BBVA Research

National Accounts in a World of Naturally Occurring Data

A Proof of Concept for Consumption

G. Buda(BSE), V. Carvalho(Cambridge), S. Hansen(UCL)
A. Ortiz(BBVA Research),T. Rodrigo(BBVA Research) & J.Rodriguez(CUNEF)

Conference on Real-Time Data Analysis, Methods & Applications (Banco de España)
Madrid, October 19th, 2023.

Table of contents

Introduction

Relevance of Consumption and Problems with Surveys

- Consumption is Relevant:
  - Largest GDP component
  - Subject to extensive study (Heterogeneity at Heart of HANK Models)
- Surveys suffer well-known problems:
  - Don't aggregate to national accounts consumption
  - Under-reporting not constant across income
  - Limited panel coverage
  - Low frequency
  - Declining response rates

Survey Responses Abraham (2023)

Survey Data	CPS[a]	CPI Housing[b]	CE Interview[c]	MEPS HC[d]	ACS-Annual
Jan 2012	90.4	66.2	71.3	61.3 (overall)	97.3 (weighted)
Jan 2014	89.5	70.8	67.0	52.8	96.7
Jan 2016	86.7	68.2	63.7	51.0	94.7
Jan 2018	84.6	65.4	58.6	46.8	92.0
Jan 2019	83.1	63.3	57.6	46.0	86.0
Jan 2020	81.7	63.9	53.2	NA	71.2
Jan 2021	78.2	52.4	43.7	NA	NA
Jan 2022	73.3	52.6	NA		

SOURCE: Response rates were found on the websites of the U.S. Bureau of Labor Statistics (for CPS, CPI Housing, and CE Interview columns, see https://www.bls.gov/osmr/response-rates/household-survey-response-rates.htm), the Agency for Healthcare Research and Quality (for MEPS HC column, see https://meps.ahrq.gov/mepsweb/survey_comp/hc_response_rate.jsp), and U.S. Census Bureau (for ACS-Annual column, see https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/response-rates/).

[a]Current Population Survey, U.S. Census Bureau.

[b]Consumer Price Index Housing Survey, U.S. Bureau of Labor Statistics.

[c]Consumer Expenditure Survey, U.S. Bureau of Labor Statistics.

[d]Medical Expenditure Panel Survey, Household Component, Agency for Healthcare Research and Quality.

This Paper

- We apply **National Accounting principles (ESA-2010)** to Transactions from major private sector bank (BBVA)
- Results in a with **Consumption Panel** characterized by:
  - A Massive Survey: Almost two million participants
  - Aggregates to national accounts consumption at comparable frequencies
  - Updated Real time (daily frequency since 2015) & Geo-localized information
  - Breakdown of consumption into COICOP categories (officially yearly, now real time!!)
- With **Many Applications**
  - Consumption in Real Time & HD (1st Distributional Accounts of Consumption)
  - Rich characterization of consumption growth
  - Inflation Heterogeneous impact, High-frequency response to monetary policy shocks & Sustainability other papers

Construction Data: Building a Massive Consumption Survey

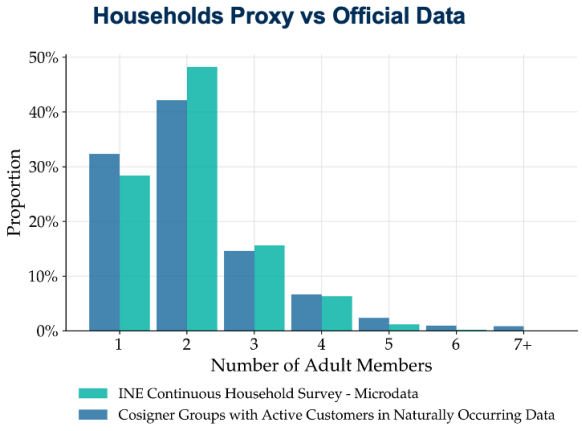
Transactions vs Household Budget Survey: Pros and Cons

- **Advantages** of Transactions vs Official Survey (HBS):
  - High Frequency
  - Observed rather than reported spending
  - Larger data both in cross section and time series
- **Disadvantages** of Transactions vs Official Survey (HBS):
  - Individual $\neq$ Households
  - Spending $\neq$ Consumption
  - Bank Clients $\neq$ Population

The Data Sample: Actual and Progress

- **Sample:** From 2015-Q1 to present
- 1.8 millions of **"Active Costumers"** (3bn individual transactions & 200 Bn euros).
  - In the sample latest 3 Years (less restrictive & Incorporating Youngers)
  - More than 10 transactions per quarter
  - Excluding clients with big Transfers among Banks (individual or Households).
  - Active Costumer drop to 1.8 million from more than 10 million clients (Likely Increase)
- Considering BBVA clients linked in a **"Household Units"** whenever:
  - They have co-signed a financial contract (bank account, loan, mortgage...)
  - They reside in the same postal code
- Households include both active and non-active clients

From Individuals to Households



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.

