

# Big Data & AI for Economics

Economic in Real Time & High Definition for Smarter Policies

---

Dr. Alvaro Ortiz  
Head of Big Data & AI Economic Analysis  
BBVA Research

Seminar at Bilkent University. April 2024

# Today's seminar

- The Spur of Big Data in Economics & Finance. Why Know and What for?
- Big Data and AI at BBVA Research: What we do and How we work
  - Real Time and High-Definition Economics
  - Big Data for Policy Analysis: Design of Smart Policies
  - AI for Economic Analysis
- Is Big Data & AI Relevant for Economists
- Questions and Answers



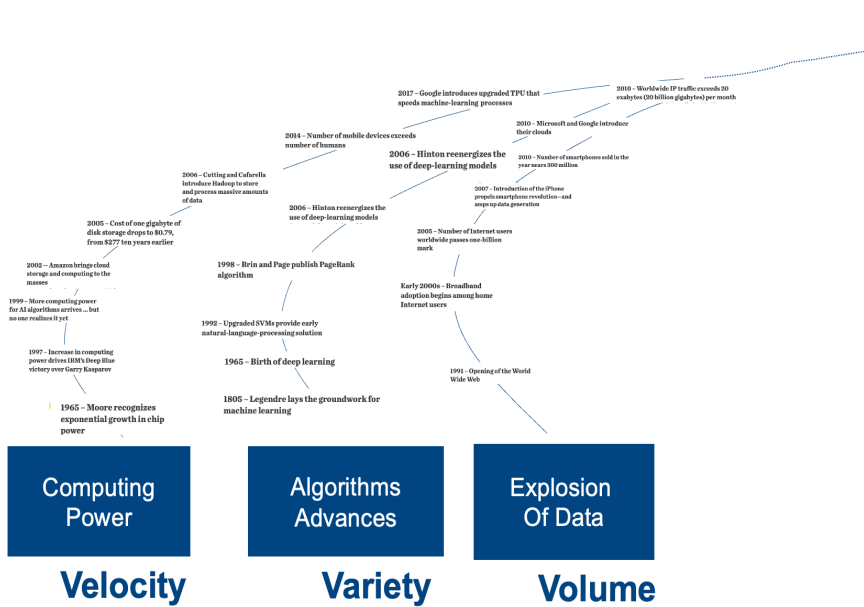
# Big Data in Economics, Finance & Social Sciences

## Why Know and What for?

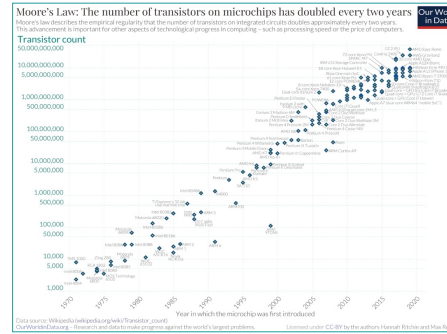
---

# Why Now? ... The alignment of several forces has triggered a powerful AI wave ... which is still well alive..

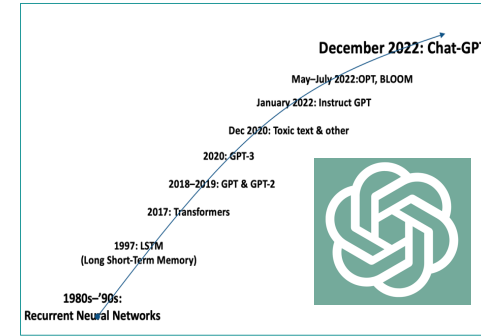
A Multi dimensional journey which is gaining momentum...



## Capacity & Speed or processing...

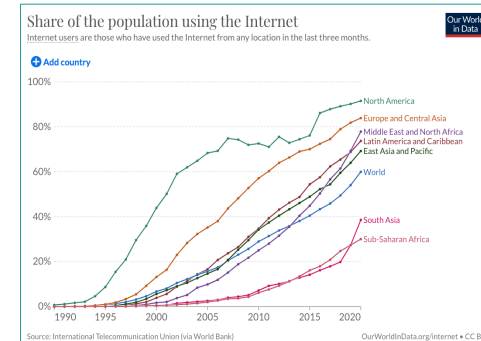
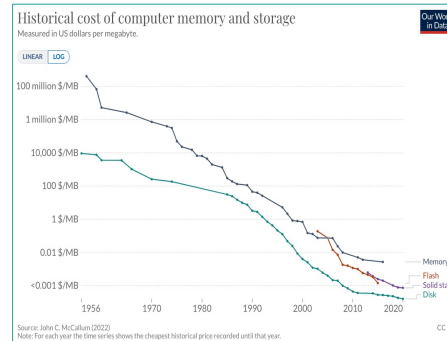


