

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
University College London and CEPR

Álvaro Ortiz
BBVA Research

Tomasa Rodrigo
BBVA Research

José V. Rodríguez Mora
University of Edinburgh and CEPR

December 19, 2022

Abstract

This paper provides a first proof of concept that naturally occurring transaction data, arising from the decentralized activity of millions of economic agents, can be harnessed to produce both traditional national accounts-like objects and novel representative economic statistics. 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 first organize the resulting 3 billion individual transactions by 1.8 million bank customers in a large and highly detailed representative consumption panel. Based on this, we then show 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. As a result of the richness of the transaction data, we additionally show that such data can produce novel, highly detailed distributional accounts for consumption which show larger consumption inequality than surveys suggest, particularly in the right tail. Finally, we use the panel nature of the data to offer a non-parametric analysis of individual consumption dynamics which feature a significant degree of mean reversion. Moreover, the distribution of consumption growth has thick tails.

JEL Codes: D30, E01, E21

Keywords: National Accounts, Naturally Occurring Data, Consumption, Consumption Inequality, Consumption Dynamics

*For helpful comments and suggestions we thank Manuel Arellano, Scott Baker, V.V. Chari, Giancarlo Corsetti, Diane Coyle, Lukas Freund, Fatih Guvenen, Kyle Herkenhoff, Christopher Kurz, Francois Lafond, Rocio Madera, Clara Martínez Toledano, Mariacristina De Nardi, Fabrizio Perri, Christopher Rauh, Amir Sufi, Gianluca Violante, Michael E. Waugh, and Nicolas Woloszko. We also thank seminar and conference participants at the Bank of Canada; Bank of Italy; CEPR Workshop on New Consumption Data; OECD; the 2022 European Micro-Macro-Midwest Conference; the UK Office for National Statistics; the Spanish National Statistical Office (INE); the Federal Reserve Bank of Minneapolis; the Bank of Spain; the Riksbank-Federal Reserve Board-Bank of Canada-Bank of Italy Conference on Non-traditional Data, Machine Learning and Natural Language Processing in Macroeconomics; and the US National Academies of Sciences, Engineering and Medicine's Panel on An Integrated System of U.S. Household Income, Wealth, and Consumption Data and Statistics to Inform Policy and Research. 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, allowing both for improved measurement of traditional national accounts objects and the computation of novel statistics that better reflect evolving societal concerns (see e.g. Bean 2016 and 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. While this system has produced a stable and reliable set of measurements for decades, there are mounting concerns that the system is coming under increasing strain from declining survey participation, rising costs of data collection combined with budget cuts or, sometimes, political interference. Further, these challenges are emerging precisely at a moment when national statistical agencies are facing increasing demands for more granular and higher frequency measures (European Commission 2010, Vinik 2017, AEA Committees on Economic Statistics and Government Relations 2020, Abraham 2022). These pressures facing national accounting systems are even more relevant in developing countries, and as a result there are serious measurement gaps. 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 naturally occurring transaction data —when organized systematically according to national accounting principles —can indeed recreate current national accounting objects with a high degree of accuracy and, moreover, allow one to construct novel measures of high economic interest. We do so by showing how a complete record of the financial transactions of millions of retail banking clients within a large private-sector bank (i) can be harnessed to produce a large-scale, high-quality and highly detailed consumption panel. In turn, this panel (ii) aggregates, on average, to the same level of final household consumption as reported in official national accounts, thereby overcoming the severe downward bias in aggregate consumption implied by traditional consumption surveys. It also (iii) allows, for the first time in the literature, for the construction of macro-consistent distributional national accounts for consumption and reveals that consumption inequality is substantially larger than that implied by traditional consumption surveys, particularly at the top. Finally, (iv) we exploit the panel dimension of the data, which is typically unavailable in traditional consumption surveys, to show how naturally occurring transaction data can provide novel insights into individual consumption dynamics. We find strong mean reversion and fat tails in individual consumption growth.

In short, we provide a proof of concept that naturally occurring data can, when suitably organized, largely reproduce existing national accounts for consumption and extend them beyond the current frontier of measurement. Such data therefore do not merely serve as useful proxies or indicators for national accounts, but are arguably competitive with centrally administered surveys for accurately measuring

¹Further, 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.

activity at scale and from the bottom-up, from individual consumption dynamics, through distributional national accounts, to more traditional aggregate national accounts objects.

The particular sample we use is formed of the universe of Spanish retail clients of Banco Bilbao Vizcaya Argentaria (BBVA), one of the largest banks in the world. 1.8 million customers are present in our panel which runs from 2015Q2. For this paper we stop the sample at 2021Q4 but our constructions can be updated daily. We observe all account outflows for these customers, which yields an unprecedentedly granular ledger of 3 billion individual transactions along with associated metadata. The general structure of the transaction data is shared by other banks in other countries, so the construction we adopt is, in principle, portable to many other contexts. We make publicly available material to facilitate this. Based on our data, we make four contributions.

First, we show how to construct a large, representative and highly detailed panel of household consumption. We focus on outflows comprised of all debit and credit card transactions (both online and offline), all direct recurrent debits, all one-off transfers and individual payments as well as all cash withdrawals, which we assume are spent as consumption. These observed raw expenditure outflows have no direct mapping to consumption. Some payments are explicitly defined in national accounts as not relating to consumption (e.g. transfers to savings account and investment vehicles, tax payments or large purchases for major house repairs). Moreover, housing services are a large component of consumption but are typically not observed as spending. To classify observed account outflows as consumption or non-consumption, we exploit metadata and counterparty information available for every transaction, along with the 2010 European System of Accounts principles (European Commission and Eurostat 2014). For observed consumption spending, we further use counterparty information to assign a harmonized spending category according to the Classification of Individual Consumption According to Purpose (COICOP) system. To impute the consumption of housing services, we follow European Commission recommendations to national statistical agencies and model the rental payments observed for a subset of households as a function of utility payments, location, and income and use it to generate out-of-sample predictions —i.e. imputations of consumption of housing services for owner-occupied dwellings. Finally, we 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.

Our second contribution is to form measures of aggregate national consumption. We do this by summing (with appropriate population weights) across the consumption expenditures of the 1.8 million Spanish adult residents present in our sample. We then compare the resulting series to the official quarterly Final Household Consumption series, as compiled by Spain's National Statistical Institute (INE, henceforth). Note that official consumption statistics are based on a myriad of heterogeneous data sources including household surveys, firm surveys, and administrative data combined with a statistical model for imputation and aggregation. Yet, despite these methodological differences, consumption levels line up well: our average coverage of aggregate consumption across quarters is 1.01 —i.e. our measure of aggregate consumption is, on average, within 1% of the level reported by INE in the quarterly national accounts —and the correlation in quarter-on-quarter growth rates is 0.982. There is also a tight correspondence between spending levels across COICOP categories including for housing services, a large but imputed category. Note that, unlike national accounts, our aggregate series can be formed on the day a quarter ends without delay and is not subject to revisions. It can also be easily subdivided into higher-frequency time units which increases its value for policy, particularly in the wake of large shocks.

We next compare aggregate consumption as measured by other approaches. In common with many countries, INE administers an annual Household Budget Survey (HBS) to capture the consumption patterns of households (its US analogue is the CEX). In addition to having a much smaller sample size than our panel (between 40K and 50K adults participate each year), the HBS has an average coverage

ratio across comparable quarters of only 0.82. This is consistent with repeated findings in other countries whereby expenditure surveys fail to aggregate to national accounts totals; see below for a discussion of this literature. One explanation often given is that households under-report their true spending in surveys, and we find some evidence for category-specific coverage ratios consistent with this. In contrast, naturally occurring data captures *actual* rather than *reported* behavior on a much larger sample. We then build an aggregate consumption measure based only on card transactions and cash withdrawals, in common with a large empirical literature. The coverage ratio is 0.45 and,² while there is a correlation of 0.941 in quarter-on-quarter growth rates, the series' average growth rate is upward biased with respect to official data. Hence, we additionally find that, while increasingly available card data might be competitive as a coincident indicator of high-frequency consumption growth, it does not appear to hold promise as a direct measure of aggregate consumption. As we show, this is because the consumption basket reflected in card purchases does not align well with the consumption basket implied by consumption expenditures across all means of payment.

An important implication of the result that our survey's consumption aggregates closely track both the levels and dynamics of published national accounts consumption is that the underlying microdata 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. 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, consumption inequality and its evolution are more welfare-relevant objects than income *per se*, this is an important gap.³

Our third contribution is therefore to construct a set of detailed distributional accounts 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 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 (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 Spanish HBS. 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 under-sampling (or under-reporting) of high-income/high-consumption households as hypothesized, for example, by Bosworth et al. (1991). 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.

²Ganong and Noel (2019) constructs individual consumption measures in the US largely using card spending and cash withdrawals from JP Morgan Chase (along with some electronic transactions). The coverage with respect to national accounts non-durable (durable) consumption is 0.66 (0.24), see table I of Ganong and Noel (2019).

³Arguably, 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. We also discuss contemporary 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. Our naturally occurring data, by aggregating up to the correct level of aggregate consumption, avoids the need to commit to these strong (and likely counterfactual) assumptions.

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 salient features of the distribution of individual consumption dynamics. Note that this exercise is unfeasible with most traditional consumption surveys as individuals are typically rotated out of the sample. Instead, we exploit the panel dimension of individual data to analyze its stochastic structure, documenting first a strong mean reversion which is incompatible with models relying on forms of the Permanent Income Hypothesis (as they would predict a martingale). The richness of our dataset allows us to show that this occurs not only on aggregate consumption, but also in the consumption of necessities, such as food. Second, we are also able to look in detail into the lumpy nature of consumption growth at the individual level, providing a non-parametric characterization. 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 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 thick tails in both the right and the left side of the distribution, albeit less so than those observed in Spanish individual income dynamics.

Our main contribution to the literature is to explicitly bring together two emerging and related ideas. On the one hand, there is a 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, National Academies of Sciences, Engineering, and Medicine 2018, Jarmin 2019, Abraham 2022, 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 increasingly complex needs of data users with growing demand for accurate, timely and granular measurement. This literature also invariably—and forcefully—suggests the use of naturally occurring data as a possible solution to alleviate such problems and concerns.⁴

On the other hand, there is a fast-growing literature that uses access to such unstructured data to build measures of individual spending, typically captured with credit and debit cards or with financial application data; see, e.g. Gelman et al. (2014), Baker (2018), Olafsson and Pagel (2018), Ganong and Noel (2019), and Aladangady et al. (2021). This literature was greatly accelerated during the COVID-19 pandemic due to societal demands for a high-frequency, granular tracking of the economy; see, for some early contributions, Baker et al. (2020), Cox et al. (2020), Carvalho et al. (2021), Chetty et al. (2020), and Andersen et al. (2022b). Vavra (2021) and Baker and Kueng (2022) are recent reviews taking stock of this literature. By and large, though, this literature has not sought to use unstructured data to directly

⁴In some contexts, national accounts data are actively manipulated under political pressure, particularly in less democratic countries (Martínez 2022). Cavallo and Rigobon (2016) shows how naturally occurring data can improve measurement for inflation when statistical offices are under pressure to misreport price changes. Our work suggests the potential of naturally occurring data to improve instead measurement of national accounts more broadly.

build national accounts consumption, even when attempts are made to build coincident indicators for particular consumption components.

Several features of our data make it uniquely placed to provide an extensive first proof-of-concept to show how unstructured data can be used to construct existing and novel national accounting objects. First, in particular relative to samples based on financial apps, our sample size is vastly larger. Second, we consider all account outflows, not just card spending and cash withdrawals. As discussed above, these cover only about half of final household consumption and the implied consumption basket is not well aligned with national accounts. Third, we work directly with BBVA rather than with a third-party data provider which gives us full control over how to use metadata to map transactions into consumption or non-consumption spending and into specific consumption categories. It also allows us to document in detail the procedures we adopt in a transparent and replicable way. Fourth, our data is observed over time and at the individual level, which allows us to capture not only the evolution of aggregate consumption, but also to construct distributional accounts and to map out individual consumption dynamics. Both Diamond and Moretti (2021) and Andersen et al. (2022a) analyze cross-sectional variation in spending across regions in the US and Denmark, respectively, from large-sample bank data with similar spending coverage, but without temporal nor individual dimensions.⁵

While these features may be unique in academic literature, they are not in the financial industry. In principle, our exercise could be repeated in any country with a private bank of sufficient national footprint and with a modern data collection system. Indeed, one of our goals is to show the value of doing so and to provide the guidance to make this feasible. A main implication of our results is that naturally occurring data can and, we believe, should be used much more widely in the construction of official national accounts as the literature has long speculated but not shown definitively until now.

Beyond this broad contribution, we also contribute to several other literatures. As described above, our bottom-up approach relies on constructing a large-scale, highly detailed consumption survey. Thus, this paper is related to the literature analyzing the methods, biases and shortcomings associated with 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), Koijen et al. (2014), Kreiner et al. (2014), Passero et al. (2014), Pistaferri (2015). 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) analyzing consumption inequality based on such data given biases in response rates along unobserved characteristics; heterogeneity in both the levels and dynamics of how particular consumption categories in surveys track official statistics; 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 consumption across all categories, at all frequencies and across various demographic characteristics.

Our paper also builds on the literature on distributional accounts which, although stretching back to the pioneering work of Kuznets and others, was not practically developed until the recent work of Piketty et al. (2018) and Alvaredo et al. (2021).⁶ This, in turn, has generated interest within national statistical agencies and international organizations, and provided the impetus for routine production

⁵These papers also mix naturally occurring data with administrative data to construct consumption. One of our goals is to study the properties of naturally occurring data alone. In many contexts, and in developing countries in particular, rich administrative data may be unavailable. Moreover, constructions involving administrative data must wait until such data is available which eliminates the advantage of the real-time nature of naturally occurring data.

⁶Also see Blanchet et al. (2022) which proposes a method for increasing the frequency of distributional income accounts, albeit not with naturally occurring data.

and dissemination of such accounts (see for instance Statistics and Data Directorate of the OECD 2020). The procedure for creating distributional accounts typically consists of 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, and without the need for generating synthetic data and adding, we believe, clarity and simplicity to the procedure. Moreover, we provide novel distributional accounts for consumption rather than income.

Our paper has a close relationship 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)). Meyer and Sullivan (2022) argues that consumption inequality may be a more relevant measure of welfare inequality than income inequality, but points out that the measurement error present in traditional consumption surveys makes characterizing actual consumption inequality difficult. By providing a novel consumption survey with better properties than traditional ones, we can help overcome this problem.

Finally, in our last section, we contribute to the vibrant literature on consumption dynamics and its relation to income dynamics (as in Arellano et al. (2017), Battistin et al. (2009), Toda and Walsh (2015) or Madera (2019)). In particular, we adapt part of the analysis of the dynamics of individual income recently developed by Guvenen et al. (2021),⁷ but instead use it to analyze year-on-year changes in consumption at an individual level. We believe our panel allows us to analyze this with much more depth and accuracy than previously employed. For example, we are able to measure the behavior of different consumption categories and frequencies of aggregation. We report a large degree of mean reversion and a distribution of consumption innovations with a large mass of agents enjoying very large consumption increases or suffering very large consumption drops (e.g. thick tails on both sides of the distribution).

The paper is organized as follows. Section 2 presents our data and explains in detail the procedure we use to move from raw transaction data to consumption. Section 3 demonstrates that the aggregation of our data reproduces Household Final Consumption Expenditure in Spanish national accounts with a high degree of accuracy, and that its distribution across consumption categories also moves in line with published accounts. Section 4 presents our distributional national accounts for consumption, and examines 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 panel structure of our data to offer a non-parametric characterization of consumption dynamics. Section 6 concludes.

2 Building a Naturally Occurring Consumption Survey

The fundamental object we build using BBVA data is a large-scale and detailed consumption survey. We do so by accessing the universe of BBVA retail accounts in Spain beginning in 2015Q2. This encompasses all consumer transactions, not just card transactions as in previous papers using BBVA data (Carvalho et al. 2021, García et al. 2021). To the best of our knowledge, this is the largest comprehensive spending dataset currently available for research.⁸

⁷This analysis has been further expanded in the GRID database project Guvenen et al. (2022), with Arellano et al. (2022) developing the Spanish case.

⁸See table 3 of Baker and Kueng (2022) which tabulates the scale of various consumer transaction databases, among which the sample from Carvalho et al. (2021) is largest. In this paper, we further expand the set of transactions from Carvalho et al. (2021) by incorporating non-card and imputed consumption.

Of course, in Spain, in common with other countries, the national statistics office (Instituto Nacional de Estadística, INE henceforth) already maintains a consumption survey —named the Household Budget Survey (HBS henceforth)—measuring individual and household consumption across different demographic groups, product and service types. Compared to the HBS, our raw transaction data presents two primary challenges. First, the HBS is constructed from surveying households that are chosen according to a well-defined sampling procedure.⁹ In contrast, as we show below, the population of BBVA clients deviates in important ways from the Spanish population along observable demographics. Second, the HBS records consumption rather than bank account outflows.¹⁰ These two concepts differ because certain types of account outflows should not be recorded as consumption according to national accounting principles (e.g. transfers to savings accounts). Furthermore, housing services are an important part of total consumption but are not directly observed in spending data.

In the rest of the section, we detail how we tackle these twin challenges. We first detail the underlying sample of BBVA clients for whom we compute consumption measures and then, in a second step, how we map from account outflows into consumption categories. Finally, we explain our approach to aggregating and sampling individual consumption in order to form representative statistics.

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 the next subsection) between 2015Q2 and 2021Q4, when our sample ends.¹¹ Information available about bank customers includes age, sex, and census tract of residence. For a subset of clients, and beginning in mid-2016, we observe income from wages, pensions, and government benefits as recorded in an internal BBVA table. By way of comparison, the resident adult population of Spain in 2021 was 39,177,710. However, many of these customers spend very 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.¹² Finally, note that while retail accounts do not cover corporate banking (i.e. firms), they do include self-employed individuals (*Autónomo* is the corresponding Spanish legal designation). We identify 181,918 self-employed customers. Since, for such customers, transactions may reflect the purchase of production inputs rather than consumption we also remove them from our sample. Finally, we additionally consider only account holders who are Spanish residents. Filtering for non-self employed, active Spanish resident customers results in a sample of 1,827,866 Spanish adults which forms the basis for all results henceforth.¹³ This is two orders of magnitude larger than the 40K-50K participants in a given year's Household Budget Survey. More broadly, our view is that naturally occurring data is likely to have far greater coverage than traditional surveys in most contexts.

Active customers' household structure is important for determining their consumption, but is not

⁹INE selects 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. Sampled households take part in the HBS for two years, with a staggered rotation in which half the sample is replaced every year.

¹⁰Households record their spending during a two-week period in standardized notebooks. Each purchase is assigned a classification based on the standard five-digit *Classification of Individual Consumption by Purpose* (or 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.

¹¹We observe irregular volatility in spending in some months between 2015Q2 and 2016Q4 which appears related to data ingestion but is resolved from 2017 onwards. Further note that, since BBVA transaction data is updated daily, in principle all our constructions can be produced and updated in near-real time.

¹²The focus on a fixed sample of clients echoes a similar choice in Aladangady et al. (2021) who hold fixed a set of retailers for tabulating firm sales.

¹³The 1,827,866 total is also after dropping a small number of outliers the procedure for which we describe below and more fully in the Appendix.

directly observed in the data. Appendix A.1 details how we form an estimate of active customers' household size using auxiliary information on co-signed financial contracts and location within Spain. As we discuss in the Appendix, this approach produces a distribution of household sizes largely in line with census data.

Another advantage of our sample over the HBS is that it forms a multi-year panel thus allowing us to track consumption at an individual level, which we explore in section 5. Note that households participate in the Spanish HBS for only two years. Further, in the publicly available micro-data there is no consistent ID that allows the study of individual consumption dynamics even on this short time scale. While not the focus of this paper, we also observe income (from wages, pensions, and government benefits) as recorded by BBVA for a substantial number of active clients, which yields a joint consumption and income panel.¹⁴

Figure 1 compares the distribution of geographic location of residence, age, gender, and neighborhood income for the sample of active clients relative to that of Spain's population extracted from the Spanish Census.¹⁵ While the distributions are clearly related, important discrepancies exist. BBVA active clients are over-represented in particular regions, among men, and among the middle aged. They are also more likely to live in higher-income neighborhoods. When we come to form either aggregate consumption measures or distributional representative data, we address these imbalances appropriately as we detail below. Before that, we turn to the task of identifying and classifying consumption.