Consumption $\neq$ Spending

- Not all Individual-to-Firm Transactions are Consumption
- We first limit attention to Individual-to-firms/organizations with tax ID an try to assign transactions to 12 official COICOP Categories
- We use ESA-2010 principles to design appropriate filters:
  - Expenditure by Non-Residents $\neq$ Consumption (Export Services)
  - Real Estate purchases & House repairs $\neq$ Consumption (Investment)
  - Direct Tax payments $\neq$ Consumption (Government Revenue)

Structure if Payments Data: Non Housing Expenditures

Card Data

- Merchant Client Code (MCC) of the counterparty firm.
- Manual Mapping to COICOPS
- Multi-product retailers. Assigned by external data on distributions.

Transfers

- String match counter-party name to commercial registry.
- If counter-party is located as a firm, we assign as above.

Direct Debit

- $\sim$ 100 internal labels.
- Manual Mapping
- When this is unclear, we read field, determine firm and use either MCC (if possible) or NACE code of firm to assign COICOP.

Cash

- Both cash and over the counter.
- Assume is consumption.
- Assumptions on distribution.

Estimating Housing Expenditures

- A major but largely imputed component
- We build household rental payments using search of text description
  - Minimum 100 EUR & 70 months
  - 32127 Households
- We regression on:
  - Income (proxied by BBVA, 6 month average)
  - Utility Payments (Direct Debits)
  - Geography: 327 Regions (postal codes)
- Good fit in-sample & out-of-sample.
- Imputed to Whole

Estimating Rental Payments

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

Consumption Categories (COICOP): Transactions (Daily) vs Official (Yearly)

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)

Weighting: From BBVA to National

Household Weighting

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

A(i): "Active" customers in i's household
O(i): "Inactive" customers in i's household

(Under Revision)

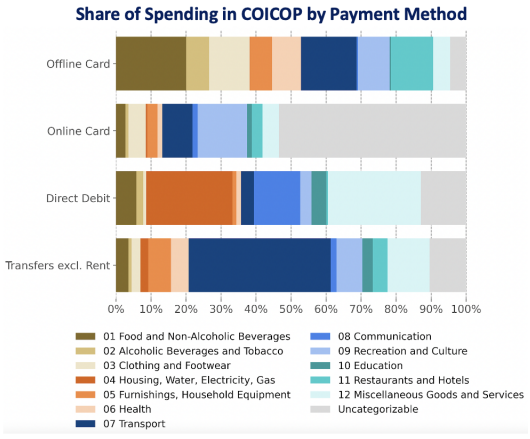
Demographic Weighting

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

(g, a, q): gender, age, within-region neighborhood income quintile cell
$\mathbf{x}^{\mathbf{INE}}$: count of Spaniards in cell in census
$\mathbf{x}^{\mathbf{BBVA}}$: count of Spaniards in cell in BBVA

Connecting Payments with Consumption

- Consumption by Means of Payments is not Homogeneous:
  - Cash: Distribution by assumptions
  - Off-Line Cards: Balanced
  - On-line Cards: Third point
  - Direct Debit: Utilities, Insurance
  - Transfers:Durables Spending (Cars,White Goods..)