## ...Improving Algorithms



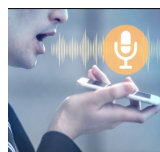
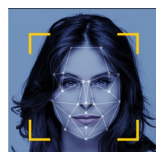
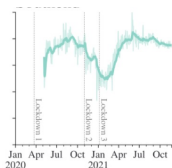
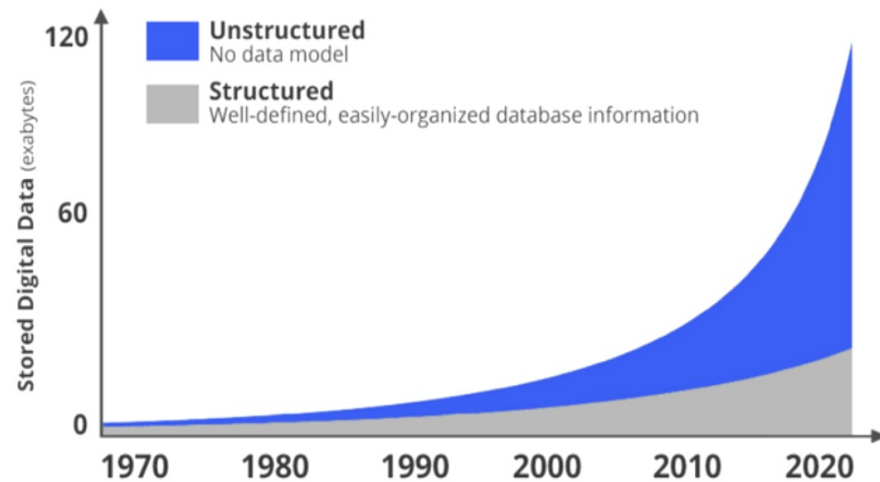
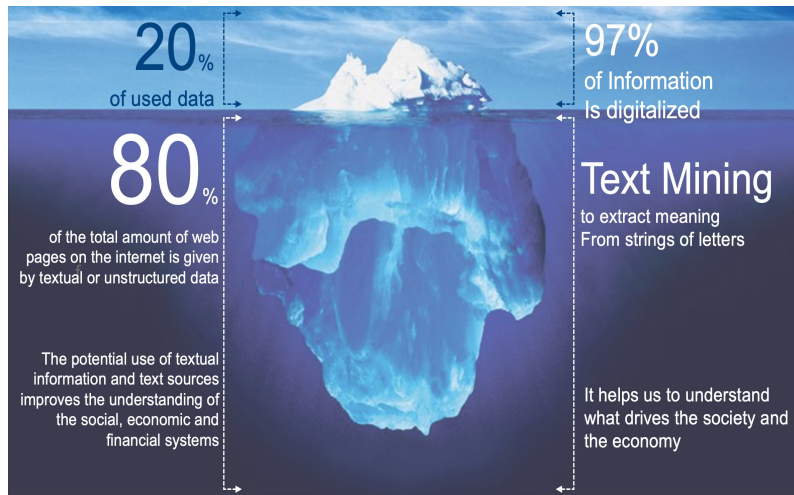
Source [ChatGPT is everywhere. Here's where it came from](#)  
...Open Source Intelligent

## Rapid decreasing Costs...



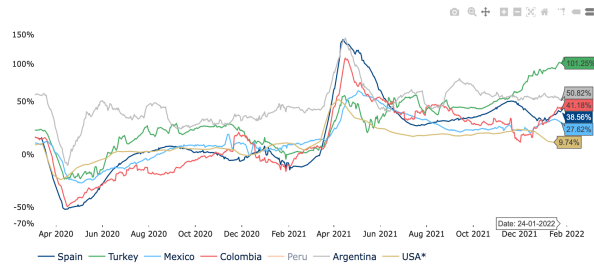
Source: Own elaboration based on [An executive's guide to AI \(Mckinsey\)](#)

# “Data from the People and for the People”: A vast amount of information growing exponentially in Non-Structured way from many devices



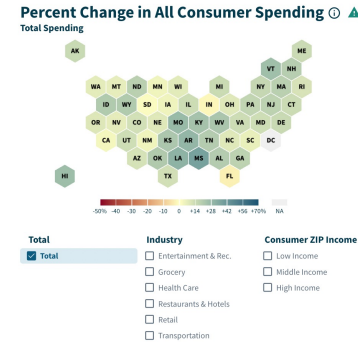
# What for? Increasing need of Real Time economics...

## Consumption Big Data: Card Data Transactions (% YoY)

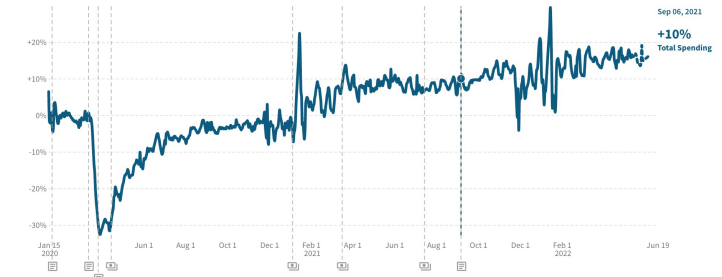


Spending at BBVA POS by BBVA and non-BBVA customers plus spending by BBVA customers at non-BBVA POS. This data considers final expenditures, so not intermediate consumption is included. Therefore, the correspondence with national accounts household consumption is not fully equivalent. Reference for the methodology can be found in the publication "Tracking the COVID-19 Crisis with High-Resolution Transaction Data" and the weekly publications "The COVID-19 Impact on Consumption in Real Time and High Definition". \*USA data provided by the Federal Bank of Chicago CARTS index.