2.2 From Account Outflows to Consumption

The next challenge is to convert individual transactions 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 them to a standard *Classification of Individual Consumption by Purpose* (COICOP henceforth) classification. In doing so, we closely follow national accounting principles from the European System of Accounts 2010 (European Commission and Eurostat 2014) 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.

For housing services consumption, we again follow national accounts practice and implement an imputation procedure based on observed rental payments. In particular, we estimate a simple regression model that predicts observed rental payments from household characteristics, for a subset of renter clients for which this information is available. We then use the model to impute monthly rental payments for all active customers.

2.2.1 Non-housing Consumption

The national accounts concept we seek to construct is *household final consumption expenditure*. According to section 3.94 of the European System of National and Regional Accounts (European Commission and Eurostat 2014), final consumption expenditure:

...consists of expenditure incurred...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.

¹⁴The number of active clients for whom we have an income record by month is stable at around 1,000,000.

¹⁵We use the census year 2018 for the comparison. The geographic distribution is reported at the level of the 19 primary regions in Spain known as Comunidades Autónomas. For age, we plot the conditional census distribution for those at least 18 years old since the BBVA sample contains only adults. Details of neighborhood income are provided in figure 1 notes.

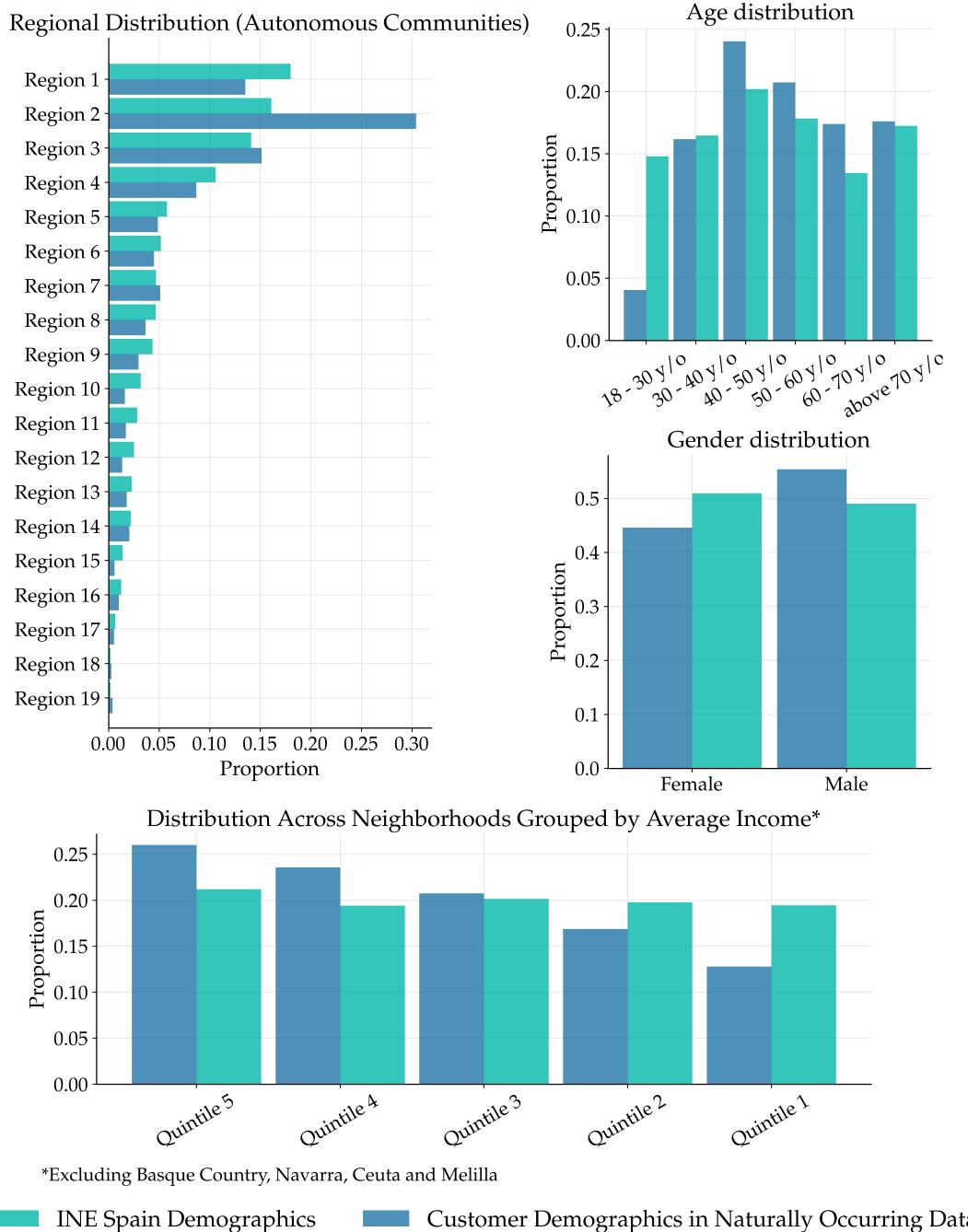


Figure 1: Demographics of Active Customers

For each BBVA customer, we observe data on age, gender, and census tract. 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, though it is very sparsely reported in Basque Country and Navarra. Ceuta and Melilla are small enclaves with very few census tracts. On the basis of average income, we group census tracts into quintiles within each of 17 Spanish regions (except for Basque Country and Navarra) 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 plot the distribution of Active Customers across quintiles.

Household final consumption expenditure is thus a sub-aggregate of final consumption, where we account solely for private individuals' expenditures but not those of non-profit institutions serving households nor government expenditures meeting 'the collective needs of the community'.

The key issue in this subsection is that, even when focusing solely on private individuals and households, many bank account outflows from the household sector—for example, transfers to investment institutions, tax payments, or major building work—do not meet the definition above. Our strategy is to filter transactions according to national accounting prescriptions in order to obtain final household consumption. To do this, appendix B provides a full list of all items that should and should not be included in consumption according to the European System of Accounts, and provides further detail on how we use transaction metadata to design appropriate filters.

An important example of this process is how we identify and discard transactions that relate to intermediate consumption by unincorporated households or transactions that reflect savings and gross capital formation by retail customers. We implement this through a number of choices. First, as discussed above we exclude all self-employed customers as well as transfers to same-named business accounts by BBVA retail customers, in order to ensure expenditures of households for business purposes are not included, as per Section 3.96 of the ESA 2010 national accounts manual (European Commission and Eurostat 2014). Second, following ESA's prescriptions, we implement a distinction between small repairs and decoration 'typically carried out by tenants or owner-occupiers' and larger gross investment housing expenses not typically undertaken by a tenant. We do this by exploiting counterparty metadata (as further described below), and then including small home repair/decoration retail outlets but excluding wholesalers of housing materials or, for example, plumbing, heating and air conditioning installation services. We further exclude real estate and residential construction firms, and any transaction denoting the purchase of dwellings and/or land. Third, we exclude savings and financial investments of BBVA retail customers, by excluding mortgage payments and transfers to savings accounts in the BBVA universe, transfers to same-named non-credit card accounts in other banks, including private pension funds, or transfers to crypto-currency wallets. Conversely, we do include all credit-card payments observed (both BBVA and non-BBVA issued) and labeled payments of consumer credit loans (either by BBVA or by specialized consumer finance arms of other financial institutions) both of which we assume to reflect consumption expenditures.

Additionally, as Appendix B further details, we are able to implement a further number of ESA's recommendations by excluding income and wealth tax payments¹⁶ and transfers to Spain's social security system, by excluding charity contributions or by including financial services' fees as well as payments by households for licences and permits. Note additionally that housing services consumption is explicitly dealt with in the next subsection. Conversely, we are unable to impute implicit service charges for insurance and pension funding services as well as correct for the purchase of consumption goods and services at 'not economically significant prices'. Rather, whenever tagged in our metadata, we include actual observed service fees. Finally, due to the nature of our transaction data, we are unable to account for in-kind income and social transfers. Overall, and while in a small number cases we cannot exactly replicate national accounting prescriptions, we are able to do so for the bulk of transactions.

After determining which transactions belong in consumption, we attribute wherever possible a COICOP classification at the two-digit level, which Table 1 displays. We do this by extensively exploiting rich metadata associated to each transaction. In particular, there are three separate transaction modes in the data—card spending, direct debits, and irregular bank transfers—and each has a distinct form of

¹⁶Note, however, that according to national accounting principles the prices of final uses do include VAT and other sales taxes and duties. Thus, we follow the OECD's recommendation in Lequiller and Blades (2014) and include these in household consumption expenditure.

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 1: 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.

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, which we assume are ultimately used for consumption. 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. Appendix A.2.3 contains further details.

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/91cab2zajijxltn/nace_to_coicop.csv.

Transfers contain the least relevant metadata, and even determining whether the counterparty is a firm requires care. Conditional on identifying the counterparty as a firms, we categorize transfer-based transactions using our manual mapping above (full details in appendix A.2).

Table 2 tabulates the number and volume of transactions made by active clients that we classify as related to consumption, broken down by transaction type. We separate cash withdrawals from other transactions, and do not include transfer payments 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.

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 2: 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.

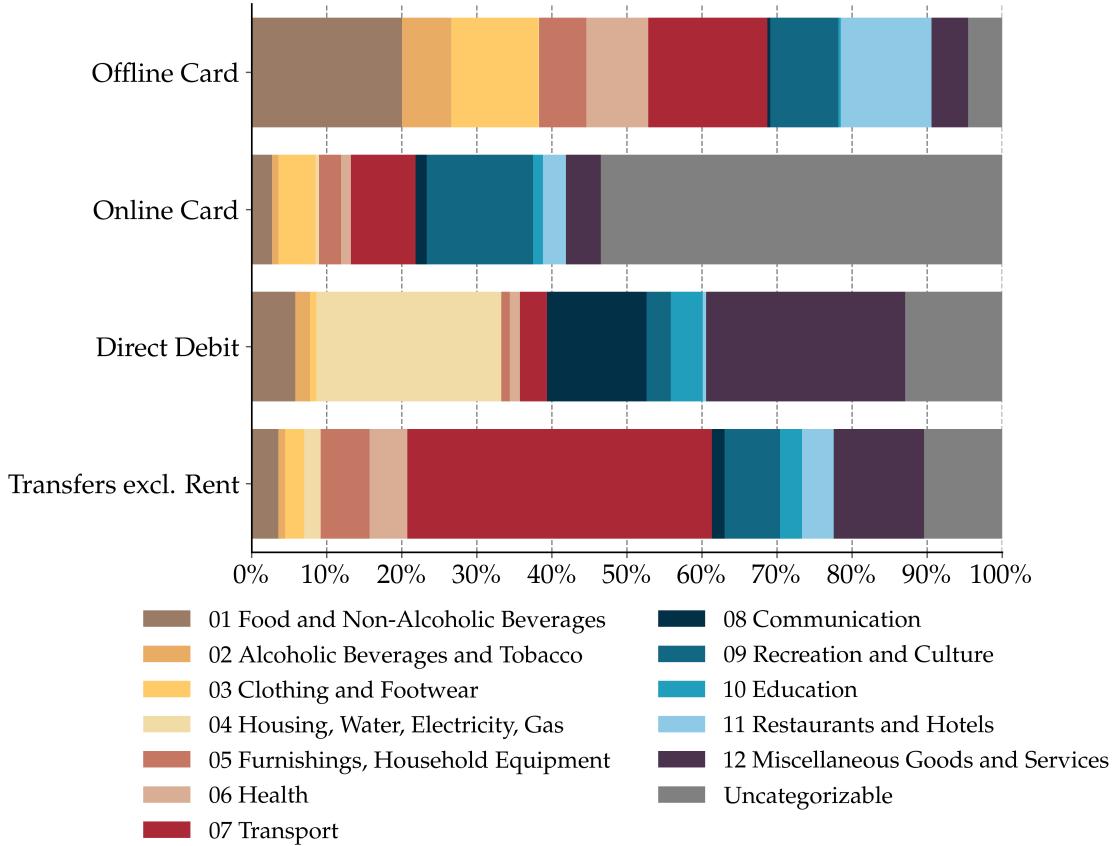


Figure 2: COICOP Shares by Payment Method

This figure shows the percentage of the total value of consumption spending across transaction mode that we manually allocated to COICOP categories based on payment codes present in metadata.

Figure 2 shows the distribution across consumption categories by payment mode.¹⁷ 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, in part due to car purchases, while utility payments mainly come via direct debits. Table 2 and Figure 2 make clear that expanding transaction beyond cards has a material impact on total measured consumption as well as its distribution across categories. We revisit this issue in the next section.

Finally, we remove active customers from the sample whose non-housing consumption is extreme

¹⁷We omit cash from the Figure as we assume its distribution across COICOP categories mimics that of offline cards. Further, as before we exclude housing rents from transfers.

relative to their census tract average income. Appendix A.3 details the procedure. This ensures that the properties of the consumption distribution analyzed below are not driven by outliers.

2.2.2 Housing Consumption Services

As per European Commission regulations (Official Journal of the European Union 2021), Spain’s INE is mandated to account for the consumption of housing services, via a two-step process. First, the European Commission states that National Statistical Institutes should prioritize a “stratification method based on actual rentals, which combines information on the housing stock, broken down by various strata, with information on actual rentals paid in each stratum.” Second, based on these strata (or cells), the European Commission then mandates extrapolating “the average actual rental per stratum (...) to all dwellings in that particular stratum.” We mimic this two-step procedure by INE with our data.

Thus, we begin by fitting the following regression model to actual observed rents.

$$\text{Rent}_{h,t} = \alpha_{s(h),t} + \beta_1 \text{Utility}_{h,t} + \beta_2 \text{Income}_{h,t} + \varepsilon_{h,t}. \quad (1)$$

Appendix A.4 provides a full description of the construction of all variables and the sample of households for which we fit (1). $\text{Rent}_{h,t}$ is the total rental payments made by members of household h in month t . Such payments can only be identified whenever individuals choose to populate an optional free-text field that describes the purpose of direct debits and irregular transfers. We form $\text{Rent}_{h,t}$ based on rental payments of all household members whether or not they are active clients.

We model $\text{Rent}_{h,t}$ as a function of three household characteristics. The first is location. $\alpha_{s(h),t}$ is a fixed effect for spatial unit s in month t . Spatial units are aggregations of postal codes such that a minimum of 30 households are present per unit. $\text{Utility}_{h,t}$ measures household utility payments, which are a proxy for house size (which INE observes but we do not). When houses are larger, they require more utility services to operate, which in turn increases utility payments. $\text{Income}_{h,t}$ is a measure of household income collected by BBVA. This proxies house quality, since higher-income households will tend to live in higher-quality housing units.

The number of households for which all this information is available over many months is relatively small compared to the total BBVA sample. There are 16,977 households for which data on all variables can be constructed for at least 70 of the 81 months, and we use these households as the sample for estimating (1). At the same time, this number of households is not far from the number that participate in the HBS each year (see Table ??).

Table 3 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.¹⁸

Finally, we use the estimated model (1) 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

¹⁸ Appendix table A.1 provides summary statistics of all variables for the baseline sample for estimating (1).

Variable	Model	Test set
Utility _{i,t}	0.0884 (0.0008)	
Income _{i,t}	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 3: Regression for Rent

We regress household-level rental payments each month on spatial unit fixed effects, household utility spending, and household income. The ‘Model’ column contains (standardized) point estimates from estimating (1) on a sample of 16,977 households for which these variable are observed in at least 70 months. The ‘Test set’ column provides goodness-of-fit information for the estimated model’s performance out-of-sample. The test set is formed by drawing a random month of rental payments for households with between 50 and 70 monthly observations.

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 Taking Stock: Outflows vs Consumption

Before proceeding, it is useful to take stock and compare how raw BBVA spending compares to consumption, as defined by our filters implementing national accounting principles. To analyse this, Table 4 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, including transfers to self). 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 4: Impact of filtering spending to consumption (2019)

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

In short, there is a large distance between raw account outflows 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 to national accounts objects.

2.3 Household and Demographic Weighting

The above steps yield both a non-housing consumption measure c_i^{NH} for each active customer i , and a housing services consumption measure $c_{h(i)}^{\text{H}}$ for each household h . As discussed earlier, the remaining challenge is to weight these consumption measures appropriately in order to ensure unbiasedness relative to Spain's adult population demographics.

First, accounting for household structure is important because part of each active client's consumption potentially benefits other (non-active) household members and vice versa. To account for this, we adopt the following weighting scheme. Let $A(i)$ ($O(i)$) be the number of active (other, respectively) customers in i 's household including himself. Household-weighted consumption is then:

$$c_i = \frac{\sum_{j \in A(i)} c_j^{\text{NH}} + c_{h(i)}^{\text{H}}}{A(i) + 0.5O(i)}. \quad (2)$$

Suppose first that a household is composed only of active customers. (2) 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, which is unobserved to us. The down-weighting by $0.5O(i)$ accounts for these competing forces. We explore the impact of using weights other than 0.5 in the next section.

Second, 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.¹⁹ Since the neighborhood 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 (2) 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 $x_{g,a,q}^{\text{INE}}$ be the total count of Spanish adults according to census data in cell (g, a, q) in 2018. Also, let $x_{\tau(t);g,a,q}^{\text{BBVA}}$ 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 is²⁰

$$c_{t;g,a,q}^W \equiv c_{t;g,a,q} \times \left(\frac{x_{g,a,q}^{\text{INE}}}{x_{\tau(t);g,a,q}^{\text{BBVA}}} \right) \quad (3)$$

From here one can form arbitrary data aggregates by summing over the appropriately weighted cells. Aggregate consumption is the sum over all cells; regional neighborhood quintile 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.

Third, while this weighting procedure accounts for demographic imbalances when producing aggregate consumption measures, it does not produce a nationally representative sample of individual level consumption. In the analysis below, for both distributional national accounts and consumption dynamics, this is necessary. In order to obtain this, for each demographic cell (g, a, q) in year τ we draw with

¹⁹Recall 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.

²⁰In principle one could use year τ census data, but $x_{g,a,q}^{\text{INE}}$ is quite stable over time so we avoid the computational cost by using a single reference year 2018.

replacement from the population of active customers $x_{g,a,q}^{\text{INE}}$ 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.

2.4 Sources of Measurement Error

While we implement a principled construction of consumption from the information available in our sample, we acknowledge that sources of measurement error and bias likely remain.

First, we treat cash withdrawals as consumption. In practice, consumers may withdraw cash for other reasons, and during crisis periods be particularly inclined to do so to avoid risk in the financial system. Moreover, we observe the day on which cash is obtained but not the day it is spent to acquire consumption goods. Allocating cash to consumption based on the date of withdrawal in high-frequency time series is therefore potentially problematic. Even where cash is used for consumption, we do not observe which consumption categories it is spent on. As described below, we use offline card spending patterns to proxy the distribution of cash spending across consumption categories. A more satisfactory approach—especially in economies where cash plays a large role in consumption—would augment the data with surveys on cash spending patterns.

Second, while we re-balance the sample using the weighting procedure in equation (3) to match the demographic distribution of the Spanish census, there may be other unobserved dimensions along which the active client sample diverges from the Spanish population. For example, active clients may on average have higher incomes than the populations within their neighborhood income quintiles. A related concern is that we do not capture the consumption of those whose economic activity takes place wholly outside the financial system. However, according to the World Bank,²¹ over 98% of adult Spaniards have a bank account, and financial inclusion levels are generally high.

Third, we do not observe consumption spending of BBVA clients mediated solely through other banks. Part of the motivation for focusing on an active client sample is to limit attention to those people who regularly use their BBVA accounts and who are thus less likely to generate significant consumption spending in other banks. We also observe many cases of active clients using their BBVA accounts to pay invoices arising from other financial institutions' products, like credit and loyalty cards. As long as spending using such products is ultimately withdrawn from a BBVA account, it is included in our measure. Still, there is invariably some amount of ‘missing’ consumption in the sample. The alignment of aggregate consumption levels according to our procedure and to national accounts (discussed in the next section) suggests the missing amount is not large, but we are unable to provide a full characterization.

A related concern is that, at the individual level, active clients may use their BBVA accounts more or less intensively during the sample period, for example if they begin to shift their spending from or to other financial institutions. This would generate larger measured consumption changes over time for these clients relative to their true consumption changes, and create the illusion of fat-tailed individual consumption growth distributions. One of the motivations for our paper is the lack of existing large-scale administrative data on individual consumption growth that one could use to study this problem. There is, however, administrative data on the distribution of individual *income* changes for Spain provided by Arellano et al. (2022) as part of the larger GRID project (Guvenen et al. 2022) on international income inequality. In Section A.5 we detail how we replicate the steps of Arellano et al. (2022) for the set of clients for which we observe income to create a distribution of income changes between 2017 and 2018

²¹<https://www.worldbank.org/en/publication/globalindex/Data>

(the earliest period for which we have two full calendar years of income). Figure 3 plots this distribution against the one-year distribution of income changes provided by the GRID project for Spain in 2015 (the closest year available in GRID to 2017) downloaded from <https://www.grid-database.org/>.