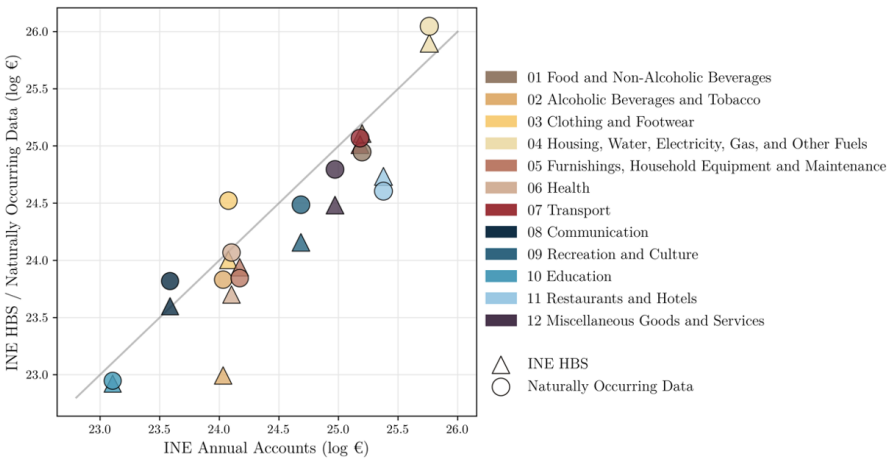
Aggregate Results

Aggregate Quarterly Results: Levels and Growth

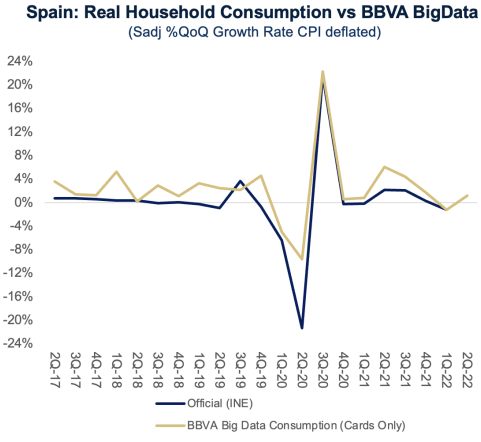
Aggregate Naturally Occurring Consumption vs. National Accounts

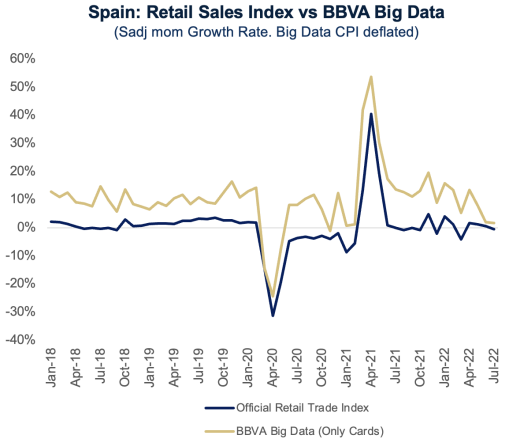
These figures compare quarterly aggregate household consumption according to official INE data and to naturally occurring data. To seasonally adjust both series, we use the Jdemetra+ application and apply X-13ARIMA-SEATS. The plot on the left shows the total level of consumption. The plot on the right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