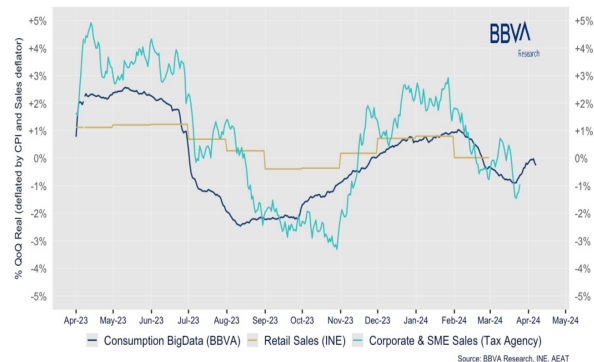
## US: Consumption Big Data: Card Data Transactions (% YoY)



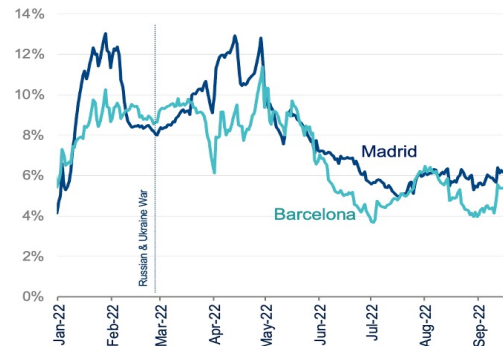
In the United States, as of September 06 2021, total spending by all consumers increased by **10%** compared to January 2020.



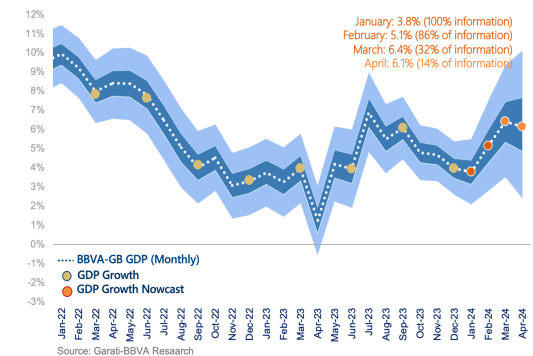
## Consumption Big Data , Retail sales & Corp. Sales (% QoQ, Sadj)



## Spain: Urban consumption in Real-Time (% YoY, Nominal 28d moving average)



## Turkey: Nowcasting GDP with Big Data (% YoY)

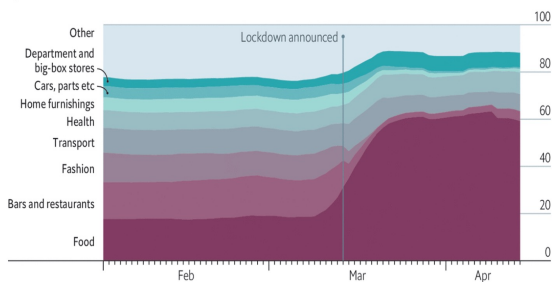


# What for?... Information in High Definition

## Spain: Consumption by Activity during Covid (% total Consumption)

### Locked down, loading up

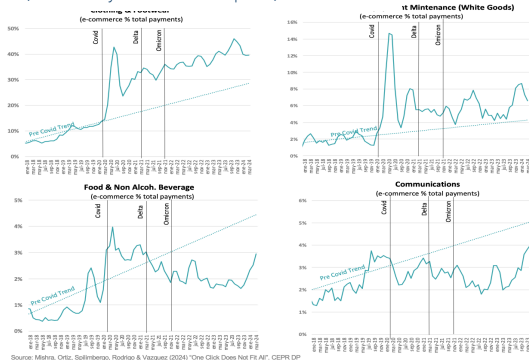
Spain, selected retail sales, % of total\*



Source: "Tracking the COVID-19 Crisis with High-Resolution Transaction Data," by Carvalho et al., Cambridge Working Papers in Economics (2020)

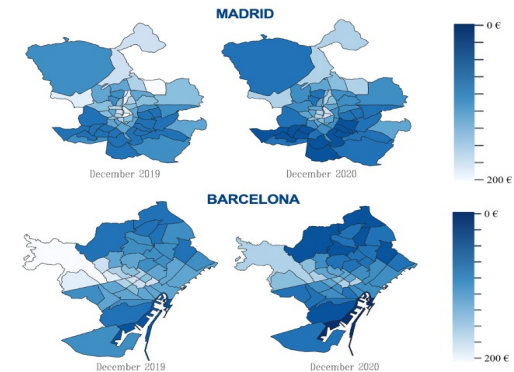
\*Seven-day moving average

## Spain: E-commerce by Category (% total Payments Consumption)



Source: Mahra, Ortiz, Spilberg, Rodrgo & Vazquez (2020) "One Click Does Not Fit All". CEPR DP