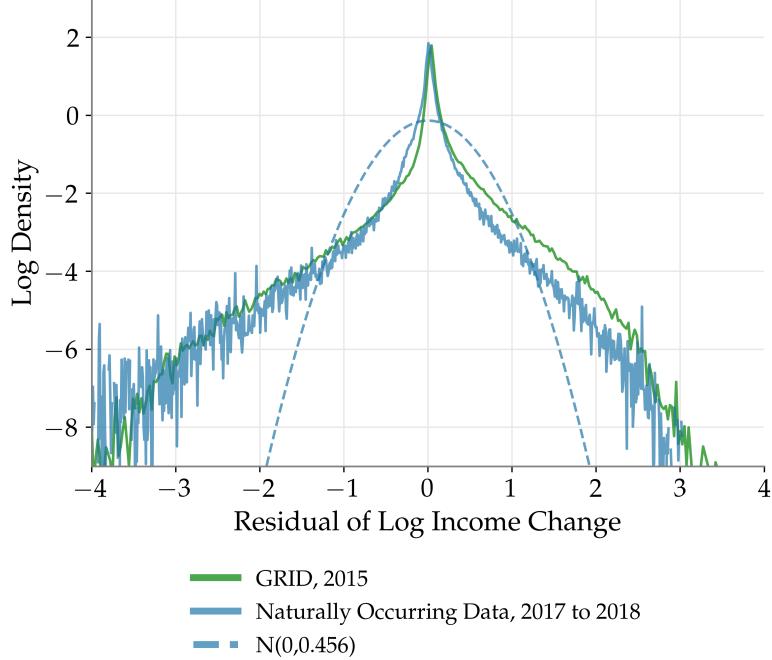


Figure 3: Log density of one-year income growth: GRID vs. BBVA

This Figure presents the log density of income growth residuals for both the GRID and BBVA databases. Both samples include only individuals in the 25-55 age range, resident outside the Basque Country and Navarra, and earning above a minimum income threshold. See Arellano et al. (2022) for full details for the construction of the GRID distribution and section A.5 for the construction of the BBVA distribution. The blue dashed curve is the normal density with the same mean and variance as the naturally occurring data distribution.

Overall, the two distributions overlap substantially and are notably heavy-tailed. (The blue dashed curve is the normal density with the same mean and variance as the naturally occurring data distribution.) One important difference is that the GRID data measures pre-tax income, while the BBVA data measures post-tax income to the extent that tax paid through retained earnings does not enter bank accounts. Accordingly, the BBVA distribution is somewhat less skewed, particularly for positive income changes.

For the key measurement error question of whether tail behavior is distorted by switching into and out of bank accounts, we find reassuring results. In line with the GRID project, we fit a regression line through the points that form the lower tail of the BBVA income distribution from [-4, -1] to form an estimate of the power law exponent that describes left tail behavior. The slope of the regression is 1.52 (which corresponds to a lower tail power law exponent of -0.52 in the CDF), compared with a slope of 1.58 reported for the GRID data.²² Similarly, right-tail regressions on the [1, 4] range produce slopes of -2.7 (which corresponds to an upper tail power law exponent of 1.7) in BBVA data and -2.44 in GRID data. Figure C.3 plots these regression lines for the BBVA data. Moreover, both distributions have similar kurtosis (GRID: 14.99; BBVA: 12.52). Overall, then, we find little evidence that consumers' switching bank accounts drives excessively fat tails in the income growth distribution. We return to this point when we discuss the tail behavior of the consumption growth distribution in section 5.

²²Throughout this paper, when reporting power law exponents, we follow the convention that these exponents are defined by the CDF of a power law distribution. This implies (by differentiation of CDF) that the implied exponents of the corresponding PDF will be given by the CDF exponent-1.

3 Measuring Aggregate Consumption

The first application of our naturally occurring consumption panel is to generate a measure of aggregate consumption per time period. To this end, we sum $c_{t;g,a,q}^W$ across all weighted cells at quarterly frequency to match the frequency at which INE publishes official statistics on aggregate final household consumption (*Gasto en Consumo Final de los Hogares*). Due to data ingestion problems that affect some transactions in the early part of the sample, we here focus on data from 2016 and beyond.

In contrast to the relatively simple and transparent procedure we use to build our aggregate consumption measure, official national accounts consumption as published by INE and other National Statistical Institutes is a complex object. It relies on dozens of underlying data series, some in levels, some in index form, and includes both household and firm surveys—with widely different samples across different goods categories—as well as administrative data obtained from government agencies. Table C.2 contains a full tabulation of the information sources that INE reports using for constructing final consumption expenditure. These data series are then passed through statistical models which aggregate them, perform imputations at various points, and satisfy adding up constraints that ensure national accounts measured via the expenditure, income or production side add up to the same value. While documentation exists with some description of the underlying models,²³ it is not made available in sufficient detail to allow us to fully replicate national accounts. The relationship between our more direct measure of consumption and national accounts is therefore not clear *ex ante* and the two alternatives are in no sense mechanically related.

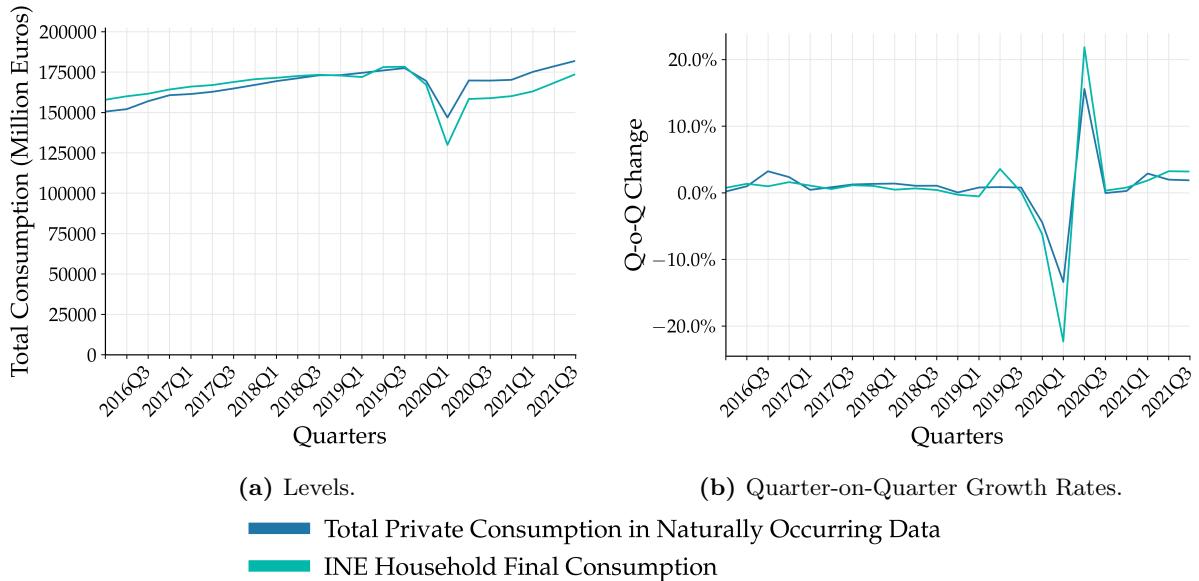


Figure 4: Aggregate Naturally Occurring Consumption vs. National Accounts

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 displays the growth rate in aggregate consumption from quarter $t - 1$ to quarter t .

Figure 4 plots the level of consumption according to INE national accounts and our approach in the left panel, and quarter-on-quarter growth rates in the right panel. Both series are seasonally adjusted using the same procedure that Eurostat officially recommends.²⁴ One first remarkable result is that the level of

²³<https://ec.europa.eu/eurostat/documents/3859598/5936013/KS-GQ-13-004-EN.PDF/3544793c-0bde-4381-a7ad-a5cfef5d8c8d0>

²⁴We use the J-Demetra open source software, made available by Eurostat and that implements the European Statistical

consumption according to both approaches is quite similar in spite of the largely different methodologies underlying them. The average coverage ratio across quarters of our level series with respect to national accounts is 1.01. Our series is thus a direct measure of national consumption, not simply a coincident indicator. Second, as the panel on the right renders clear it also serves as a highly informative indicator if one is interested in predicting consumption movements: the correlation in quarter-on-quarter growth rates is 0.982. Again, we stress that this high correlation is not generated by an explicitly formulated statistical model designed to nowcast consumption. Instead, as discussed in the previous Section, it is the byproduct of carefully considered choices for filtering raw expenditure data that we made to align with national accounts consumption concepts wherever possible. Table C.3 (in the appendix) reports more summary statistics for the series underlying Figure 4.

Of course, some of our design choices were motivated by considerations other than creating an appropriate expenditure classification system. One question is to what extent these particular choices drive the results in Figure 4. In Table C.3 we explore this issue. For example, equation (2) down-weights active client i 's consumption by $A(i) + 0.5O(i)$, and, as previously remarked, the choice of 0.5 is subjective. We therefore consider the alternative values 0.3, 0.4, 0.6, and 0.7. By construction, the level of consumption will decline monotonically in the weight, and the choice of 0.5 produces a level of consumption whose mean is most in line with that of national accounts. Even with the other weights, though, consumption levels are similar to national accounts: the coverage ratio with a weight of 0.3 (0.7) is 1.09 (0.95). The correlation with national accounts (levels and growth rates) is largely unaffected by the weight. Second, we form aggregate consumption directly from c_i rather than weighting by Spanish demographic data as in (3). This mechanically generates a consumption level far below national accounts, but again the correlation in growth rates is high. Finally, we separately drop various consumption components: imputed rent, online spending, and cash withdrawals (which may or may not be ultimately used for consumption). Both housing and cash contribute substantially to aligning levels, with online cards less important. On the other hand, dropping online spending better aligns the means of the consumption growth series. Since online spending grows faster than other forms of spending in our sample, this suggests that online spending may be over-represented in our data relative to that covered by Spain's INE. Overall, we conclude that the high correlation in growth rates is stable across design choices, while the baseline choices are important for aligning levels.

Although the coverage ratio in levels is approximately one on average, there is also variation across quarters. In 2016Q1, the coverage ratio is 0.96 and then increases quarter-by-quarter until stabilizing between 0.99 and 1.01 between 2018Q2 and 2020Q1. A notable gap opens in consumption levels in 2020Q2 at the onset of the pandemic with the coverage ratio reaching 1.13. This then declines over time and falls to 1.05 by the final quarter 2021Q4. A natural question is: what drives the opening of the gap in 2020Q2?²⁵ We explore this question below after benchmarking our measures of category-specific consumption.

We next compare how our COICOP-specific consumption measures compare to INE's national accounts COICOP disaggregation in calendar year 2019.²⁶ One issue is the presence of non-categorized consumption which creates a generic downward bias for every consumption category. By far the largest con-

System's Guidelines on Seasonal Adjustment. For details, see https://ec.europa.eu/eurostat/cros/content/download_en

²⁵The persistence in the gap once it opens is less surprising since, as discussed above, Spain's quarterly national accounts are in part built from underlying data series that take the form of indices. These inform the estimation of current growth rates which are then applied to previous levels, meaning that a given quarter's consumption level estimate will mechanically influence levels estimates several quarters into the future.

²⁶INE publishes a breakdown of consumption by COICOP as part of its annual reporting of national accounts; in Spain there are no quarterly consumption accounts by category. This breakdown is not equivalent to the breakdown provided by the HBS. We compare the distribution of consumption across categories according to national accounts and according to the HBS below.

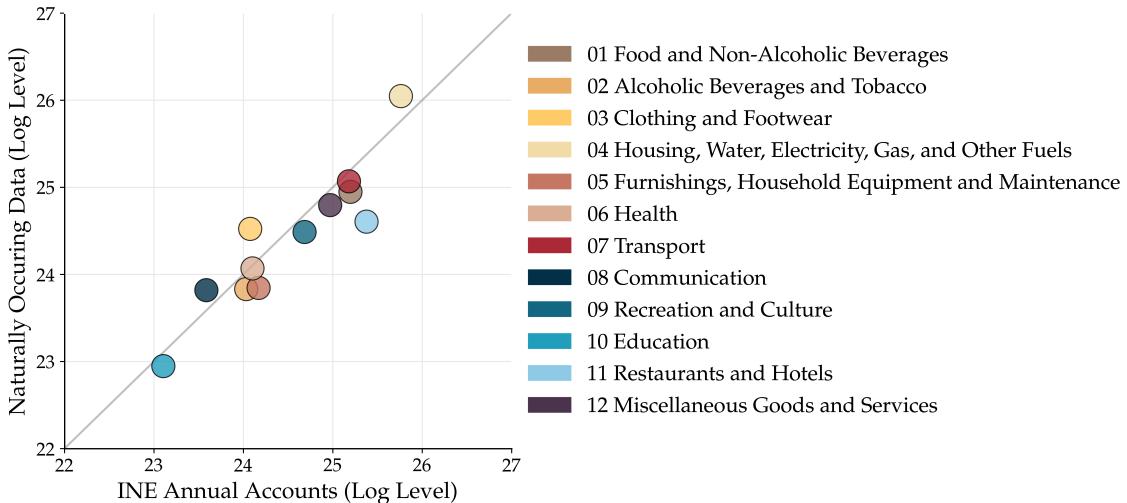


Figure 5: Distribution of Spending across COICOP Categories

This figure compares the (log) levels of consumption across COICOP categories from Spanish national accounts (x-axis) and from naturally occurring data (y-axis) for calendar year 2019. 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.

tributor to non-categorized consumption is cash, and our approach is to distribute cash across COICOP categories in proportion to offline card spending on those categories (see Figure 2). The motivation is that cash and offline card spending are substitutes and so should be spent on related items. After categorizing cash, 93% of total consumption is assigned a COICOP.²⁷ Because our aggregate measure including all consumption essentially exactly covers aggregate national accounts in 2019, this implies that the coverage ratio considering only classified categories also falls to 93%. Figure 5 plots the level of category-specific 2019 consumption according to INE’s annual accounts and our measure, in log space. As with aggregate consumption over time, there is a strong correlation (0.92) between the two measures across categories, and consumption levels also line up well despite the inherent downward bias arising from non-classified consumption.

Our COICOP-specific constructions also allow us to revisit the divergence between aggregate consumption in 2020 apparent in Figure 4 between official and naturally occurring data. From Table C.2 we identify input data sources used by INE that appear to align closely to COICOP categories, but vary in their sample size. For COICOP category 3, ‘Clothing and Footwear’, a primary source of input data in national accounts is sales declared to the Spanish tax authority by 1.1 million firms. From these declarations, we take the value of domestic sales in the textile and footwear industries.²⁸ For COICOP category 1, ‘Food and Non-Alcoholic Beverages’, INE’s input data comes from much smaller-scale surveys: a Retail Trade Index built from 12,000 firms, only some of which sell food products; and a Food Consumption Panel involving 12,500 households self-reporting on their relevant consumption. Figure 6 plots quarter-

²⁷Total directly categorized consumption in 2019 equates to 493,931 million euros, cash equates to 156,253 million euros, and remaining categorized consumption equates to 48,412 million. These numbers differ from those in Table 4 because in Table 4 we report unweighted spending totals, whereas the totals we report here are after household and demographic weighting.

²⁸Domestic textile industry sales includes sales of clothing retailers based in Spain, but also sales of Spanish firms’ intermediate textile products to downstream firms also based in Spain. We do not know the exact breakdown of textile sales into final and intermediate goods nor whether INE has access to this information. Furthermore, INE makes available the time series for sales of the textile industry but not the raw microdata. We therefore do not know exactly how many firms underlie the time series for textile sales, but believe it to be nearly comprehensive.

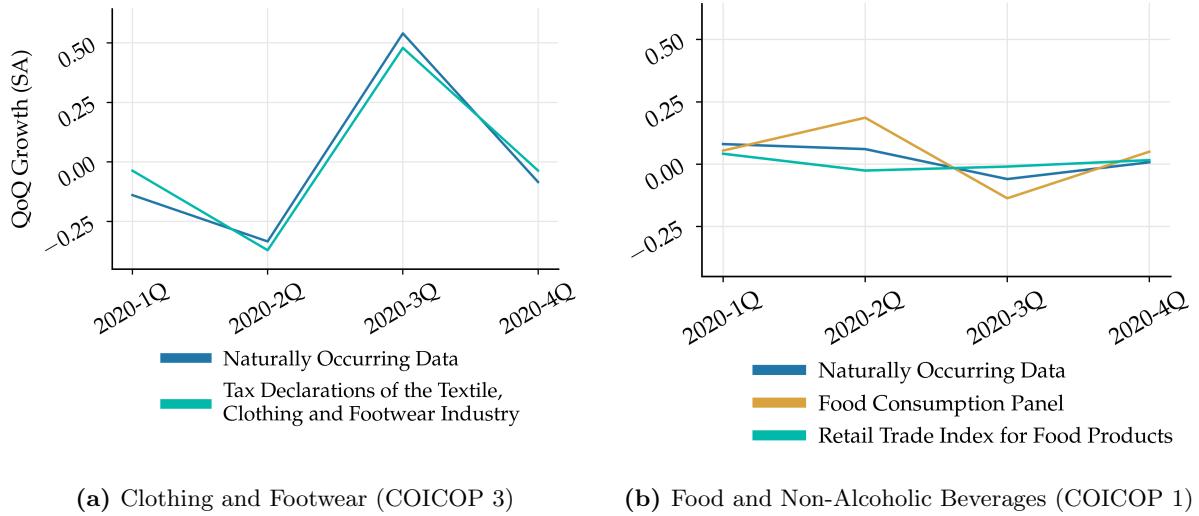


Figure 6: Comparison with Specific Data Sources used in Official National Accounts during 2020

The left panel plots quarter-on-quarter growth during 2020 in the naturally occurring measure of COICOP category 3 ‘Clothing and Footwear’ as well as quarter-on-quarter growth in the value of domestic sales for Spanish textile and footwear businesses as reported to the tax authority. The right panel plots quarter-on-quarter growth during 2020 in the naturally occurring measure of COICOP category 1 ‘Food and Non-Alcoholic Beverages’ as well as growth rates from two alternative survey-based measures of at-home food consumption. The first is the Retail Trade Index (RTI) which is based on a survey of 12,000 firms. The RTI for food products arises from the reported sales of the subpopulation of RTI firms that sell food products. The second is the Food Consumption Panel, which is administered to 12,500 individuals and includes a category for ‘food at home’ that we use to build the series. All time series are seasonally adjusted.

on-quarter growth rates for these different data sources during 2020. Note the tight correspondence between the large-sample measure of textile sales from official data and the naturally occurring measure of consumption growth in the ‘Clothing and Footwear’ category. On the other hand, the small-sample surveys for at-home food consumption diverge from each other and from our own measure —based on 1.8 million adults’ actual transactions—for ‘Food and Non-Alcoholic Beverages’ consumption. While not a definitive analysis, this exercise suggests that the 2020 divergence may arise, at least in part, from categories of consumption that INE measures with limited information, such as relatively small panels of self-reported consumption in specific categories.

Finally, our measure of consumption is nominal since we observe the total value of transactions but not price and quantity separately. A real version of consumption can be readily obtained by applying INE’s COICOP-by-Autonomous-Community monthly price indices. That is, we compute c^W by month, COICOP category, and Autonomous Community and then multiply each object by the corresponding price index.²⁹ For completeness, Figure C.4 in the appendix, plots the resulting real consumption levels and growth rates for our sample. Unsurprisingly, given our previous results, there is a tight correspondence between official and naturally occurring measures.

3.1 Comparison to Household Budget Survey

As discussed in the previous section, INE additionally runs a longstanding Household Budget Survey (HBS), the Spanish equivalent of the US Consumer Expenditure Survey. A natural question is how aggregate consumption measured by summing up reported consumption in the HBS compares to the aggregate measures reported above, both relative to INE’s published national accounts and our own naturally occurring measure.

²⁹The price indices are available at <https://www.ine.es/dynt3/inebase/es/index.htm?padre=8423&capsel=8428>. For uncategorized consumption, we apply the generic regional price index.