Consumption Categories (COICOP): Naturally Occurring vs Official



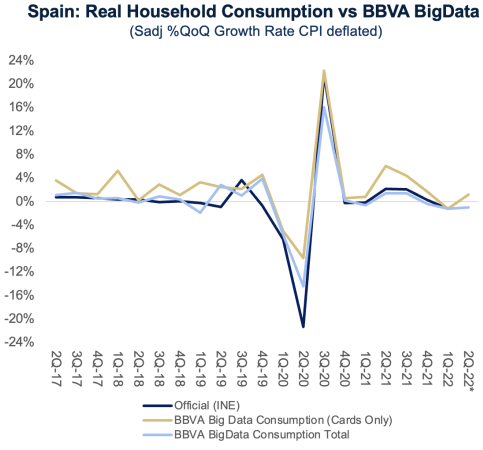
Total Consumption: Not Only Cards



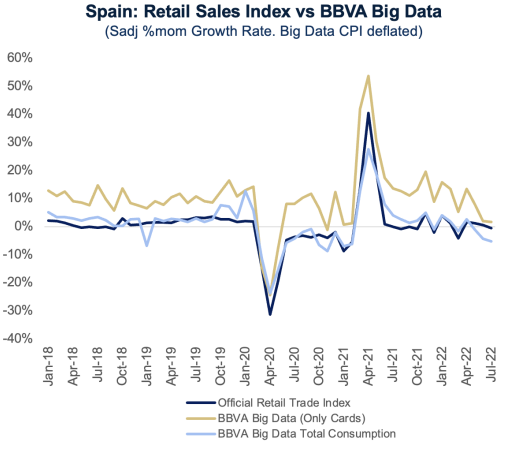


Source: Buda et Al (2022) BBVA Research & INE

Total Consumption: Not Only Cards



Source: Buda et Al (2022) BBVA Research & INE

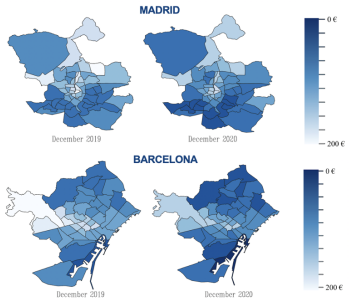


Total Consumption: The High Definition Component

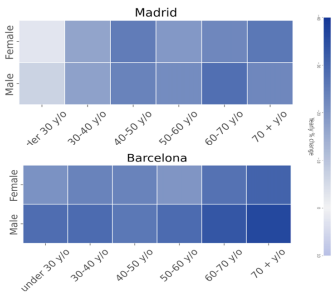
Urban Big Data Consumption: Barcelona & Madrid
(Moving Average 28D. YoY Nominal)

Source: Buda et Al (2022) BBVA Research

Consumption in Restaurants & Hotels in 2020
(Avg Consumption per person December 2020 & December 2019, %



Change in Consumption Restaurants & Hotels in 2020
(December 2020 vs December 2019, % YoY)



Aggregate Results: Take Home

National accounts are vital economic statistics but face pressure:

- Declining survey participation
- Budget cuts for national statistics agencies
- Pressure for more timely and granular data

In many countries they are sparse to non-existent (Silungwe et. al. 22):

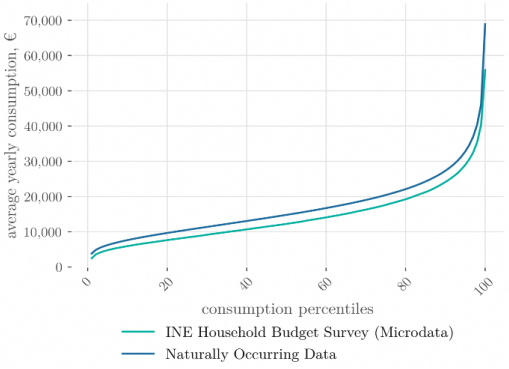
- 1/3 of countries do not publish quarterly accounts (1/2 in Africa)
- Only 4 European and 5 Asian countries produce quarterly acc. within 30 days
- 1/4 of countries have no Household Budget Survey (1/2 in Africa)

Proof of Concept: Widely available financial transaction data is competitive as a national accounts aggregate consumption measure.

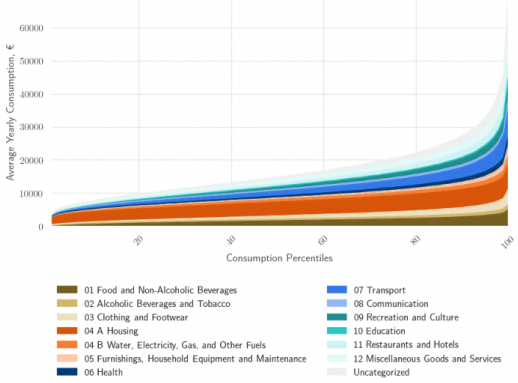
Use Cases

Consumption Inequality

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

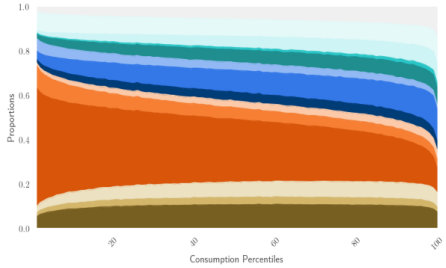


Consumption distribution disaggregated by COICOP
(consumption categories by percentiles adult distribution 2017)

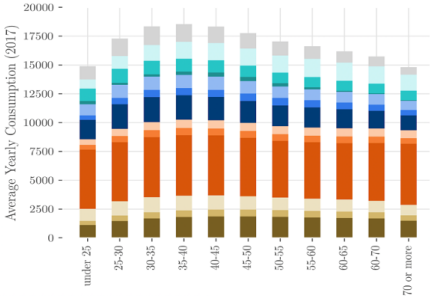


Consumption Distributional Accounts

Consumption distribution by COICOP Categories (Share)
(consumption categories by percentiles adult distribution 2017)