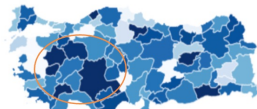
## Spain: Consumptions Restaurants & Hotels (avg Consumption per person during December 2019-2020)



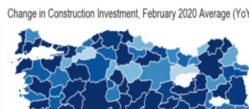
## Turkey: Investment During Covid

(avg Consumption per person during December 2019-2020)

Change in Machinery Investment, February 2020 Average (YoY)      Change in Machinery Investment, April-May 2020 Average (YoY)



Covid trigger GVA problems In Textile & Cars Export platform Provinces (Center Anatolia)...



Legacies with 2018- 2019 Financial Crisis with permanent effects in Construction

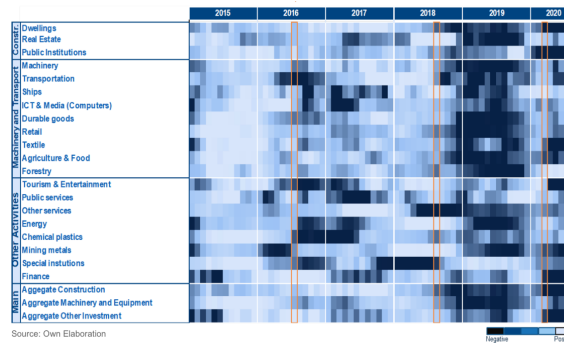
\*White stands for YoY rates >50% nominal & Dark Blue <-30% nominal declines

\*White stands for YoY rates >50% nominal & Dark Blue <-30% nominal declines

Source: Barlas, Guller, Ortiz & Rodrigo. Investment in Real Time and High Definition

## Turkey: Investment Heat Map

(avg Consumption per person during December 2019-2020)

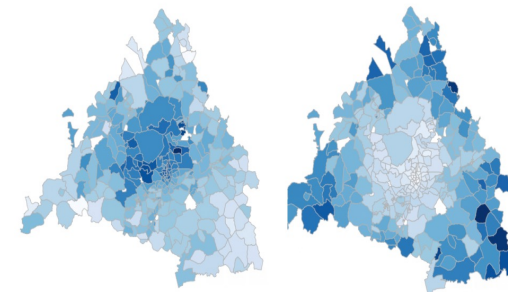


Source: Own Elaboration

## Madrid: Lifestyles

(Intensity of Lifestyles Go Bar & Restaurants & Car Trips)

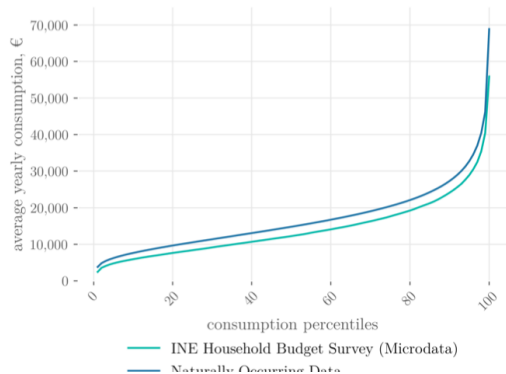
### Go Bars & Restaurants



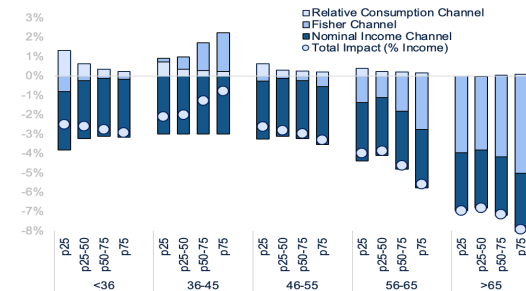
Source: BBVA Research & MIT Media Lab (forthcoming)

# What for? "Smart Policies" to deal with heterogeneity of People & firms

## Consumption Inequality: Big data vs Official (Average yearly consumption by percentile)

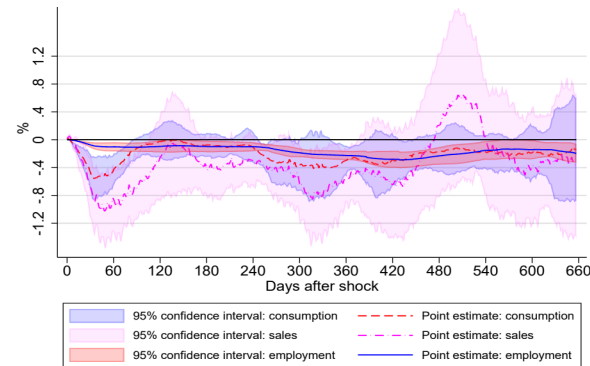


## Heterogeneous Impact of Inflation (through Fisher, Income and Consumption)

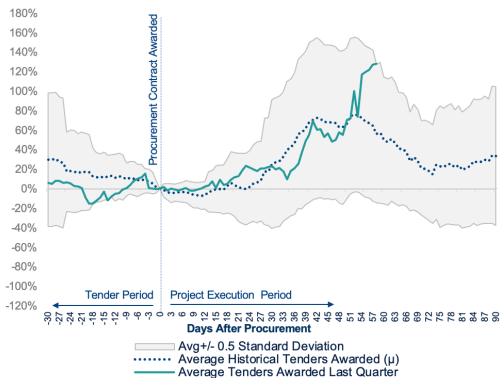


Source: Ferreira et al. (2023). The heterogeneous impact of Inflation in Households Balance Sheet. BIS Working Paper 1152