The top panel in appendix Figure C.5 plots the analogue of Figure 4a but for the HBS aggregate consumption measure in place of ours. The main finding is that aggregate consumption as measured by the HBS has substantially lower coverage of national accounts consumption compared to our measure. The average coverage ratio over 2016-2021 is 0.82. Again, this low coverage is in line with the aforementioned findings that, in many other countries, consumption surveys also fail to aggregate to national accounts (e.g. Barrett et al. 2014, Passero et al. 2014, Pistaferri 2015).

The bottom panel in appendix Figure C.5 adds detail to Figure 5, by displaying COICOP-specific aggregate consumption for calendar year 2019 according to the HBS. Overall, across categories, the HBS lines up less well with national accounts than even its lower aggregate coverage ratio would imply: the average coverage ratio across categories for the HBS is 0.77, while for our measure it is 0.94. Interestingly, the largest category-specific error made in the HBS is for the ‘Alcoholic Beverages and Tobacco’ category where its coverage ratio with respect to national accounts is just 0.35 (whereas ours is 0.82). This is consistent with the HBS containing measurement error due to misreporting. This category has perhaps the largest social stigma and is by implication the one that consumers are the least likely to want to reveal.

In short, the level of consumption recovered overall and category-by-category in national accounts is significantly more in line with our naturally occurring data measure than with the HBS. This suggests that a key benefit of naturally occurring data is that it captures actual rather than reported behavior, apart from producing substantially larger samples with a panel structure at a much higher frequency.³⁰

3.2 Consumption by Transaction Mode

Another important aspect of the data is that several transaction modes contribute to it. Many papers emerged during the COVID-19 crisis that tracked consumer spending using mainly card purchases (among them Andersen et al. 2022b, Chetty et al. 2020, Carvalho et al. 2021, Vavra 2021), and card spending remains the most widely-available form of payments data. However, despite increasing reliance on card data, the relationship between card-based spending measures and national accounts remains unclear in the literature, and our approach can help to clarify the link. Figure 7 is analogous to Figure 4 but only considers those payments made using card, including cash withdrawals made with cards. All other steps are maintained, e.g. the removal of non-consumption spending, conditioning on a sample of active customers, demographic re-weighting, etc. This allows us to isolate the contribution of different payment modes to national accounts coverage.

Figure 7a shows that, perhaps not surprisingly, the coverage of national accounts’ aggregate consumption level obtained with card data is limited. Across all quarters, the average coverage ratio is 0.45.³¹ As a proxy of the *level* of consumption, then, card data alone is quite poor. On the other hand, Figure 7b shows that card data holds more promise as an indicator of consumption *growth*: the correlation between QoQ growth in national accounts consumption and card-based consumption is high at 0.941, though still somewhat lower than our more comprehensive measure of aggregate consumption. Still, even for growth rates, biases appear to exist in card data with respect to national accounts. The average rate of quarter-on-quarter consumption growth in national accounts over this period is 0.7%. Our baseline measure slightly exceeds this and has an average growth rate of 0.9%, with the excess growth being

³⁰We have also explored the relationship with the Bank of Spain’s Survey of Household Finances (*Encuesta Financiera de las Familias*) which is conducted every three years and asks, among others, questions related to spending at the household level. The number of surveyed households is lower than in the HBS, with 6,413 participating in 2017, and the questions related to spending are broader than in the HBS with no COICOP breakdown. However, the survey does have a panel structure that allows households to be tracked over time. At the time of writing, though, the 2020 micro data was not available and the only year that covers our sample is 2017. In this year, total aggregate spending estimated from the survey is 295 billion EUR, which is substantially lower than aggregate 2017 consumption derived from the HBS.

³¹The final column of Table C.3 reports summary statistics pertaining to the aggregate series in figure 7.

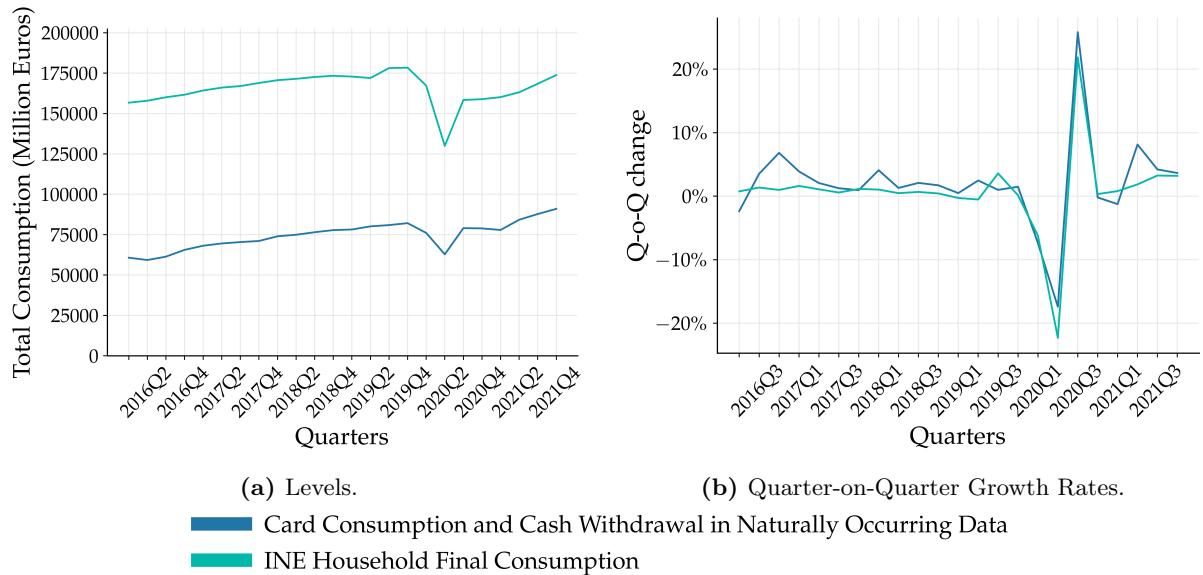


Figure 7: Aggregate Card-Based Consumption vs. National Accounts

These figures compare quarterly aggregate household consumption according to official INE data and to naturally occurring card 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 displays the growth rate in aggregate consumption from quarter $t - 1$ to quarter t .

mostly attributable to the COVID period. On the other hand, over the entire sample, the card series has an average growth rate of 2.0%, nearly three times that implied by official aggregate consumption.

These findings are not just of academic interest. In Spain, the Ministry of Economy recently developed indicators of economic growth that it views as more timely, and potentially more relevant, than traditional national accounts. These indicators are in part made up of card spending (lainformacion.com 2021), but not the full set of transactions made by Spanish consumers. Recently, traditional national accounts and the new indicators began to markedly diverge, with the new card-based indicators showing higher growth than official data.³² Our analysis suggests part of the divergence may have arisen due to an upward bias in the growth rate of card spending relative to total consumption. The larger issue is that in enormous databases there are also enormous degrees of freedom in how to process and represent data. When these choices have policy and political consequences, they are all the more important to consider carefully.

One reason for the discrepancy between card-based measures and aggregate consumption is that cards measure a distorted consumption basket relative to the aggregate. Appendix Figure C.6 compares the baseline and card-only consumption measures across COICOP categories in 2019. While coverage of any particular category for cards is generally poor since the aggregate measure has poor coverage, there is notable variation across categories. Certain categories, like ‘Clothing and Footwear’ and ‘Restaurants and Hotels’ have reasonable alignment. The reason is that—as shown in Figure 2—cards account for the bulk of consumption in these categories. On the other hand, communication has quite poor coverage with cards since most consumers pay their phone and internet bills via direct debit. Housing is hardly represented in cards since the vast majority of the category is imputed. Overall, card spending alone appears to give a distorted picture of aggregate consumption shares.³³

Note finally that, more generally, our naturally occurring consumption survey also contains, as a

³²See elEconomista.es (2022) for further details on this episode.

³³Relatedly, Table 2 in the appendix of Ganong and Noel (2019) tabulates the coverage ratio across categories of consumer spending from JP Morgan/Chase to US national accounts. Consistent with a large fraction of spending in that study arising from cards, there is substantial heterogeneity in coverage ratios across categories.

byproduct, useful information on the mode of payment. Similar to the Federal Reserve System’s Survey and Diary of Consumer Payment Choice (see Greene et al. 2017) or the ECB’s survey on payment attitudes of consumers in the euro area (European Central Bank 2020), our transaction data allows us to resolve trends in, for example, cash usage or online payments. Unlike these surveys however, our data allows for a real-time, large-scale tracking of mode-of-payment choices by consumers where, by design, the total across payment modes sums to aggregate consumption. Figure C.7 in the Appendix illustrates the value of naturally occurring data in this setting, tracking the evolution of transaction mode over our sample period. Consistent with trends documented in both the US and Europe, we observe a steady increase in the use of online card transactions to satisfy consumption needs, and a steady decline in the use of cash. At a higher frequency, we also observe a drastic decline in the use of cash during the first quarter of the pandemic. 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 seen in the previous Section, our transaction data, when organized and classified via national accounting principles, provides a high-quality aggregate match with the national accounts for consumption published by Spain’s National Statistical Institute. Importantly, this result immediately implies that our underlying micro-data 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 the dynamics of consumption inequality.

While—following the seminal work of Piketty et al. (2018)—distributional national accounts for *income* already exist for a large number of countries, to the best of our knowledge, distributional national accounts for *consumption* are virtually non-existent.³⁴ 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 the dynamics of this distribution. Specifically, Section 4.1 below provides a detailed discussion of the macro-consistent distribution of consumption across adults in Spain in 2017 and then benchmarks it against (i) existent distributional accounts for income and (ii) the distribution of consumption implied by the Spanish Household Budget Survey. In Section 4.2 we then turn our attention to the distributional accounts of consumption growth, where we 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

³⁴The 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.

weighting the micro-data as we aggregate. Here, however, we are interested in analyzing 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 in Section 2.3, where sampling weights reflect the corresponding population weights for a given cell. We also note that the removal of outliers in the consumption distribution at the census tract level (see Section 2 and Appendix A.3) implies that our estimates of consumption inequality are likely conservative.

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

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

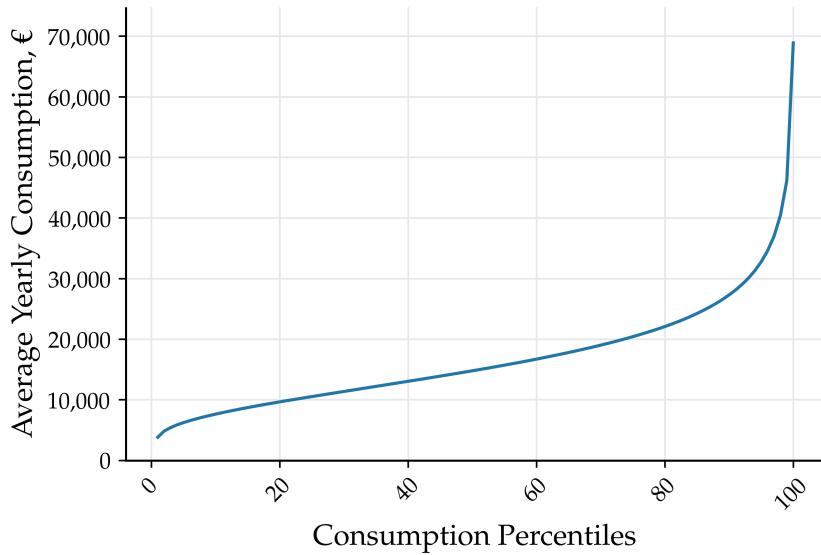


Figure 8: Distribution of Yearly Adult Consumption in 2017

This figure gives the distribution of consumption in the year 2017. Y-axis gives annual average consumption per adult (in 2017 Euros) for the corresponding X-axis percentile of consumption. To form the distribution we apply the sampling (with replacement) procedure described in Section 2.3 such that the integral of this distribution gives aggregate consumption in 2017.

Figure 8 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 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 con-

sumption 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 28,115 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 68,893 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 128,907 Euros in 2017 while the top 0.01% consumed roughly double that, at 242,490 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 2017.³⁵ 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 7,869 Euros, roughly half of the median adult consumer and 3.6 times less than the top 10% adult.

A 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.³⁶ 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 9 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 9 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 (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

³⁵Note that Figure 8 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.

³⁶Available at <https://wid.world/country/spain/>

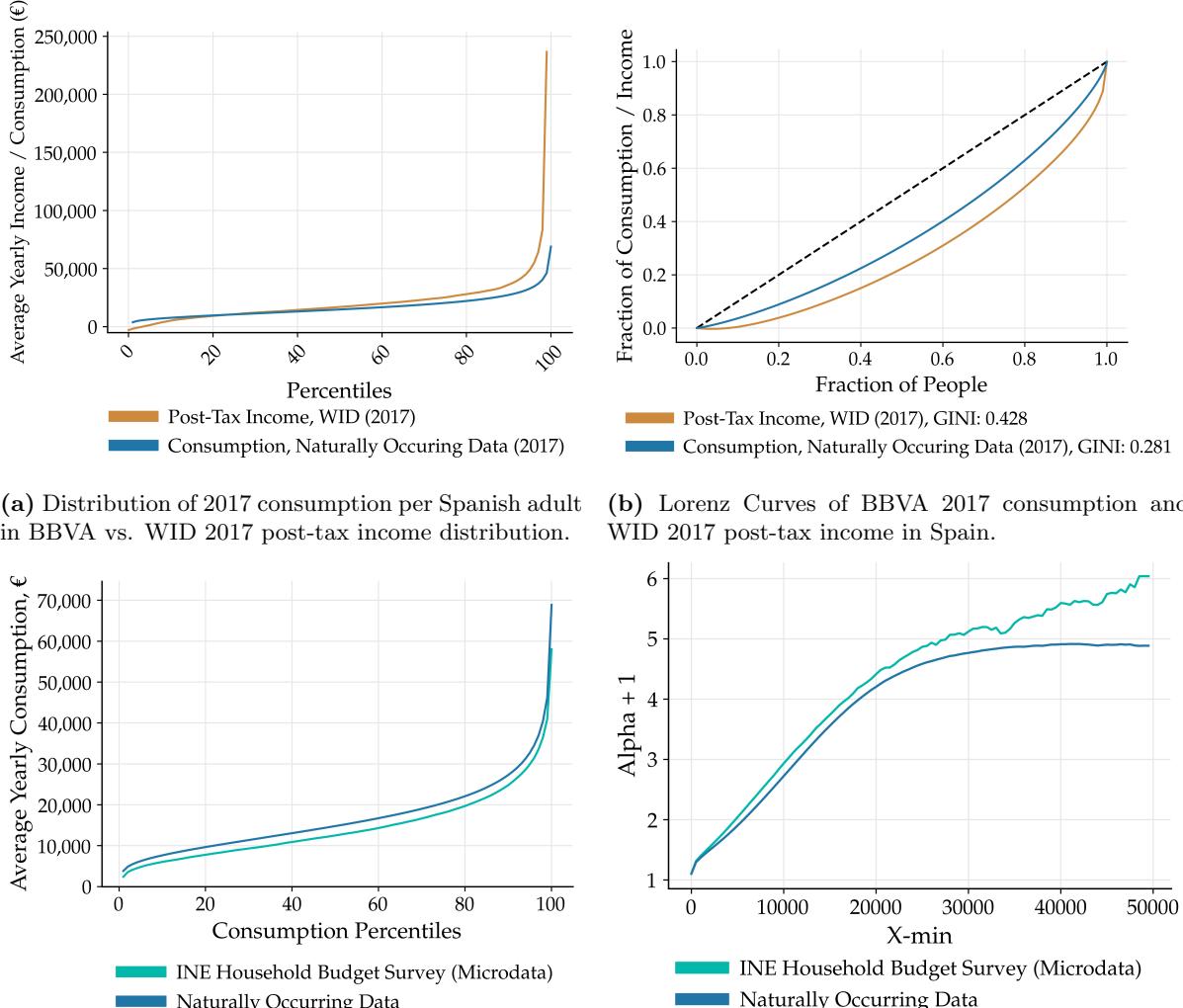


Figure 9: BBVA Consumption Inequality vs Income Inequality (Top Panels: a,b) and HBS Consumption Inequality (Bottom Panels: c,d)

The top left panel (a) of this figure compares the 2017 distribution of adult consumption from transaction data against the 2017 distribution of total post-tax income in Spain as given by the World Inequality Database, made available at <https://wid.world/country/spain/>. Y-axis gives annual average consumption or income per adult (in 2017 Euros) for the corresponding X-axis percentile of consumption or income. The top right panel (b) displays the corresponding Lorenz curves associated to the consumption and income distributions in panel (a). The bottom left panel (c) compares the 2017 adult consumption distribution from transaction data against the 2017 adult consumption distribution implied by the Spanish Household Survey. The bottom right panel (d) displays the estimated shape ($\alpha + 1$) parameter for best fit power law distribution as a function of scale parameter (x_{min}), for the two distributions in panel (c). Throughout we apply the sampling (with replacement) procedure described in Section 2.3 such that the integral of the consumption distribution from transaction data corresponds to aggregate consumption in 2017.

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.³⁷

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 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 9 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 (as anticipated by e.g. Bosworth et al. 1991, Kocherlakota and Pistaferri 2009) and, in turn, implies that some familiar inequality ratios (such as p99.9/p50) would be 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 the maximum-likelihood methods of Clauset et al. (2009) 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

³⁷The finding that consumption inequality is smaller than income inequality also suggests that the post-tax savings distribution is more unequal than the 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.

alternative parametric distributions.

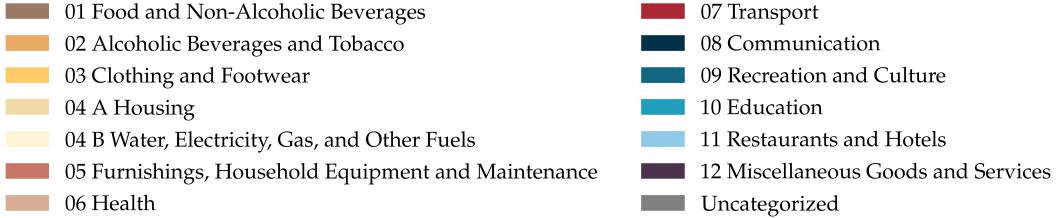
For our BBVA transaction data the maximum-likelihood estimates of the power law shape parameter is 3.91, holding at the tail of the distribution, for observations above a minimum scale of 38,000EUR. Consistently with the discussion above, comparing income and consumption distributions, this behavior 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 point estimates 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 9 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.

Finally, note that concurrent with our 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 do 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-up the household survey data, at the level of COICOP category subaggregates, and then apply the same correction factor (given by the ratio of aggregate national accounts to that implied by a consumption survey, in a given COICOP category) 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 population, meaning that potential under-reporting or sampling errors are evenly distributed among the population.”³⁸

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 imputing missing consumption in the HBS under the assumption of evenly distributed sampling errors, the proposed

³⁸Further, as a recognition of this very strong assumption, OECD-Eurostat warns that the accuracy of the implied results may be “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.”



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

Figure 10: Consumption Distribution Disaggregated by COICOP Consumption Categories

The left panel (a) of this figure gives the distribution of consumption levels in the year 2017 disaggregated by COICOP categories. Y-axis gives annual average consumption per adult (in 2017 Euros) for the corresponding X-axis percentile of consumption. To form the distribution we apply the sampling (with replacement) procedure described in Section 2.3 such that the integral of this distribution gives aggregate consumption in 2017. The right panel (b) of this figure gives the implied consumption shares of each COICOP category across the consumption distribution in 2017.

OECD-Eurostat experimental accounts will necessarily impart bias in inequality calculations.³⁹

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 10 we present results on the decomposition of consumption across major COICOP consumption categories.⁴⁰ Panel (a), on the left of Figure 10 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 10, 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

³⁹Note 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 disproportionately more luxury items. This, in turn, suggests that consumption inequality of certain types will be relatively more distorted than others by the proposed imputation.

⁴⁰In this exercise, we distribute cash across categories using offline card shares by consumption percentile.

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 subaggregates, necessities vs luxuries. We include in consumption necessities, Food and Non-Alcoholic Beverages (COICOP category 01), Alcohol and Tobacco (02) Clothing and Footwear (03), Housing and Utilities spending (04) 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.⁴¹ Clearly, given the coarse COICOP classification we work with, any such disaggregation will be fraught with some amount of misclassification. For example, the COICOP 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 top panel of Figure 11 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,

⁴¹Note 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 2,500EUR a year.