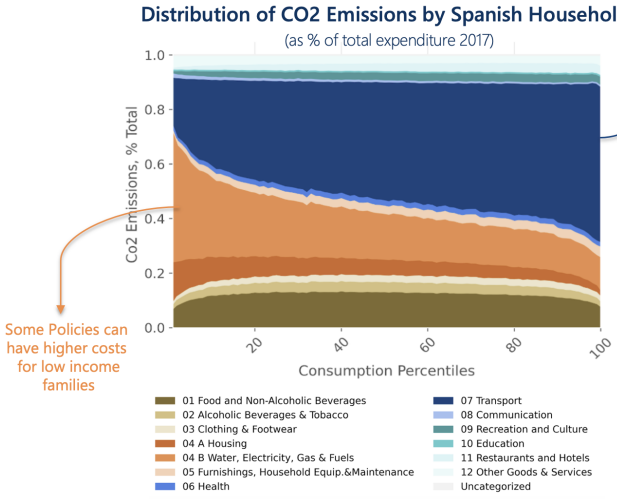


Consumption distribution by Category & Age
(consumption categories by age cohorts 2017)



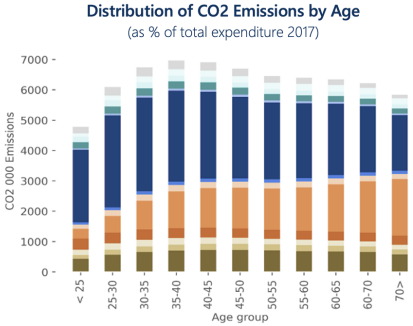
- 01 Food and Non-Alcoholic Beverages
- 02 Alcoholic Beverages and Tobacco
- 03 Clothing and Footwear
- 04 Housing, Water, Electricity, Gas
- 05 Furnishings, Household Equipment
- 06 Health
- 07 Transport
- 08 Communication
- 09 Recreation and Culture
- 10 Education
- 11 Restaurants and Hotels
- 12 Miscellaneous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

Sustainability: Inequality on Households Carbon Footprint



There is a Relevant Margin for Heterogeneous Policies

Some Policies can have higher costs for low income families

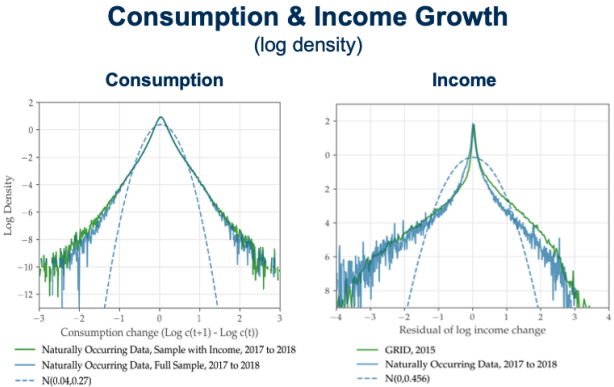


Consumption Growth Analysis

- Distribution of Consumption Growth
  does not look Gaussian not Homogeneous
  - Thick Tails
  - Albeit less than income
  - Very Non-Linear & Lumpy process
  - Very strong mean reversion

.

- Potential Causes
  - Income Process
  - Lumpiness of purchases

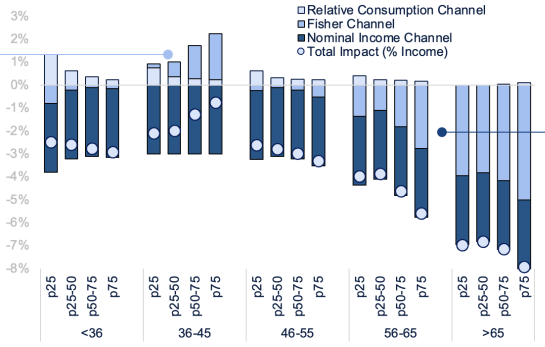


Heterogeneous Effects of Inflation Shock

Effect of "Inflation Shock" on Household by Income & Age

(Unexpected Inflation Shock by Age & Percentiles of Income. in % of Income)