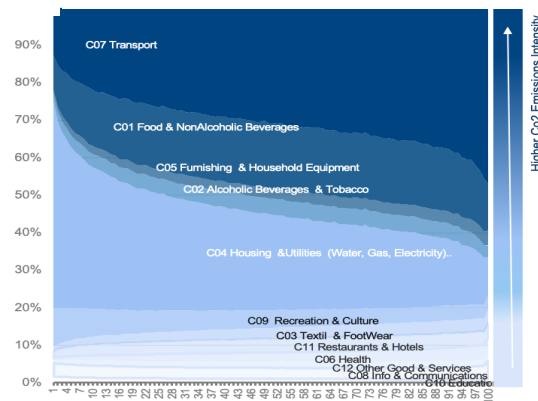
## Monetary Policy: Short and Variable Lags (Percentile & Categories as % of Total Consumption)



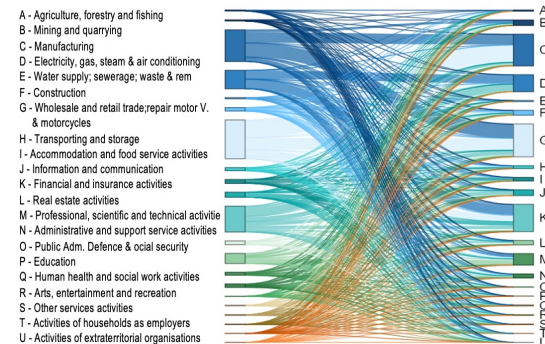
## Impact of Procurement Funds on New Loans (Loans Awarded vs Non-Awarded ratio evolution)



## Spain: Co2 Emissions in Consumption (By categories and percentiles of consumption)



## Financial Firm-to-Firm Transactions IO Table (BBVA Intra sector Firm-to-firm Transactions)



Source: BBVA Research

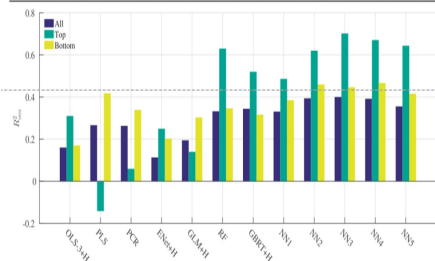


# And Last, But not Least, there is a huge potential for Finance and Banking Applications

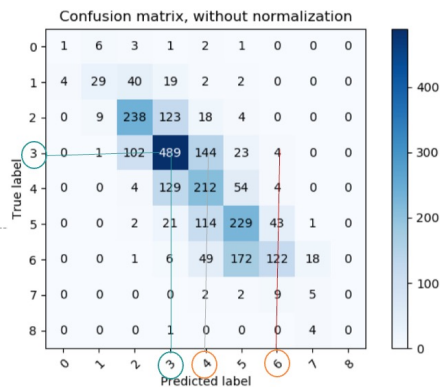
## Horse Racing On Returns Models

Table 1  
Monthly out-of-sample stock-level prediction performance (percentage  $R_{008}^2$ )

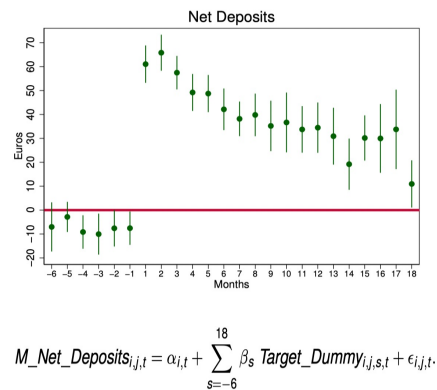
	OLS +H	OLS-3 +H	PLS +H	PCR +H	ENet +H	GLM +H	RF +H	GBRT +H	NN1	NN2	NN3	NN4	NN5
All	-3.46	0.16	0.27	0.26	0.11	0.19	0.33	0.34	0.33	0.39	0.40	0.39	0.36
Top 1,000	-11.28	0.31	-0.14	0.06	0.25	0.14	0.63	0.52	0.49	0.62	0.70	0.67	0.64
Bottom 1,000	-1.30	0.17	0.42	0.34	0.20	0.30	0.35	0.32	0.38	0.46	0.45	0.47	0.42