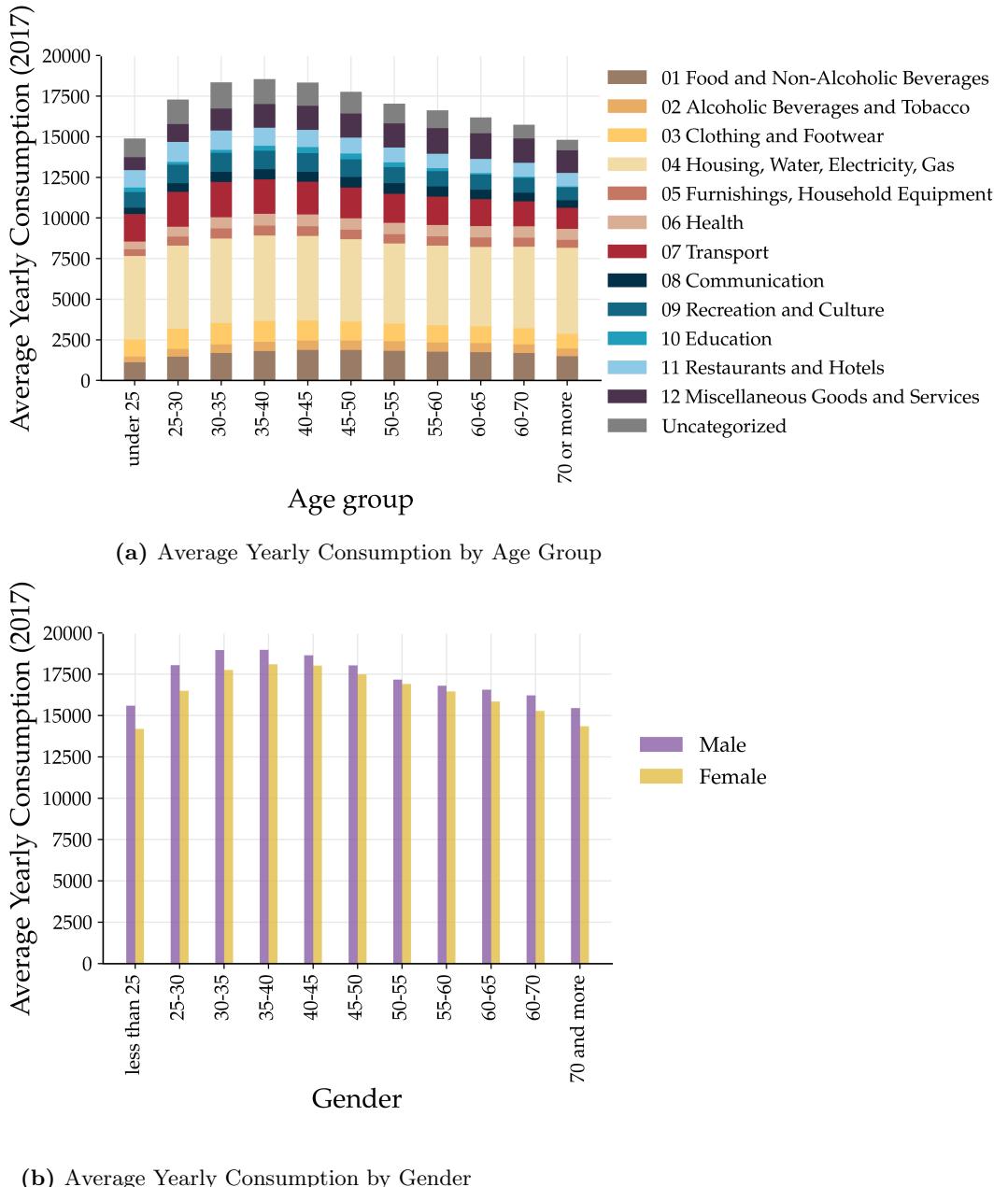


Figure 11: The Distribution of Consumption by Age and Gender.

The top panel (a) of this figure gives the distribution of consumption in the year 2017 by age group, disaggregating by COICOP categories within each age bin. The bottom panel (b) of this figure compares the life cycle profile of (total) consumption by gender. In both figures the Y-axis gives annual average consumption per adult (in 2017 Euros) for the corresponding X-axis age bin. To form the consumption distributions we apply the sampling (with replacement) procedure described in Section 2.3.

consumed around 18,500 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 than 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 bottom panel in Figure 11 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 households (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 11 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 household unit), where results are unaffected by assumptions on household-level consumption and its distribution across members and, in particular, gender.⁴² 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.⁴³

⁴²Note 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.

⁴³Quantitatively, 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

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 –two 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.⁴⁴

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.

To do this, Table 5 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 weekly, to monthly 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.⁴⁵ 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 12 complements these results by showing the associated Lorenz curves over the entire

⁴⁴To 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 find consumption inequality when there is none at a five year horizon.

⁴⁵The 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.

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

Table 5: Inequality of Consumption across Different Time Windows

This table displays two measures of consumption inequality, the Gini index and the variance of log consumption, when the distribution of consumption is measured at different frequencies. For daily frequency, and given memory constraints, we sample the daily distribution of consumption, uniformly at random, for 30 days in 2017 and then compute the corresponding average of daily inequality, for each of the two measures. For weekly, monthly and quarterly, we average each corresponding inequality measure over all weeks, months or quarters of 2017. For the measures referring to yearly, 2017-2019 or 2017-2021, we report the inequality measures implied by the distribution of total consumption over these time periods. To form the consumption distributions at every frequency above, we apply the sampling (with replacement) procedure described in Section 2.3.

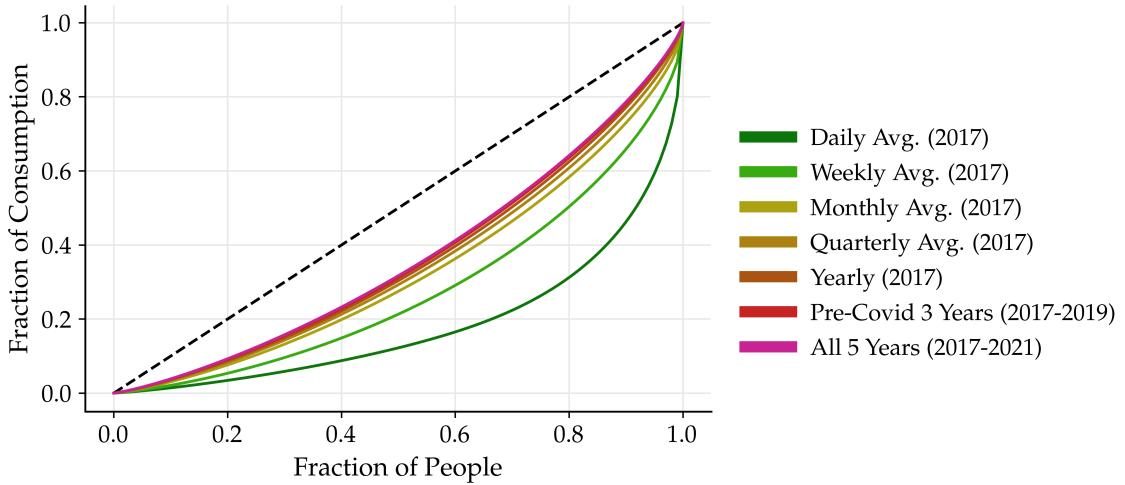


Figure 12: Lorenz Curves of the Distribution of Consumption across Different Time Frequencies.

This figure displays Lorenz curves when the underlying distribution of consumption is measured at different frequencies. For daily frequency, and given memory constraints, we sample the daily distribution of consumption, uniformly at random, for 30 days in 2017 and then compute the corresponding average of daily Lorenz Curve. For weekly, monthly and quarterly, we average the corresponding Lorenz curves over all weeks, months or quarters of 2017. For the distributions referring to yearly, 2017-2019 or 2017-2021, we plot the Lorenz curve implied by the distribution of total consumption over these time periods. To form the consumption distributions, we apply, at every frequency, the sampling (with replacement) procedure described in Section 2.3.

distribution of aggregate consumption, implied by the different sampling frequencies.⁴⁶

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 (COICOP 1) vs that of Furnishings and Household Equipment (COICOP 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 C.3 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 may be counter-factually 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 differences in consumption patterns across Spanish adults but also to document the evolution of these differences over time, while still being consistent with the level and dynamics of national accounts macro-aggregates. Thus, the distributional properties of the data allow us to go beyond Figure 4, and analyze the micro-level *distribution* of growth rates of consumption, leveraging the microstructure of our data to determine who benefits from this growth, and by how much.

Arguably, the most influential result that has arisen from the extant literature on distributional accounts is 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. While our data is unable to resolve long-term secular trends, it does include a particularly

⁴⁶We distribute imputed housing services uniformly over the days of each month. If a week falls on 2 different months, the weekly housing consumption will be proportionate to the number of days falling in one or the other month.

interesting period, the COVID-19 pandemic. As is well known, this period experienced what is one the largest consumption shocks ever recorded, with significant reallocation of consumption both across agents and consumption categories. In the remainder of this Section, we thus provide a distributional accounting exercise for consumption over the period 2017-2021.

Figure 13 gives the main result, plotting the distribution of consumption growth across the consumption distribution. Specifically, in Figure 13, and following Piketty et al. (2018), we index the percentiles of (base year) consumption on the horizontal axis, and on the vertical axis we plot the total growth in consumption, over the relevant period and for each consumption percentile.⁴⁷

To better understand Figure 13 take, for instance, the green solid line in the graph. This plots the distribution of consumption growth across all percentiles, over the entire sample period (2017-2021). It maps, to each percentile in the horizontal axis, the four-year growth rate in the total consumption for Spanish adults who were in that percentile in 2021, relative to the consumption of the adults who were in the same percentile four years earlier. To do this, we obtain the (total, across all categories) consumption of all Spanish adults in our sample (appropriately sampled, as per our discussion in Section 2.3) during the year 2017. We then order individuals by total consumption levels and assign them to percentiles in this base year. 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.⁴⁸ Finally, still focusing on the 2017-2021 period, the horizontal dotted green line in the Figure gives the implied (cross-sectional) average growth rate of nominal consumption over this 5-year period, obtained as an average across all 2017 percentiles. Thus, if for a given percentile the solid green line lies below the dotted line (as we observe for the lowest percentiles), the share of total consumption for this percentile must have decreased over the 2017-2021 period; if it lies above it, it implies that the corresponding share has increased over the same period.

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

For the pre-pandemic years, we can observe that consumption growth is larger in the top 10% of the distribution than in the bottom 10%, indicating that the p90/p10 ratio increased. 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, making it difficult to argue that overall consumption inequality increased during these “normal”, pre-COVID years. This is because, while the share of the lowest percentiles indeed increased less than at the median⁴⁹ or the average (the horizontal line), the share of consumption in the highest percentiles decreased with respect to the median, and simply kept pace with average consumption growth in Spain. Indeed, we find that the Gini Index decreased slightly from 0.281 in 2017 to 0.279 in 2019. Thus, for the three years preceding the pandemic, the overall picture is rather different from the one that appears when considering the longer-term evolution of the income distribution in Spain, as documented in Alvaredo et al. (2019): it is not the case that the top percentiles of consumption were decoupling in relation to the rest of society.⁵⁰

⁴⁷Notice we 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 the panel dimension of our data is in fact much richer than the synthetic cross-sectional data typically used to construct distributional accounts, therefore allowing for the analysis of individual (and not only aggregate) consumption dynamics.

⁴⁸As we will see in the next section that, in general, these are not the same individuals as there is mobility in the consumption distribution).

⁴⁹The p50/p10 ratio increased from 1.902 to 1.909

⁵⁰We 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.

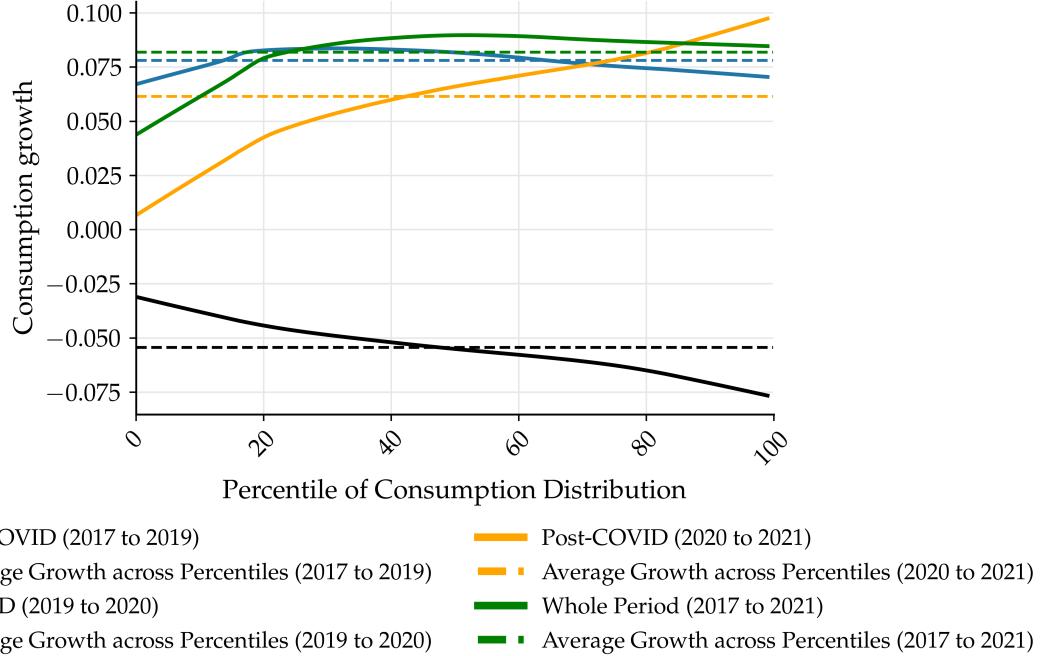


Figure 13: Growth Rate of Consumption per Percentile of the Consumption Distribution (Yearly Cross-Sections)

This figure gives the total growth rate of consumption (on the Y-axis) corresponding to each percentile of the consumption distribution (on the X-axis), across different time horizons (green line: 2017-2021 period; blue line: 2017-2019; black line 2019-2020; yellow line: 2020-2021). Solid lines are smoothed versions of the raw data where we applied a Locally Weighted Scatterplot Smoothing (LOWESS) procedure. For each time horizon, the dotted lines give the (cross-sectional) average growth in consumption in the corresponding period, obtained by averaging growth across all percentiles. To construct this figure, for each year in the sample, we assign adults in our sample to a given consumption percentile and then compute average consumption growth of that percentile from base to end year. Note that, as in Piketty et al. (2018), the identity of individuals in a given percentile is not generically the same from base to end year. To form consumption distributions for each year, we apply the sampling (with replacement) procedure described in Section 2.3.

As we saw above, we can also generate disaggregated distributional accounts at the consumption category (COICOP) level, and we can equally measure its evolution, category by category. We do so in Appendix Figure C.9, where we present the same growth allocation per percentile of total consumption, for each of the COICOP categories separately. Again focusing on pre-COVID times (in blue), it is clear that consumption inequality is stable or actually decreasing slightly across all categories except “Communication” (COICOP 08) and, to a lesser extent, “Education” (10)⁵¹.

Against this background of pre-COVID stability—or even slight decline—in consumption inequality, the COVID crisis brought with it sharp changes in the distribution of consumption growth, both across consumption percentiles and across consumption categories. The yellow and black lines in Figure 13 give a graphical representation of the distributional effects of the COVID pandemic on total consumption. Again, unlike previous measurements, they aggregate into changes in consumption in national accounts.

The black line graphs the distributional change of aggregate consumption between the end of 2019 and the end of 2020, i.e. the 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 pandemic year, consumption inequality decreased unambiguously, as the consumption shares of all percentiles above the median monotonically decreased and those of all percentiles above it increased.

⁵¹Recall that this does not include public education provision as, per national accounting principles, this is included in Government expenditure rather than private consumption. Education expenditures in our data thus reflect private expenditures over and above publicly provided education services

This finding is consistent with previous analysis of card-only expenditures in Carvalho et al. (2021), showing that the restrictions imposed by lockdowns had a larger impact on the consumption of better-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 top of the distribution, as documented in the Section 4.1.1. This is also apparent in Appendix Figure C.9, 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* across all groups during this year, but it also became more concentrated at the top of the distribution.⁵²

The yellow lines show the distributional consumption dynamics during the recovery from the pandemic (in 2021), for both aggregate consumption (in Figure 13) and, in the Appendix, for disaggregated consumption categories (Figure C.9). Perhaps not surprisingly, we see that total consumption during this recovery year increased sharply across all percentiles. However, it is also apparent that 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. Breaking down this overall pattern by COICOP category reveals that it is largely due to a recovery in the consumption of luxuries at the very top of the consumption distribution. To see this, note that in Appendix Figure C.9, the increasing profile of the yellow curves is readily apparent in the “Recreation” and “Miscellaneous” categories (and, incidentally, in “Education”, suggesting a reallocation toward private education expenditures in the year following the pandemic). 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. 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, returning to our full sample period analysis, the green lines plot the overall change from 2017 to 2021. Despite the massive COVID shock in between, the overall evolution during the whole period is not substantially 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, as in the pre-COVID period, 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.

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 (private) “Education” expenditures and “Communications”. At the same time, for the bulk of consumption categories – including both necessities and other luxury goods – the overall trends are similar to the

⁵²The implied decrease in consumption inequality in this data is also *smaller* than that implied by card data alone. This is because of two reasons: (i) as we saw above goods and services not paid with cards or cash, represent consumption commitments (e.g. utilities, rents etc, typically settled by direct debit or wire transfers) that were not as affected by COVID restrictions to economic activity and (ii) their share in total consumption is flatter across 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.

pre-COVID ones, with a general tendency for a slight decrease in the inequality of consumption in these categories.

5 Individual Consumption Dynamics

Our final contribution is to leverage the panel dimension of our naturally occurring consumption survey in order to produce a detailed analysis of the process of *individual-level* consumption growth. This goes beyond distributional national accounts—which stratifies the population into percentiles separately for each cross-section of data—and, instead, establishes salient properties of individual consumption risk while preserving desirable aggregation properties.

Recall that economic theory suggests that the consumption profile of risk-averse agents with access to well-functioning financial markets should be characterized by consumption smoothing of income shocks. In turn, this should result in not only smaller volatility (relative to income) but also in a consumption path that is unpredictable (i.e. its stochastic path should follow a martingale as per Hall 1978) and unrelated to the level of current consumption. As is well known, though, imperfections in the financial sector are bound to alter this behavior, thereby inducing some correlation of consumption and income paths. Furthermore, these imperfections can also be expected to affect households heterogeneously, yielding a rich distribution of marginal propensities to consume out of income and carrying with it important implications for economic policy.

Despite its importance, the ability to track the distribution of individual consumption over time remains limited.⁵³ This is certainly the case relative to recent developments in the literature documenting first-order features of earning dynamics (e.g. among others Arellano et al. 2017, Guvenen et al. 2021, 2022). This, in turn, is largely because existing consumption surveys suffer from the limitations discussed above: they rely on self-reported measures, with likely under-sampling of high-income/high-consumption households; limited coverage or reporting biases in specific consumption categories; and a limited panel dimension.

Against this background, this Section deploys our large-scale consumption panel to document salient features of the distribution of individual consumption growth, allowing for a non-parametric analysis of mean-reversion and higher order moments of consumption growth and extensive heterogeneity, both over the life cycle and the consumption distribution.

5.1 Mean Reversion

To illustrate the importance of conditioning on a fixed set of individuals when computing consumption growth across periods, consider Figure 14. This is similar to Figure 13 but the populations that make up each percentile are now formed based on the level of 2017 consumption and held fixed across the years for which we compute consumption growth. The most striking finding is the strong mean reversion in consumption at the individual level that Figure 13 had masked. There exists a clear negative relationship between the consumption level in 2017 and subsequent consumption growth between 2017 and 2019 (described by the solid blue curve). Those at the bottom of the consumption level distribution in 2017 increase their consumption by 30% over this period, while those above the 90th percentile in 2017 have negative consumption growth in subsequent years. This pattern then attenuates over time, with a smaller negative relationship between 2017 consumption levels and consumption growth between 2019 and 2020 (described by the solid black curve) and a flat relationship for consumption growth between 2020 and 2021 (described by the orange line).

⁵³But see, for example, Battistin et al. (2009), Toda and Walsh (2015), and Arellano et al. (2017).

