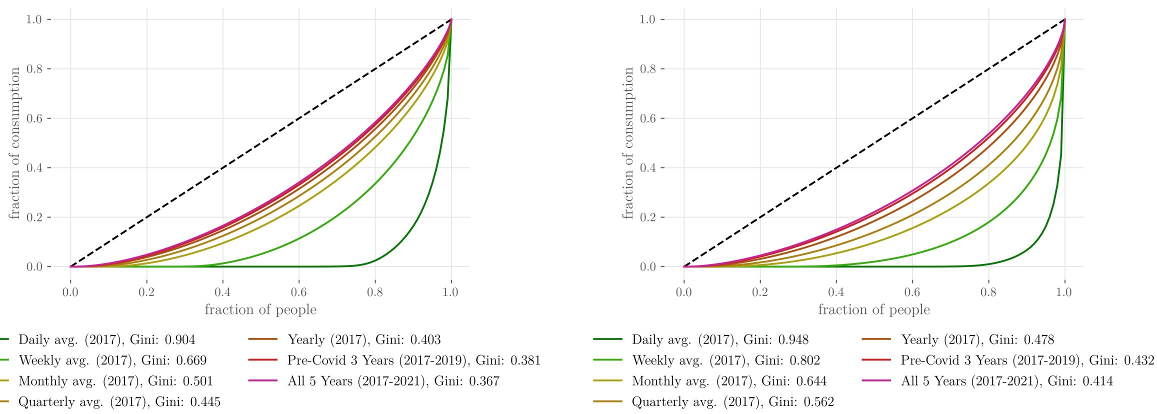
f. *voluntary transfers in cash or in kind by households to charities and relief and aid organisations.*

   We exclude such donations, e.g. MCC 1437 for 'charity contributions'.

C Distributional Analysis across Time Frequencies and Consumption Categories: Food vs. Furniture

Figure C.2 depicts Lorenz Curves and Gini Coefficients implied by the distribution of 2017 consumption for selected COICOP categories across time frequencies. Panel (a) on the left shows the results for measured inequality in Food and Non-Alcoholic Beverages (COICOPS category 1) as a function of time aggregation. Panel (b) on the right does the same for Furniture and Household Equipment (COICOPS category 5).

Notice that at very high frequencies and at this level of disaggregation across consumption categories, zero individual consumption is a pervasive feature of the data (while it is not for aggregate consumption), particularly for COICOPS category 5. This is as it should be: not every household purchases a sofa or a piece of household equipment on a given day (or week) in 2017. This pervasiveness of zeros in turn justifies the very high (near 1) Gini indexes we find at the daily frequency. Note also that the level of measured inequality is always higher - for whatever time frequency - for Furniture and Household Equipment (a luxury).



(a) Food and Non-Alcoholic Beverages (COICOPS 1) **(b)** Furniture & Household Equipment (COICOPS 5).

Figure C.2: Lorenz Curves and Gini Coefficients implied by the distribution of consumption of selected COICOP categories across time frequencies. Panel (a): Food and Non-Alcoholic Beverages. Panel (b) Furniture and Household Equipment.