## Building Risk Models



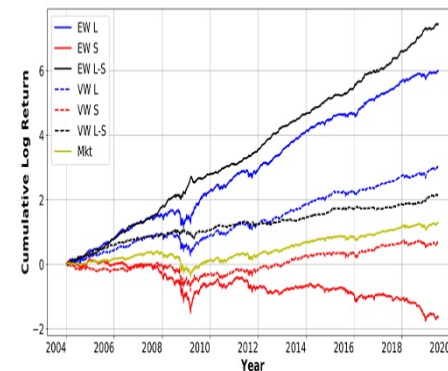
## Casual Inference (Robo Advising)



Gargano & Rossi (2021). Goal Setting and Saving in the FinTech Era

## Portfolio Strategy based on LLM

Figure 3: One-day-ahead Portfolio Performance based on OPT



Chen Kelly & Xiu (2023) Expected Returns and Large Language Models

Gu, Kelly & Xiu (2021) Empirical asset pricing via machine learning. NBER WP

Provenzano et al. (2019) An Artificial Intelligence approach to Shadow Rating

# Big Data is also contributing to consolidate a revolution in Economic Analysis to a more Empirical, Team-Building & Multi-Disciplinary science

## Six Decades of Top Economics Publishing: Who and How?

DANIEL S. HAMERMESH\*

Distribution of Number of Authors > 3  
(Groups >3-4 and > 4-5)



% of Articles on Empirical Own Data & Experimental  
(rest: Theoretical and Borrowed Data)



Six Decades of Top Economics Publishing: Who and How?



Enormously impressed by the webinar by Raj Chetty today

[bcf.princeton.edu/event-director...](http://bcf.princeton.edu/event-director...)

An amazing combination of how to put big data together, test hypotheses, and draw highly relevant policy conclusions. This is 21st century economic research. I feel old, but excited. Watch it.

Traducir Tweet

### % of Citations from Other Disciplines

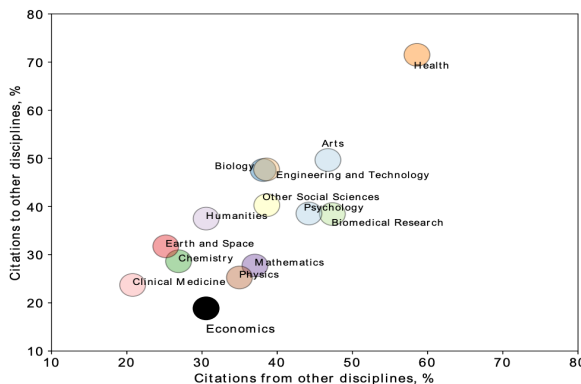


Figure 1. Citations in and out of disciplines by discipline. Source: Van Noorden (2015).

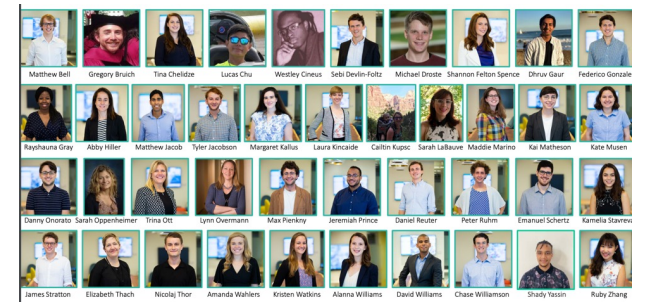
...

**How Did COVID-19 and Stabilization Policies Affect Spending and Employment?**  
A New Real-Time Economic Tracker Based on Private Sector Data

Raj Chetty, Harvard  
John N. Friedman, Brown  
Nathaniel Hendren, Harvard  
Michael Stepner, Univ. of Toronto and the Opportunity Insights Team

June 17, 2020

9 data providers and Near 45 authors !!!



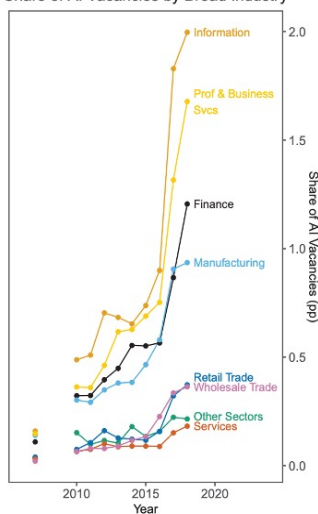
The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data

# ... and is already having an impact on "Human Resources" in terms of opportunities and challenges for Economists ...

## AI & JOBS: EVIDENCE FROM ONLINE VACANCIES

(Source: Acemoglu, D, Autor, D, Hazell, J & Restrepo NBER wp28257)

Share of AI Vacancies by Broad Industry



AI By Sector



Measure Felten et al Webb SML

## THE DEMAND FOR ECONOMISTS IS GROWING IN THE PRIVATE AND PUBLIC SECTOR..

**aws** **Economist, AWS Central Economics**  
Amazon Web Services (AWS)  
Seattle, WA  
5 connections work here  
3 days ago - **13 applicants**

**aws** **Senior Economist - Forecasting**  
Amazon Web Services (AWS)  
Seattle, WA  
5 connections work here  
6 days ago - **14 applicants**

**aws** **Senior Economist, AWS Central Economics, DSI Team**  
Amazon Web Services (AWS)  
Cologne, North Rhine-Westphalia  
5 connections work here  
1 week ago - **24 applicants**

**aws** **Economist**  
Amazon Web Services (AWS)  
Seattle, WA  
5 connections work here  
3 days ago - **8 applicants**

**Amazon researchers honored by two esteemed academies**  
Staff writer  
May 20, 2022  
Guido Imbens elected to the National Academy of Sciences; Alberto Abadie elected to American Academy of Arts and Sciences.