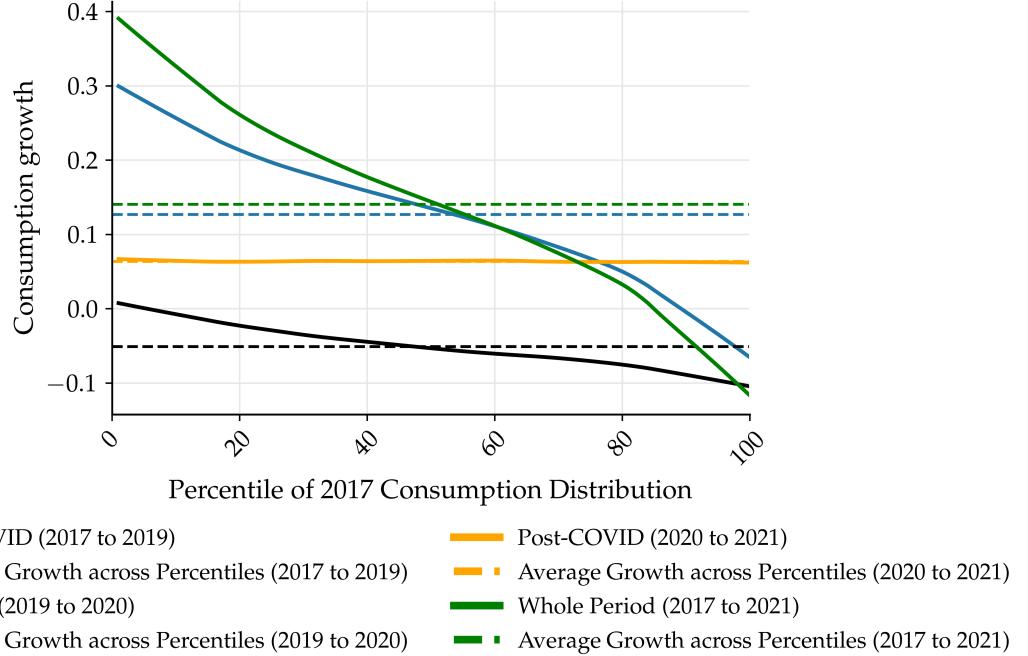


Figure 14: Growth Rates Conditional on 2017 Consumption Percentile

We first group individuals into percentiles based on total consumption in 2017. Then, for each percentile, we compute consumption growth between different years and plot these in solid lines. Dashed lines represent the average growth rates computed across percentiles. Unlike in Figure 13, the population within each percentile is fixed when computing consumption growth between any two years.

There are two natural explanations for this mean reversion pattern. First, financial market failures may move agents away from the random walk paradigm such that, to the extent that income is mean reverting, the consumption process may simply inherit features of (heterogeneous) income dynamics across the population. Second, and related to the discussion in Section 4.1.3, recall that some part of consumption is lumpy and characterized by infrequent purchases of durable (or storable) goods of high value. The consumption distribution across individuals in a given base year is therefore driven in part by the idiosyncratic process that generates such purchases. An individual whose consumption is high due to purchasing a car, for example, will tend to have lower measured consumption in future years as she enjoys its service stream without needing to replace it.⁵⁴ At the same time, the growth rates in Figure 14 are computed at annual frequency or longer, where consumption is already substantially less lumpy than at daily or weekly frequency (as seen above). While disentangling the relative importance of these two channels is beyond the scope of the paper, in Appendix Figure C.10, we further document patterns of mean reversion by disaggregating across consumption categories and considering arguably durable items (such as furniture) vs. non-durable ones (such as food and non-alcoholic beverages). While, as expected, the latter exhibits less mean reversion than the former, individual consumption growth exhibits non-trivial mean-reversion in both categories, suggesting that this first-order feature is not a result of lumpy consumption alone.

⁵⁴See Madera (2019) for a discussion of related effects.

5.2 Tail Behavior in the Distribution of Individual Consumption Growth

To explore in more detail the nature of consumption growth at the individual level, we follow the methodology of Guvenen et al. (2021)⁵⁵ in Figure 15 where we plot the full distribution of individual consumption changes from 2017 to 2018. More specifically, we record for each individual i in the bootstrap sample the value of $\log(c_{i,2018}) - \log(c_{i,2017})$ where $\log(c_{i,t})$ is total consumption measured in year t . The blue solid curve in figure 15 represents the log density of this distribution.

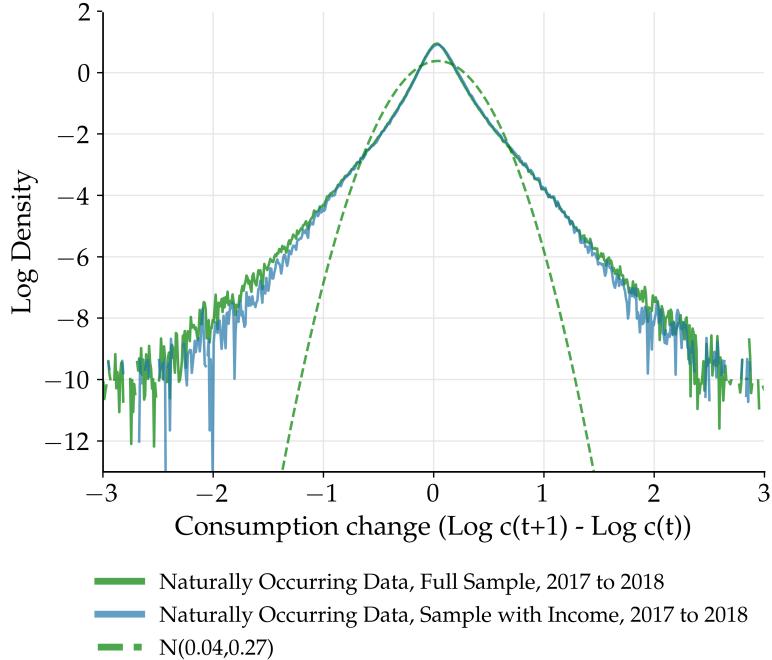


Figure 15: Log Density of Individual Consumption Growth Rates from 2017 to 2018

The solid blue curve is the log density of consumption growth rates between 2017 and 2018 for all individuals in the sample. The dashed blue curve is the density of a normal distribution with the same mean and variance as the blue density. The green curve is the log density of consumption growth rates for the subset of BBVA active clients for whom we construct the income growth distribution in Figure 3.

The most striking feature of the consumption growth distribution is its tail behavior. The dashed blue curve is a normal density with the same mean and variance as the consumption growth distribution, and clearly this does not describe well the distribution of consumption growth.⁵⁶ Given the linearity in the tails of the log density, we can, as in Section 2.4, form indicative estimates of the lower- and upper-tail power-law exponents by regressing the points that form the $[-4, -1]$ and $[1, 4]$ portions of the log density, respectively. We obtain the estimates -2.17 for the lower tail and 1.94 for the upper tail.⁵⁷ Thus, a substantial mass of individuals experiences large declines and rises in consumption year-on-year.

One concern is that part of the volatility might arise from consumers switching into and out of their BBVA accounts. To address this concern, the green empirical density in Figure 15 describes the consumption growth distribution based on the set of clients for which we compute the income growth

⁵⁵While we do not fully characterize the non-parametric distribution of consumption as Guvenen et al. (2021) does for income, we adopt part of its analysis in the rest of the section for constructing individual growth distributions.

⁵⁶In Section 4 we show that the consumption *level* distribution has heavy tails. Here the finding is instead that the consumption *growth* distribution has heavy tails.

⁵⁷Toda and Walsh (2015) uses CEX survey data from the US and also finds a power law describes the tail of the consumption growth distribution, albeit with exponents of -4 and 4 in the lower and upper tails, respectively, i.e. with less extreme behavior than we observe. Unfortunately, because the public microdata for the Spanish Household Budget Survey does not have a consistent household ID, we cannot generate an analogous survey-based consumption growth distribution for Spain to compare to the naturally-occurring-data-based one.

distribution in Figure 3 in Section 2.4. These clients have a stronger relationship with BBVA than the average active client since they receive regular income into their BBVA accounts, and for this reason are arguably more likely to maintain stable spending patterns from their BBVA accounts. Moreover, as Figure 3 in Section 2 showed, the income growth distribution constructed from these clients' BBVA income records largely matches that constructed from administrative income data. It is reassuring, then, that the consumption growth distribution for this subset of active clients matches closely the distribution for all active clients. In particular, the estimated power-law exponents for the lower and upper tails are the same as for the distribution based on all active clients.

While the tails of the consumption growth distribution are notably thick, it is interesting to compare them to those of the income growth distribution in figure 3. On both sides, the income growth distribution has thicker tails than the consumption growth distribution, so that income changes do not appear to generate one-for-one consumption changes. The divergence is particularly notable in the lower tail, suggesting an interesting asymmetry in which individual consumption reacts more to positive than to negative income shocks.

Overall, then, the stochastic process for consumption is complex, non-Gaussian, and with thick tails in both sides of the distribution, featuring a non-trivial mass of large consumption declines and a non-trivial mass of large increases. Not only we are far from observing a martingale in consumption (as we saw above), but the magnitude of the shocks provides a challenge to traditional theories of consumption smoothing.

5.3 Decomposing the Consumption Growth Distribution

Finally, we decompose the consumption growth distribution to more carefully examine its life-cycle patterns. To do so, for each year $t = 17, 18, 19, 20$, we compute total consumption per active client over years $t - 1$ and t , as well as the log consumption change from year t to $t + 1$. As in Guvenen et al. (2021), we then form moments of the consumption growth distribution by total consumption percentile and age bin (under 40, 40-49, 50-59, and over 60). Finally, we average the moments over the four years and plot the results in Figure 16.⁵⁸

Beginning with the mean of the consumption growth distribution, there are two salient features. The average growth rate of consumption decreases with age independently of the initial consumption percentile of the agent. Additionally, all age groups feature mean reversion in consumption. As alluded to above, in Appendix Figure C.10, we repeat the analysis separately by COICOP 1 (food and non-alcoholic beverages) and by COICOP 5 (furniture and household equipment). Again note that, while weaker, the mean-reverting pattern is still present in food consumption, which is a non-durable good purchased frequently. Overall this suggests that basic pattern of mean reversion documented above is a robust feature of consumption behavior.

The standard deviation is slightly higher for younger relative to older adults, but these differences are small in comparison to the differences between agents with low and high consumption. The standard deviation is roughly flat at around 0.25 through the 70th percentile, but then rises sharply and reaches between 0.45-0.5 for the highest percentiles. At the one-year horizon, those with high consumption therefore face the highest consumption risk (in addition to forming the group whose consumption is expected to decline the most). This contrasts with the patterns in Spanish income data (see Arellano et al. 2022), where the standard deviation of income growth decreases with income.

Consumption skewness also behaves differently from that of income reported by Arellano et al. (2022),

⁵⁸We performed the same exercise only with the years before the COVID pandemic and obtained the same qualitative results.

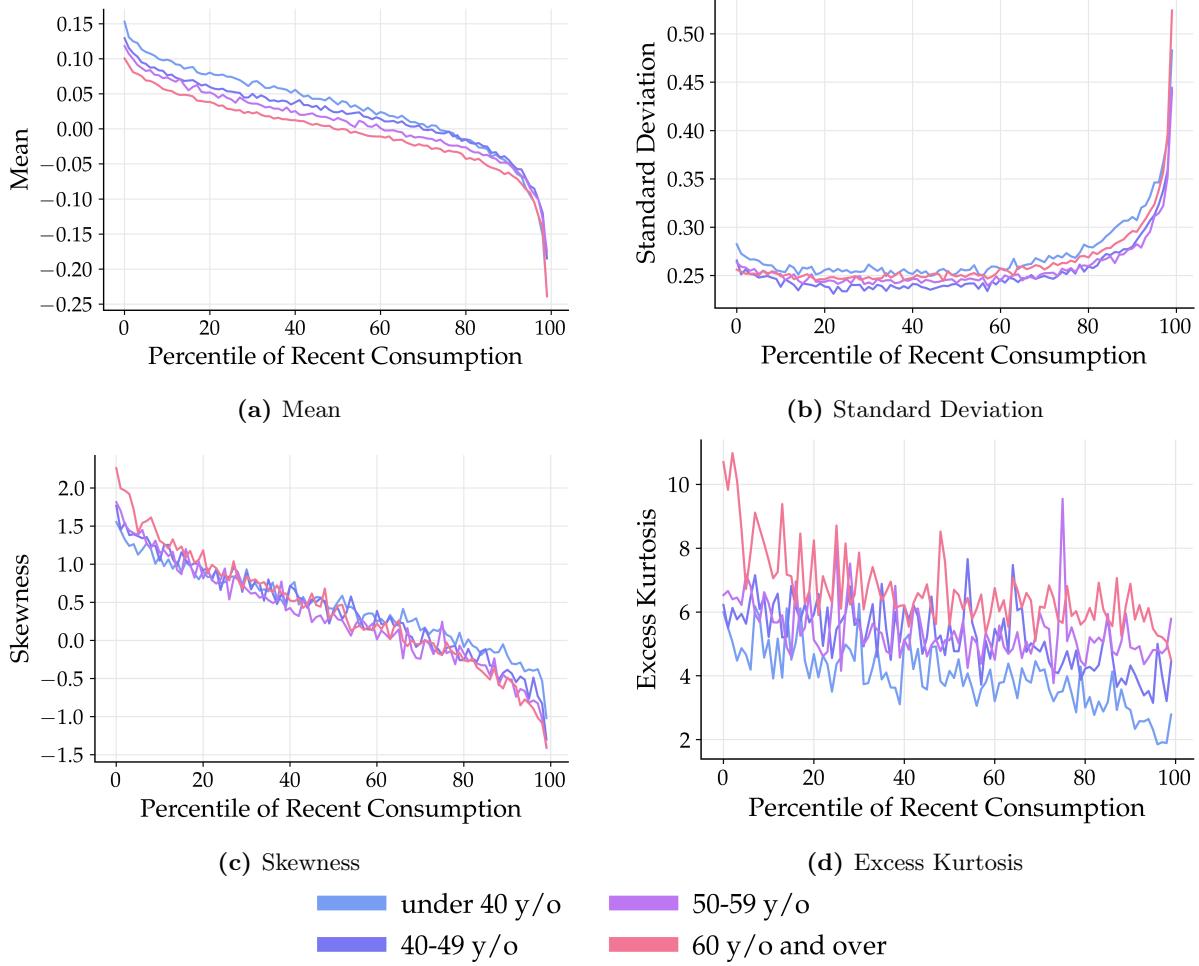


Figure 16: Moments of Consumption Growth Distribution by Consumption Percentile and Age Group

For each year $t = 17, 18, 19, 20$, we compute total consumption per active client over years $t - 1$ and t , as well as the log consumption change from year t to $t + 1$. We then form moments of the consumption change distribution by total consumption percentile and age bin (under 40, 40-49, 50-59, and over 60). Finally, we average the moments over the four years and plot the results in this Figure. Note that this construction is based on the population of active clients, not the bootstrap sample.

which documents negative skewness for the income growth distribution at all percentiles of current income. Instead, individuals with low consumption levels have a positively skewed consumption growth distribution, which implies that their subsequent rise in average consumption levels tends to be driven by large but relatively infrequent positive consumption outcomes. The opposite pattern holds for individuals with high consumption levels.

Finally, we find positive excess kurtosis (Figure 16d) in the distribution of consumption growth for almost all percentiles and age groups, suggesting fat tails. The degree of kurtosis is decreasing in consumption percentile, and higher for older agents. Again, the comparison with income is striking. Arellano et al. (2022) reports that the excess kurtosis of income growth is increasing in 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.

Overall, our results seem difficult to reconcile with traditional theories of consumption behavior and call for further research to understand these salient features of individual consumption growth. An obvious future avenue is to jointly model nonnormalities and nonlinearities in the joint process for income and consumption to better understand the channels through which our observations arise.

6 Conclusion

Especially since the COVID-19 crisis, interest in the use of naturally occurring data for the production of economic statistics has grown rapidly (e.g. Haldane and Chowla 2021) but against a background of some skepticism in the public debate on whether such data are robust enough to produce high quality, national-accounts-like, objects (e.g. The Economist 2020). Nevertheless, naturally occurring data are increasingly being used by national statistical agencies, with the Spanish government recently announcing they will be incorporated into the production of official statistics (lainformacion.com 2022). It is therefore increasingly pressing for the academic literature to establish whether and how these new data sources can be harnessed to produce robust, granular measures of the economy.

Our paper provides a first proof-of-concept that financial transaction databases arising from private banks can indeed be usefully transformed into national accounts. We do so by constructing a large-scale consumption panel that, on average, aggregates to the same level of consumption as reported in official data, thereby overcoming the severe downward bias in total consumption present in typical consumption surveys. This aggregation property then allows us to exploit the width of the panel to present the first distributional national accounts for *consumption* in the literature. Finally, we use the micro-data in the consumption panel to establish new facts about the consumption growth distribution—strong mean reversion and fat tails—that challenge traditional theories of how consumption evolves at the individual level. In other words, we both reconstruct existing national accounts objects and create new ones using naturally occurring data.

In principle, our constructions can be implemented whenever one has access to the payments records held within large private banks. We transparently detail the full steps in our processing, and make available crosswalks, to facilitate the broader use of naturally occurring data going forward. This may be particularly valuable in countries where national accounts are only sparsely populated or are otherwise unreliable.

One natural next step is to use transaction data for firms, which is also available within BBVA and other large banks, and produce national accounts for investment and input-output tables, where the latter is one of the slowest-moving official data series. Another is to generate a full consumption and income panel to examine the drivers of consumption dynamics across different groups of individuals. A vast

literature exists that addresses this question, but often the measures used for consumption are limited and not in line with national accounts. Numerous downstream applications of our consumption panel are also immediately possible. For example, Cardoso et al. (2022) use it to determine the heterogenous impact of inflationary shocks.

More broadly, we hope our work establishes that naturally occurring data can and should be treated as seriously as traditional surveys for generating measures of the economy. In their raw form, financial transactions clearly have biases and inconsistencies, but many of these can be mitigated by the rigorous application of national accounting principles. Moreover, the surveys that underlie national accounts also have important limitations. Ultimately, XXI-century national statistics are likely to fuse multiple data sources, but we are confident that, in light of our work, financial transactions' data will play a key role.

References

- Abraham, K. G. (2022). Big Data and Official Statistics. *Review of Income and Wealth*, forthcoming.
- 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., Gouille-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., Huber, K., Johannessen, N., and Straub, L. (2022a). Disaggregated Economic Accounts.
- Andersen, A. L., Hansen, E. T., Johannessen, N., and Sheridan, A. (2022b). Consumer responses to the COVID-19 crisis: Evidence from bank account transaction data. *The Scandinavian Journal of Economics*, forthcoming.
- Arellano, M., Blundell, R., and Bonhomme, S. (2017). Earnings and Consumption Dynamics: A Non-linear Panel Data Framework. *Econometrica*, 85(3):693–734.
- Arellano, M., Bonhomme, S., De Vera, M., Hospido, L., and Wei, S. (2022). Income Risk Inequality: Evidence from Spanish Administrative Records. *Quantitative Economics*, forthcoming.
- 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.
- Baker, S. R., Farrokhnia, R. A., Meyer, S., Pagel, M., and Yannelis, C. (2020). How Does Household Spending Respond to an Epidemic? Consumption during the 2020 COVID-19 Pandemic. *The Review of Asset Pricing Studies*, 10(4):834–862.

- Baker, S. R. and Kueng, L. (2022). Household Financial Transaction Data. *Annual Review of Economics*, 14(1):47–67.
- 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.
- Battistin, E., Blundell, R., and Lewbel, A. (2009). Why Is Consumption More Log Normal than Income? Gibrat's Law Revisited. *Journal of Political Economy*, 117(6):1140–1154.
- Bean, C. (2016). Independent review of UK economic statistics: Final report. Technical report.
- Blanchet, T., Saez, E., and Zucman, G. (2022). Real-Time Inequality.
- Bosworth, B., Burtless, G., and Sabelhaus, J. (1991). The Decline in Saving: Evidence from Household Surveys. *Brookings Papers on Economic Activity*, 22(1):183–256.
- Cardoso, M., Ferreira, C., Leiva, J. M., Nuño, G., Ortiz, Á., and Rodrigo, T. (2022). The heterogeneous impact of inflation on households' balance sheets. Working Papers 176, Red Nacional de Investigadores en Economía (RedNIE).
- 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.
- Cavallo, A. and Rigobon, R. (2016). The Billion Prices Project: Using Online Prices for Measurement and Research. *Journal of Economic Perspectives*, 30(2):151–178.
- 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 Kousta, 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 Tsikas, O. (2022). Distributional national account estimates for household income and consumption: Methodological issues and experimental results. Technical report, Eurostat.
- Cox, N., Ganong, P., Noel, P., Vavra, J., Wong, A., Farrell, D., Greig, F., and Deadman, E. (2020). Initial Impacts of the Pandemic on Consumer Behavior: Evidence from Linked Income, Spending, and Savings Data. *Brookings Papers on Economic Activity*, pages 35–69.
- Diamond, R. and Moretti, E. (2021). Where is Standard of Living the Highest? Local Prices and the Geography of Consumption.
- 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.