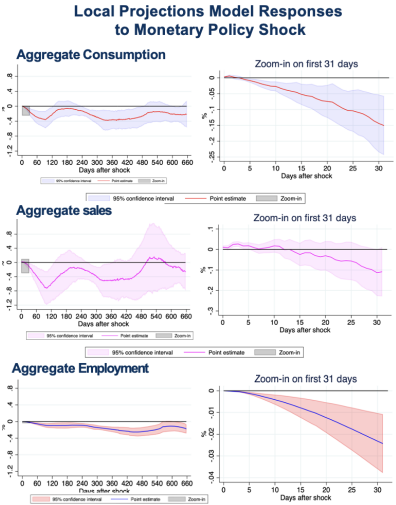
Indebted Mid-Age Balanced the impact (Fisher Effect)

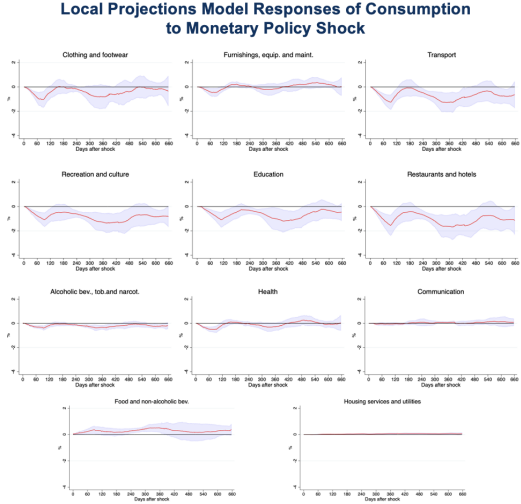


Older & High Income Cohorts relative higher impact (No Debt)

Source: Forthcoming Cardoso et al (2023)

Monetary Policy: Short & Variable Lags + Heterogeneous response

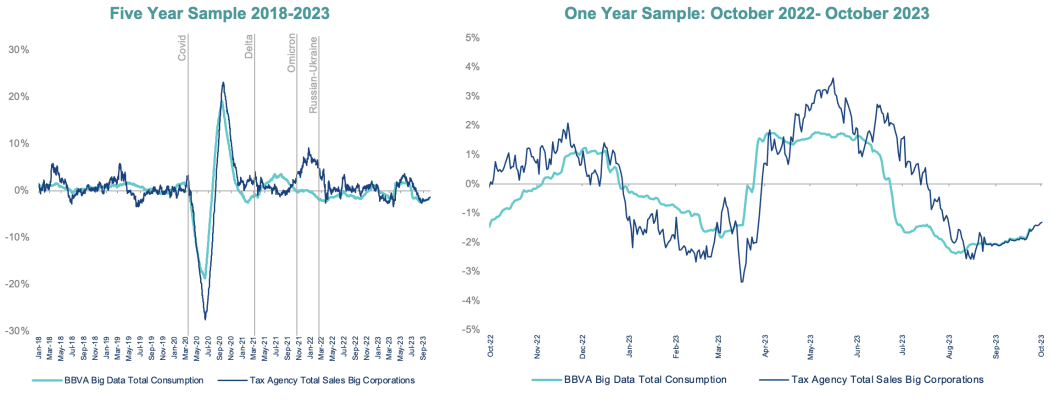




Nowcasting Consumption & GDP

Spain: Big Data Total Consumption BBVA & Tax Agency Total Real Sales

(SA . Adjusted, Cumulative 91 days % Quarterly Growth Rate. Big Data Consumption CPI deflated)



Source: Own source and AEAT

References

- Buda, Gergely, Stephen Hansen, T Rodrigo, VM Carvalho, Á Ortiz, and JVR Mora. (2022). **"National Accounts in a World of Naturally Occurring Data: A Proof of Concept for Consumption"**. CEPR Press Discussion Paper No. 17519.
- Buda, Gergely, Stephen Hansen, T Rodrigo, VM Carvalho, Á Ortiz, and JVR Mora. (2022). **"The Spanish Households' Carbon Footprint Inequality in High Definition & Real Time"**. BBVA Research Economic Watch
- Buda, Gergely and Carvalho, Vasco M. and Corsetti, Giancarlo and Duarte, João B. and Hansen, Stephen and Ortiz, Álvaro and Rodrigo, Tomasa and Rodríguez Mora, José V. (2023). **"Short and Variable Lags"**. CEPR Press Discussion Paper No. 18022.
- Cardoso, Miguel, Clodomiro Ferreira, José Miguel Leiva, Galo Nuño, Álvaro Ortiz, and Tomasa Rodrigo (2023). **"The Heterogeneous Impact of Inflation on Households Balance Sheets"**. Forthcoming in Bank of Spain Working Paper Series

Thanks...