### Two Amazon-affiliated economists awarded Nobel Prize

Amazon Scholar David Card wins half the award, while academic research consultant Guido Imbens shares in the other half.

### Amazon gets an edge with its secret squad of PhD economists

By Lydia DePillis, CNN Business  
Updated 1039 GMT (1839 HKT) March 13, 2019

Our job vacancies

Search jobs | What we offer | Graduates and Students | How you can join us

---

**Business Analysis Specialists - Information Systems**  
Information Technology | Information Systems | 22-Jun-2022

---

**Compliance and Governance Officer - Legal Analysis**  
Legal Analysis & Advice | Secretariat | 23-Jun-2022

---

**Data Science Experts - Securities & Financial Markets Data (ESCB/IO)**  
Knowledge, Information, Data&Statistics | Statistics | 08-Jun-2022

---

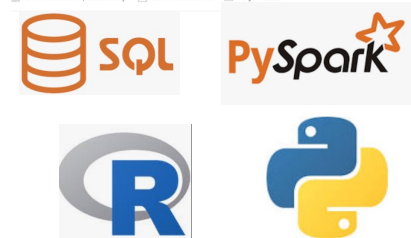
**Financial Risk Expert (ESCB/IO) - Risk Analysis**  
Knowledge, Information, Data&Statistics | Risk Management | 09-Jun-2022

---

**Financial Stability Experts (ESCB/IO) - Secretariat of the European Systemic Risk Board**  
ESRB Secretariat | 14-Jun-2022

---

**Head of the Design & Digital Division**  
Communication, Leadership | Communications | 15-Jun-2022



# ... It's a very interesting field for economics and maybe you are in the right place and the right moment... just ask to your AI friend !!



## You

Tell me some reasons why Big Data and AI are very interesting field of Study for the young Economists.



## ChatGPT

Big Data and AI are particularly intriguing fields of study for young economists for several compelling reasons:



•**Data-Driven Decision Making:** The integration of Big Data allows economists to analyse vast datasets for insights, leading to more informed and effective economic policy and business strategy decisions.

•**Improved Economic Forecasting:** AI techniques like machine learning can process complex and large datasets to forecast economic trends more accurately and quickly than traditional models.

•**Innovation in Economic Models:** Big Data enables the development of more nuanced economic models that reflect real-world complexities, enhancing the predictive power and relevancy of economic theories.

•**Efficiency in Market Analysis:** Automation and advanced data analytics can drastically reduce the time required to analyse market trends, competitive dynamics, and consumer behaviour, making economic research more efficient.

•**Understanding Complex Systems:** The ability to process and analyse large-scale data helps economists understand and model complex systems like global trade networks, financial markets, or urban economies.

•**Career Opportunities:** Proficiency in Big Data and AI opens up a range of career opportunities in both the public and private sectors, including roles in governmental agencies, international organizations, financial institutions, and tech companies.

•**Cross-disciplinary Research:** These tools enable cross-disciplinary research involving psychology, sociology, and computer science, providing a more holistic view of economic problems.

•**Social Impact:** Economists can use AI to address significant challenges such as poverty, inequality, and climate change, using data to craft solutions that are effective at a large scale.

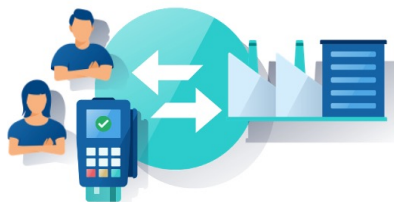
•**Technological Advancement:** Continuous advancements in AI and data analytics ensure that the field remains on the cutting edge of technology, continually offering new tools and methodologies for economic analysis.