- elEconomista.es (2022). Economía trata de imponer al INE un cálculo de PIB sin aval de Eurostat. <https://www.eleconomista.es/economia/noticias/11842711/06/22/Economia-trata-de-imponer-al-INE-un-calculo-de-PIB-sin-aval-de-Eurostat.html>.
- European Central Bank (2020). Study on the payment attitudes of consumers in the euro area (SPACE). Official Publication.
- European Commission (2010). Report on Greek Government Deficit and Debt Statistics. Technical report, European Commission, Brussels.
- European Commission and Eurostat (2014). *European System of Accounts : ESA 2010*. Publications Office.
- 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.
- Ganong, P. and Noel, P. (2019). Consumer Spending during Unemployment: Positive and Normative Implications. *American Economic Review*, 109(7):2383–2424.
- 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.
- Greene, C., O'Brien, S., and Schuh, S. (2017). U. S. consumer cash use, 2012 and 2015: An introduction to the Diary of Consumer Payment Choice. Technical Report 17-6, Federal Reserve Bank of Boston.
- 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.
- Guvenen, F., Pistaferri, L., and Violante, G. L. (2022). Global Trends in Income Inequality and Income Dynamics: New Insights from GRID.
- Haldane, A. and Chowla, S. (2021). Fast economic indicators. *Nature Reviews Physics*, 3(2):68–69.
- Hall, R. E. (1978). Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. *Journal of Political Economy*, 86(6):971–987.
- Jarmín, R. S. (2019). Evolving Measurement for an Evolving Economy: Thoughts on 21st Century US Economic Statistics. *Journal of Economic Perspectives*, 33(1):165–184.
- Kocherlakota, N. and Pistaferri, L. (2009). Asset Pricing Implications of Pareto Optimality with Private Information. *Journal of Political Economy*, 117(3):555–590.
- 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.
- lainformacion.com (2021). Calviño ignora al INE y crea sus propios indicadores para vender la recuperación. <https://www.lainformacion.com/economia-negocios-y-finanzas/calvino-ignora-ine-crea-proprios-indicadores-vender-recuperacion-economia-espanola/2849874/>.
- lainformacion.com (2022). El INE modificará el cálculo del PIB tras los cambios en la cúpula del organismo. <https://www.lainformacion.com/economia-negocios-y-finanzas/ine-modifica-formula-calculo-pib-tras-cambios-direccion/2877198/>.
- Lequiller, F. and Blades, D. (2014). *Understanding National Accounts: Second Edition*. OECD.
- Madera, R. (2019). The Consumption Response to Tail Earnings Shocks. *SMU*, page 50.
- Martínez, L. R. (2022). How Much Should We Trust the Dictator's GDP Growth Estimates? *Journal of Political Economy*, 130(10):2731–2769.
- Meyer, B. D. and Sullivan, J. X. (2022). Consumption and Income Inequality in the U.S. Since the 1960s. *Journal of Political Economy*, forthcoming.
- National Academies of Sciences, Engineering, and Medicine (2018). *Reengineering the Census Bureau's Annual Economic Surveys*. The National Academies Press, Washington, DC.
- Official Journal of the European Union (2021). Commission Implementing Regulation (EU) 2021/1949 of 10 November 2021 on the principles for estimating dwelling services for the purposes of Regulation (EU) 2019/516 of the European Parliament and of the Council on the harmonisation of gross national income at market prices (GNI Regulation) and repealing Commission Decision 95/309/EC, Euratom and Commission Regulation (EC) No 1722/2005 (Text with EEA relevance).
- 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.
- The Economist (2020). Why real-time economic data need to be treated with caution.

- 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*.

ONLINE APPENDIX

A Further Details on Construction

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

A.1 Forming household groups

To infer household membership, we link 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.⁵⁹ 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.

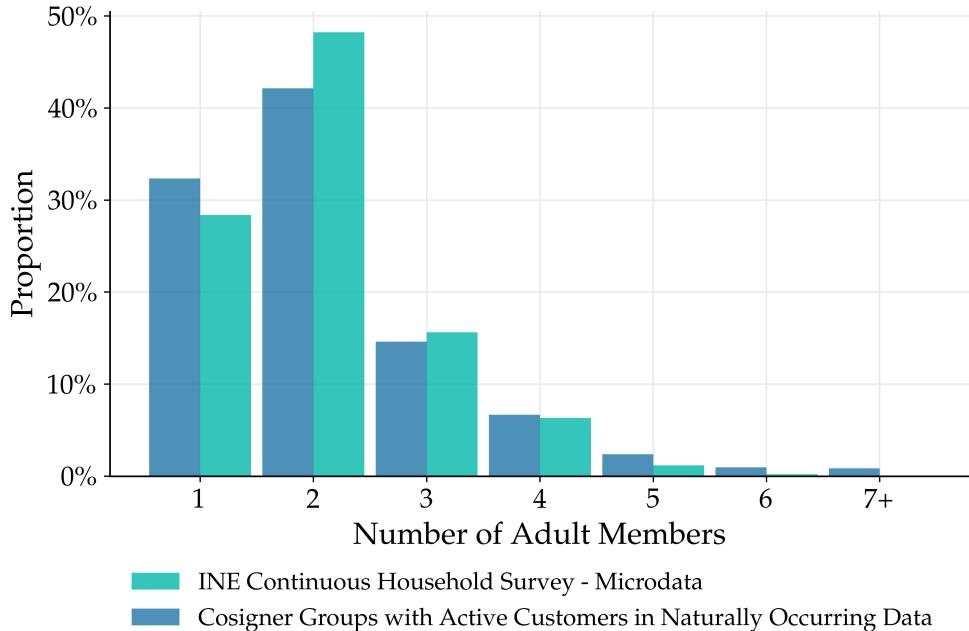


Figure A.1: Household Proxy vs Official Data

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 minors are unlikely to be BBVA customers.

Figure A.1 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

⁵⁹Postal codes are coarser than census tracts. We do not match on census tracts to avoid privacy violations.

other quite closely which suggests our grouping procedure is a viable estimate of household size in the absence of direct data.

A.2 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.2.1 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.⁶⁰ 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.⁶¹ 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.⁶² 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 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

⁶⁰The Spanish term is *Comunidad de Propietarios*. The Description field either contains this or its variants, e.g. *C.P* or *cmdad prop*.

⁶¹We identify these because the first character is a digit not a letter.

⁶²The problem of how to filter housing rental payments is addressed in the section on housing services imputation.

total value of account inflows in 2019 and 2020 and manually assign a NIF and NACE where possible.⁶³ 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.2.2 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 1, 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/91cab2zajijxltn/nace_to_coicop.csv). The sequences of steps is:⁶⁴

1. Apply the manual mapping from 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.

⁶³In 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.

⁶⁴The 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.

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.⁶⁵ 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.2.3 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 retailers⁶⁶. 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 lines⁶⁷. This source provides a broader disaggregation of retailers and products, classified based on NAICS and Product/Services Codes, respectively. These statistics allow for a more precise matching between retailer types, e.g. the ‘Sporting Goods’ label is matched with the NAICS code for ‘Specialty-line sporting goods stores’. We manually label the relevant Product/Services Codes with their corresponding ECOICOP categories.

Ultimately, we compile the corresponding product distribution for each retailer label by first relying on INE’s breakdown. If no matching retailer category is identified here—such as in the case of ‘Sporting Goods’—we fully rely on the ‘NAICS to Products’ distribution provided in US data. On the other

⁶⁵As with direct debits, we separately assign payments to the 17 consumer-credit firms.

⁶⁶<https://www.ine.es/jaxi/Tabla.htm?tpx=36388&L=0>

⁶⁷<https://data.census.gov/cedsci/table?q=EC1244&g=0100000US&tid=ECNLINES2012.EC1244SLLS1&hidePreview=false>

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.3 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.⁶⁸ 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 $\tau(i)$.⁶⁹ We then regress c_i on $y_{\tau(i)}$ for all active clients. The fitted regression line is $50138.89 + 0.534 * y_{\tau(i)}$.⁷⁰

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.⁷¹

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.4 Rental imputation model details

The construction of $Rent_{h,t}$ 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

⁶⁸This saves the computation time of extracting multiple years’ worth of census tract files, apart from the fact that such data is not available for the whole period of observation.

⁶⁹If 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.

⁷⁰We do not use BBVA’s own income records on active customers because we do not have them available for all of them.

⁷¹ 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.

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.

	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 A.1: 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 $\text{Income}_{h,t}$ 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. $\text{Utility}_{h,t}$ is 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 A.1 provides summary statistics for household-level observables in this set. Where no income or utility information is available to form a given month's record, we use the household average over all months.

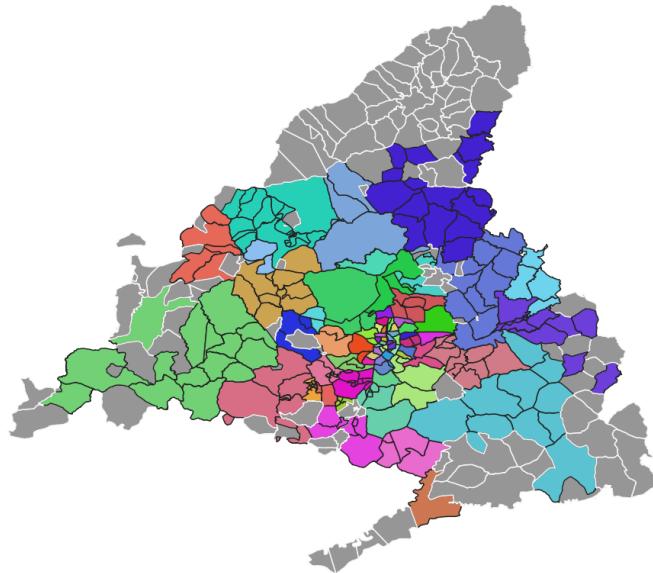


Figure A.2: Merged Regions in Madrid

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:

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.

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.

Figure A.2 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.

A.5 Construction of individual-level income growth distribution

As described in section 2.4, we construct individual-level income growth measures from BBVA data, form the distribution of these changes, and compare to the distribution of income growth as reported for Spain in the GRID project (Arellano et al. 2022, Guvenen et al. 2022). Here we detail how we formed our distributions, where all choices are made to stay as close to the GRID documentation as possible.

- Retain clients between ages 25 and 55.
- Drop clients residing in the Basque Country and Navarra (these regions have a separate tax office so income records are not available for their residents in the GRID project).
- Retain clients with a non-zero income in 2017 and 2018.
- Retain clients who in 2017 have a total income greater than 1,238 euros. The tax data that GRID uses reports whether someone is a part-time or full-time worker, and GRID retains individuals who earn more than the minimum wage for one quarter of part-time work. The 2017 minimum wage is 9.908€ yearly which we divide by 4 (to convert to quarters) and then by 2 (to convert to part time). At this stage, we have a sample of approximately 801,000 active clients.
- Scale up the sample using the bootstrap method described in section 2.3. The cell weights we use differ from those in section 2.3 because the income sample differs in its demographic distribution than the full active client sample. After bootstrapping, we have a representative sample of 25.2 million Spanish adults between 25-55 residing outside the Basque Country and Navarra.
- In each cross section of year (2017 and 2018) and gender, regress log income on a set of indicator variables for six age bins: 25-29, 30-34, 35-39, 40-44, 45-49, 50-55.
- Extract the residuals from the year by gender regressions.
- Further pare the sample to include only individuals whose 2018 income also surpasses 1,238 euros.
- Construct the distribution of the changes in the residuals for these individuals from 2017-2018.

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.2.3). 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;*

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 reimbursed 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;*

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 Additional Tables and Figures

C.1 Building a Naturally Occurring Consumption Survey: Additional Material

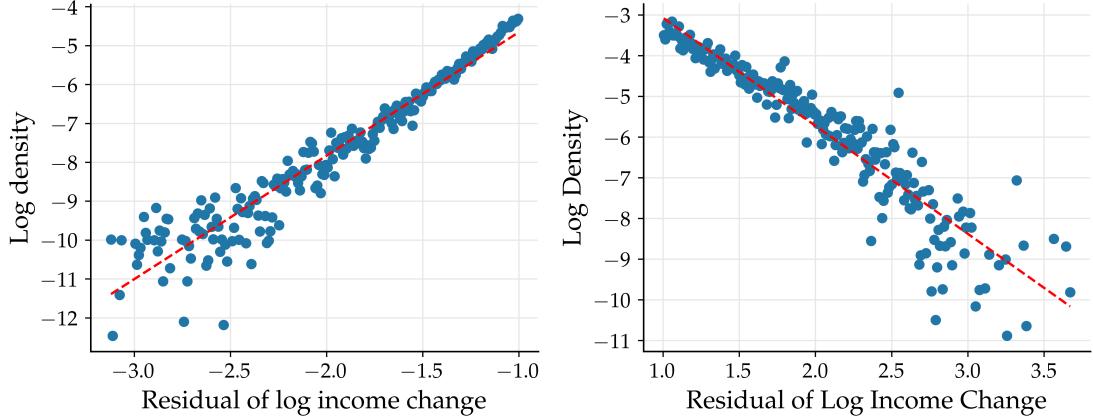


Figure C.3: Tails of 2017-18 Income Growth Distribution for BBVA Clients

The left (right) figure displays the underlying grid points for the x-axis range $[-4, -1]$ ($[1, 4]$) that are inputs into the construction of the kernel density for the BBVA-measured income growth distribution from figure 3. For each set of points, we plot the regression line that relates them. The slope for the left figure is 1.52 and for the right figure is -2.7. The equivalent slopes for GRID data are 1.58 and -2.44.

C.2 Measuring Aggregate Consumption: Additional Material

Data Source	Produced By	Data Collection	Sample Size	Consumption Categories Used For
Retail Trade Index	INE	Survey to firms	12000 firms	Food, Health, Personal Equipment, Household Equipment, Entertainment, Other Goods
Service Sector Index	INE	Survey to firms	28000 firms	Newspapers, postal services, touristic services
Sales, Employment and Salaries in Large Companies and SMEs	SEPE	Tax declarations by firms	1.1 million firms	Textile products, footwear, gambling and betting, postal services, travel agencies, education, healthcare
Vehicle Registers	ANFAC, DGT	Obligatory registration of vehicles	All registrations	Cars and motorcycle
Food Consumption Panel	MAPA	Survey to consumers	10500-12500 individuals	Food at home, Restaurants
Sales of tobacco products to retailers	CMT	Obligatory declarations by firms	All distributor firms	Tobacco
Consumption of petroleum products and natural gas	CORES	Obligatory declarations by firms	All distributor firms	Petroleum Products, Natural Gas
Households' bill tracking and survey statistics	CNMC	Survey to consumers	Around 9000 individuals	Electricity
Water Consumption	Canal de Isabel II	Sales by Canal de Isabel II	Sales by Canal de Isabel II	Water
Pharmaceutical Consumption through Medical Prescriptions	Ministry of Health	Registered medical prescriptions	All prescriptions registered including contribution of the buyer	Pharmacy products
Gambling sales	SALAE, ONCE, Ministry of Consumption	Obligatory declarations to Ministry of Consumption	All gambling firms, including online	Gambling and lottery
Residents Travel Survey	INE	Interviews to customers	8000 individuals each month	Tourist Trips
Economically Active Population Survey	INE	Survey to individuals	160000 individuals	Number of employed in Education and Healthcare
Synthetic data produced by INE	INE	Not public	Not public	Availability data of cleaning products, paper, textile products and footwear, toys, household equipment, computers, furniture, electronic products, jewelry. Synthetic data on Gross Value Added in transportation, catering and other services

Table C.2: Data Sources for National Accounts Consumption

This table summarizes the underlying data that the Spanish national statistics institute reports as using for inputs into the construction of national accounts consumption. We organize and translate the information available at https://www.ine.es/daco/daco4214/inventario_fuentes_metodos.pdf. The data sources reported in the last row are mentioned as inputs into national accounts, but we have been unable to find documentation on how they are constructed.

		Only Card and Cash Withdrawal with Card					
		Dropping Cash Withdrawal					
		Dropping Online Card Transactions					
		Dropping Imputed Rent					
		No Demographic Weights Applied					
		Weight of non-actives = 0.7					
		Weight of non-actives = 0.6					
		Weight of non-actives = 0.4					
		Weight of non-actives = 0.3					
		Baseline					
		National Accounts					
Levels		Correlation with Nat. Accts.	0.731	0.754	0.742	0.721	0.712
		Rooted MSE vs. Nat. Accts.	7185	16039	10130	8096	10990
		Mean per Adult Quarterly Cons. (€)	4288	4321	4662	4477	4188
		St. Dev. of per Adult Quarterly Cons.	263	228	237	232	224
QoQ		Correlation with Nat. Accts.	0.982	0.982	0.982	0.982	0.980
		Rooted MSE vs. Nat. Accts.	0.026	0.028	0.029	0.029	0.028
		Mean	0.007	0.009	0.009	0.010	0.010
		Standard Deviation	0.069	0.046	0.046	0.046	0.046

Table C.3: Impact of Modeling Choices on Relationship between Official and Naturally Occurring Consumption Series

This table reports the mean and standard deviation of different aggregate measures of quarterly consumption, along with their relationship to INE's National Accounts series for Final Household Consumption. The 'Baseline' series is our construction laid out in section 2. The next four columns then replace the 0.5 weight in the household weighting formula (2), with, respectively, 0.3, 0.4, 0.6, and 0.7. 'No Demographic Weights Applied' sums $c_{t,g,a,q}$ over all cells by quarter. We do not report levels since by construction this series will lie well below national consumption. The following three columns drop from the 'Baseline' series various components of consumption: imputed rent, online card spending, and cash. The final column reports consumption measured by card spending (including cash withdrawals) alone.

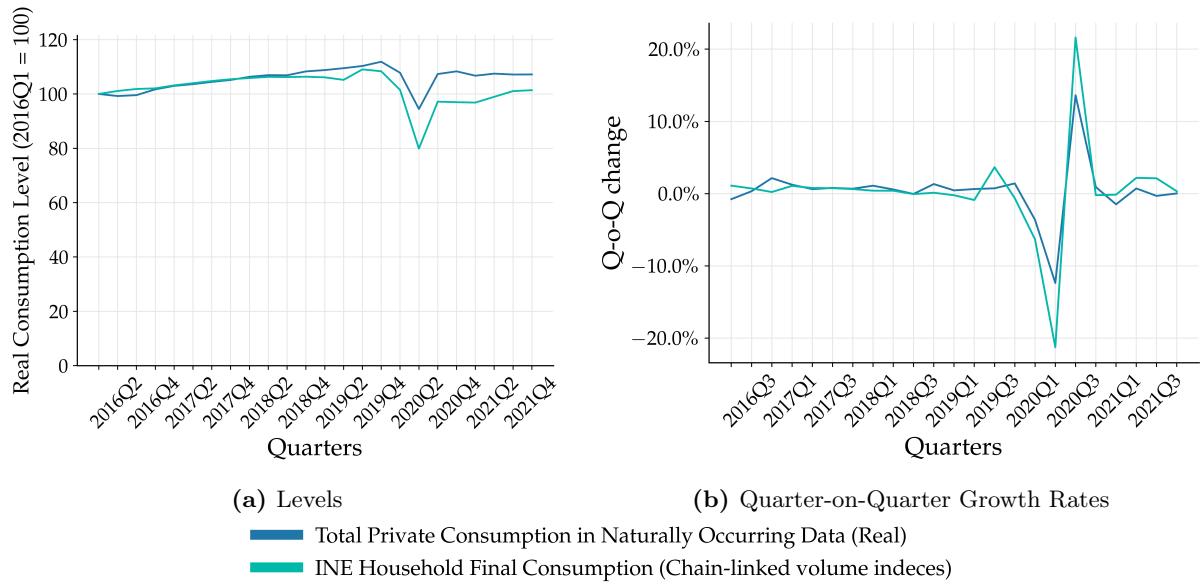
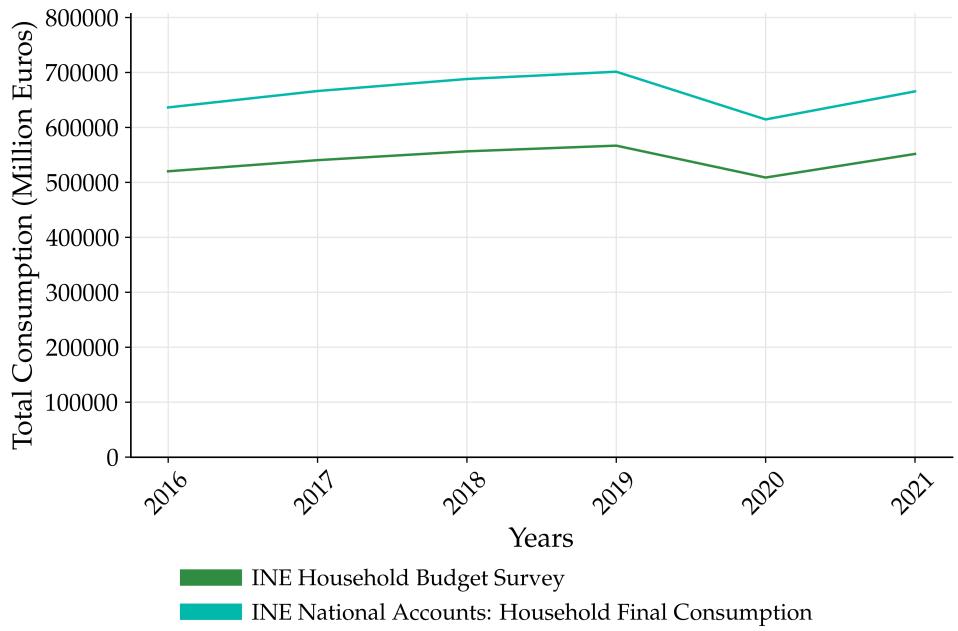
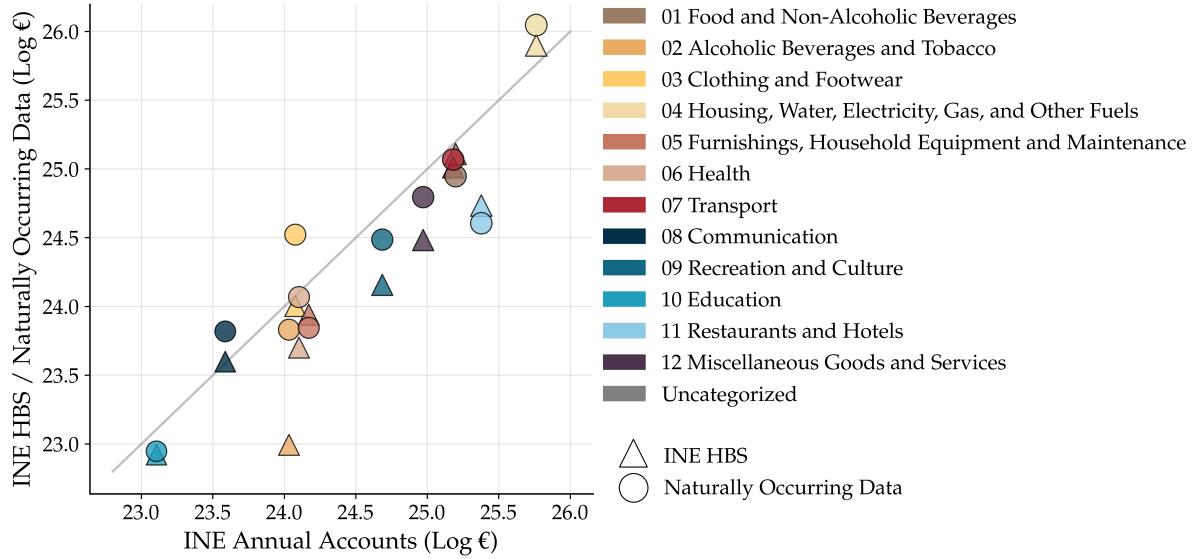


Figure C.4: Aggregate Naturally Occurring Consumption vs. National Accounts (Real)

Compared to the nominal consumption series plotted in Figure 4, 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) Total Consumption Levels



(b) Category-Specific Consumption Levels

Figure C.5: Total and Category-Specific Consumption Levels: Household Budget Survey vs. National Accounts

The top panel plots aggregate consumption at annual frequency as measured by the Household Budget Survey and National Accounts. The bottom panel is similar to figure 5 in the main text, except we add COICOP-specific consumption levels as measured by the Household Budget Survey in calendar year 2019.

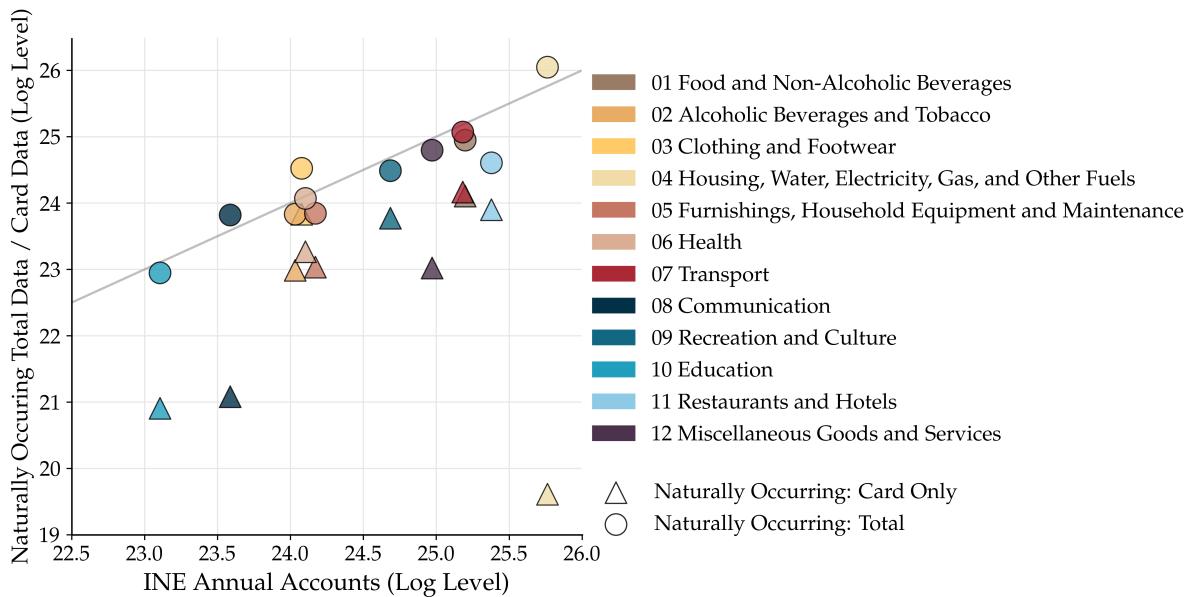


Figure C.6: Distribution of Spending across COICOP Categories: Naturally Occurring Card Data vs. National Accounts

This figure is similar to figure 5 in the main text, except we add COICOP-specific consumption levels as measured by the part of naturally occurring consumption that is purchased through cards, including cash withdrawals made with cards.

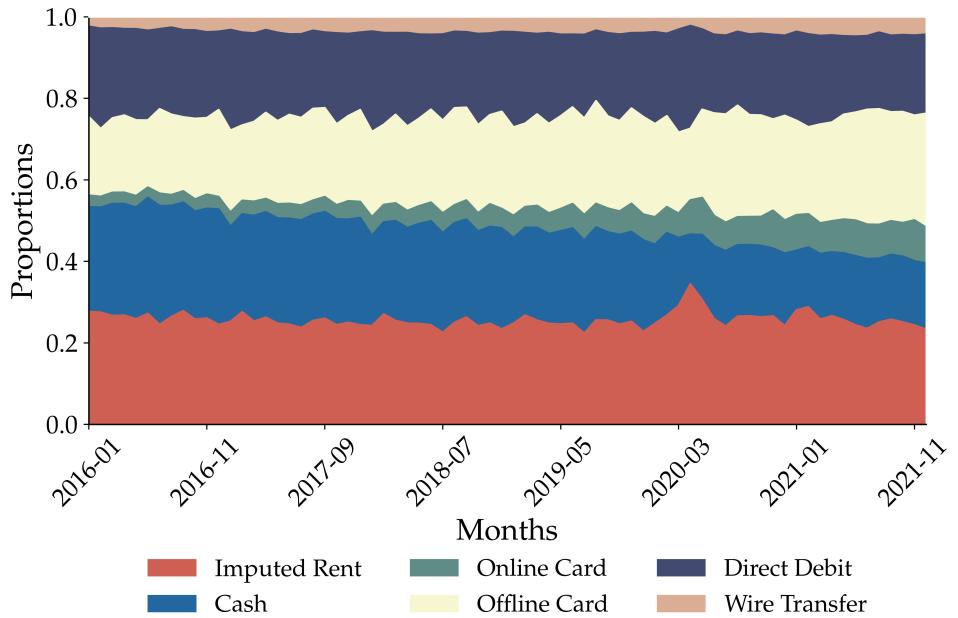


Figure C.7: Shares of different payment methods

This figure depicts shares of modes of payments at a monthly frequency from January 2016 till December 2021. For each BBVA customer we split monthly consumption expenditures into Offline Card, Online Card, Cash, Direct Debits and Wire Transfers and then an additional category to reflect the individual's imputed housing services. We then aggregate at a monthly frequency using demographic weights as discussed in Section 2, such that total expenditures across modes of payment add up to our baseline series of aggregate consumption.

C.3 Distributional National Accounts for Consumption: Additional Material

Figure C.8 depicts Lorenz Curves and Gini Coefficients implied by the distribution of 2017 consumption for selected COICOP categories across time frequencies. Panel (a) on the top shows the results for measured inequality in Food and Non-Alcoholic Beverages (COICOP category 1) as a function of time aggregation. Panel (b) on the bottom does the same for Furniture and Household Equipment (COICOP 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 COICOP 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).

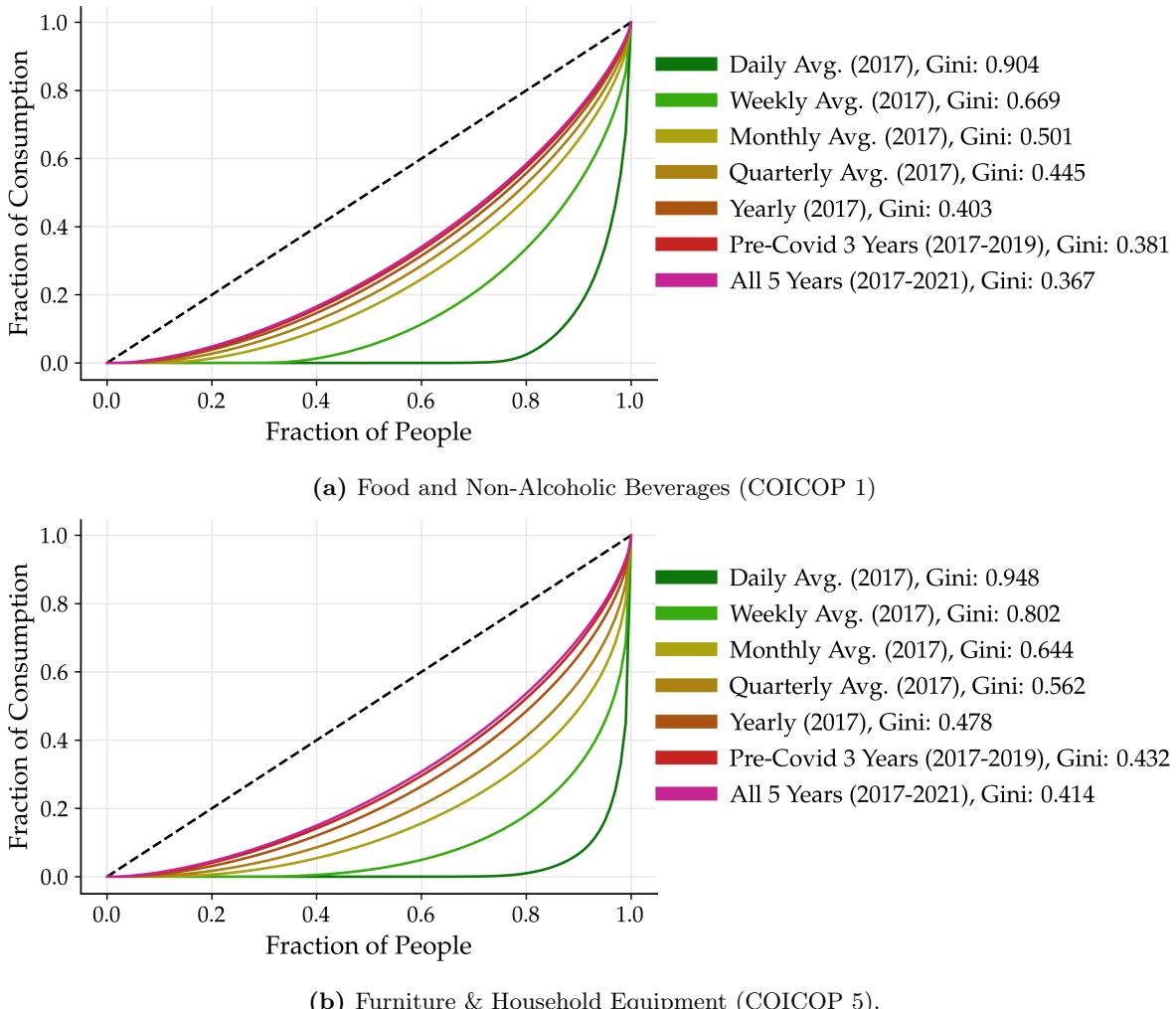


Figure C.8: Lorenz Curves and Gini Coefficients implied by the distribution of consumption of selected COICOP categories across time frequencies. Panel (a): Food and Non-Alcoholic Beverages. Panel (b) Furniture and Household Equipment.

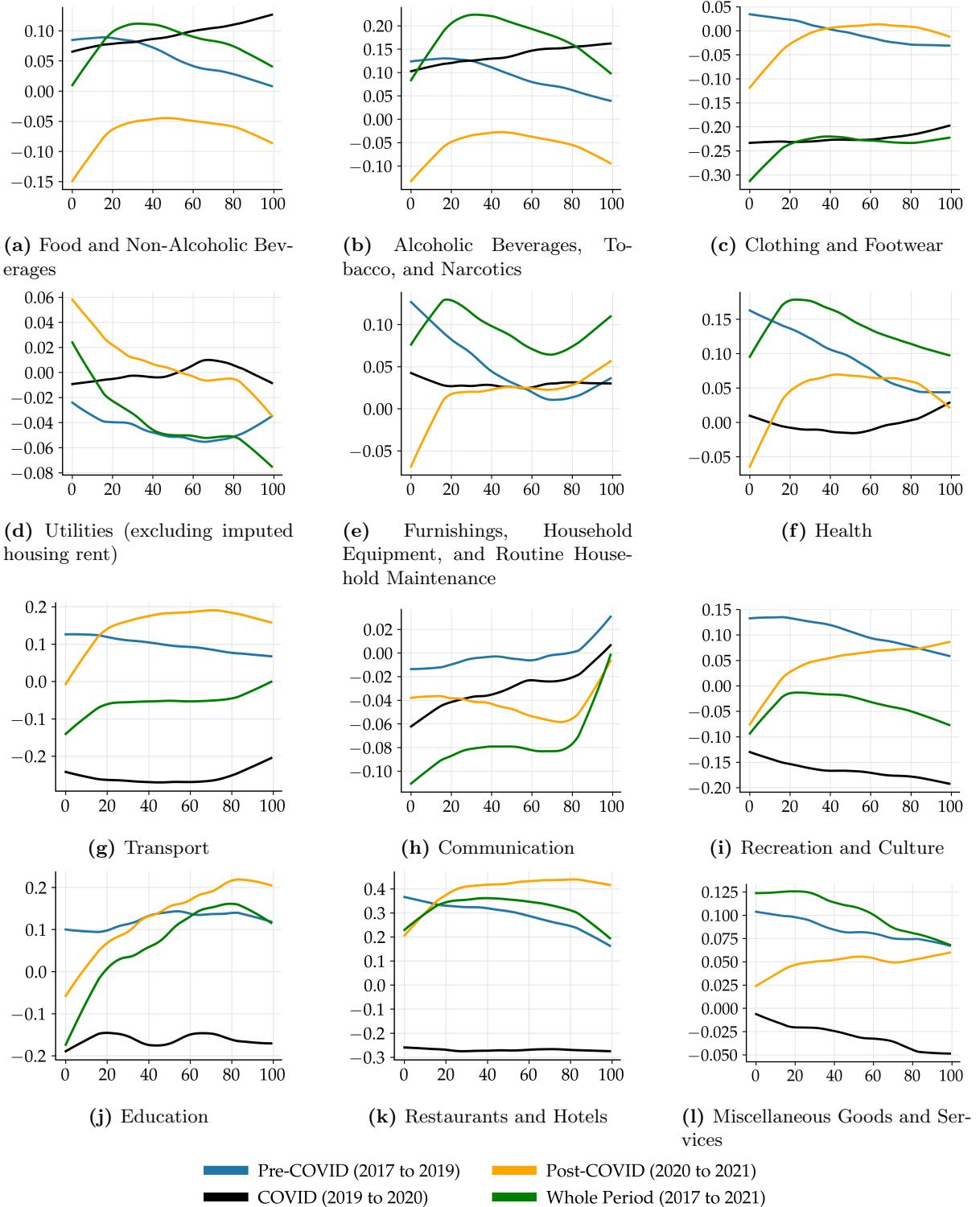


Figure C.9: Growth Rate of Consumption per Percentile of the Consumption Distribution and COICOP Consumption Categories (Yearly Cross-Sections)

This figure gives the average yearly growth rate of consumption, on the Y-axis, corresponding to each COICOP category and percentile of the total (across all COICOP categories) consumption distribution, on the X-axis. Each line corresponds to a different time horizon: green line: 2017-2021 period; blue line: 2017-2019; black line 2019-2020; yellow line: 2020-2021. The detailed construction for each category follows the same procedure as described in the note for Figure 13.

C.4 Individual Consumption Dynamics: Additional Material

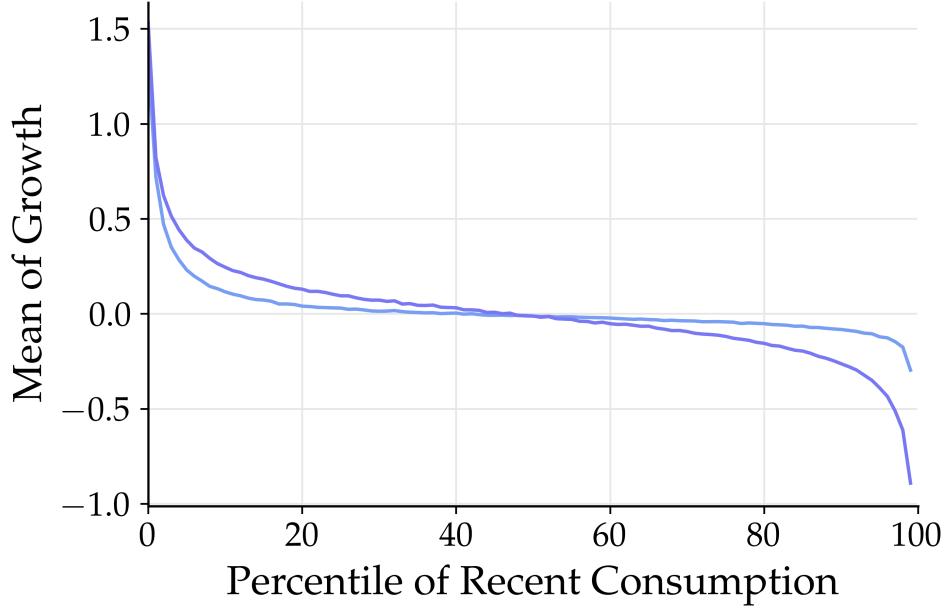


Figure C.10: Moments of Consumption Growth Distribution for COICOP 1 (Food and Non-Alcoholic Beverages) and COICOP 5 (Furniture and Household Equipment).

For each year $t = 17, 18, 19, 20$, we compute total consumption per active client over years $t - 1$ and t , as well as the log consumption change from year t to $t + 1$, for COICOP 1 and COICOP 5. We drop any client-year-category for which consumption is zero in years t or $t + 1$. We also redistribute cash to COICOPs 1 and 5 in proportion to their share of offline card spending. Finally, we average the moments over the four years and plot the results in this figure. Note that this construction is based on the population of active clients, not the bootstrap sample.