# What do we Do at BBVA Research? and... How do work ?

---

# Big Data at BBVA Research: What do we do?



**National Accounts  
in “Real Time” &  
“High Definition”**



**Policy Analysis  
In real Time &  
High Definition**



**Sustainability for  
Household,  
Firm & Cities**

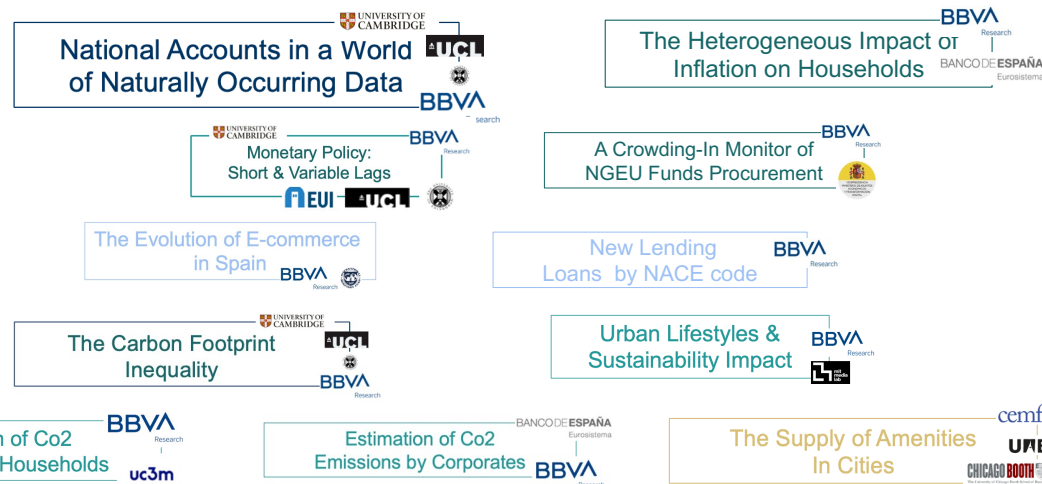


**AI  
NLP & LLM for  
Social Sciences**

# How do we work: The BBVA Research Model & Topics

Macroeconomics  
 Economic Policy  
 Banking & Finance  
 Sustainability

Financial Health  
 Urban Economics





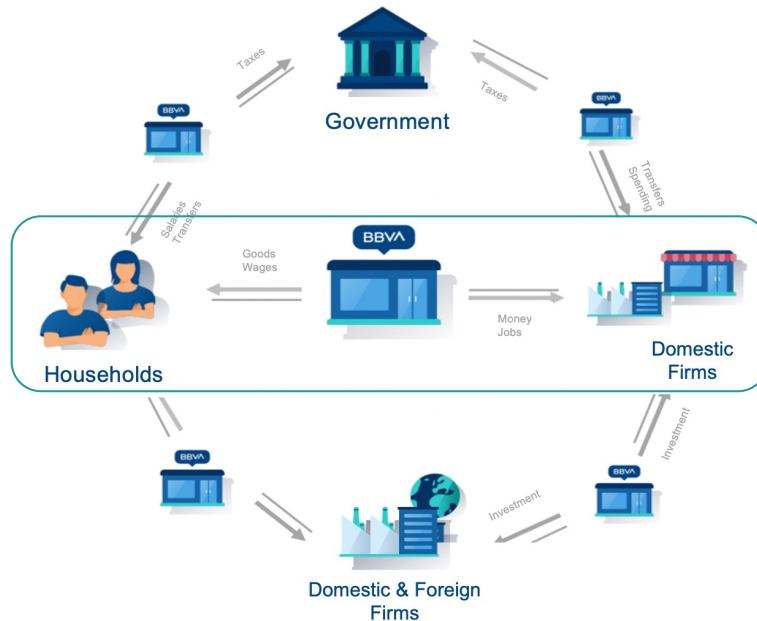
# Economics in Real Time and High Definition

---

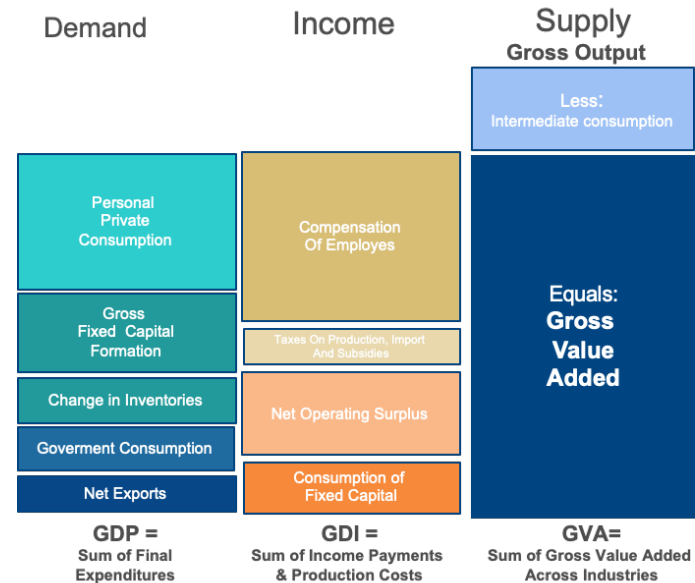


# Our Core Project: Replicating National Accounts through Non-Traditional Financial transactions... (A hard task by the way!!!)

## Financial Transactions



## The National Accounts System



# A social need (i.e. COVID) stimulated the development of new tools to know what was going on in real time...

National Accounts in a World  
of Naturally Occurring Data



## Financial Transactions & Consumption



# 38%

Total Consumption Payments



### Tracking the COVID-19 crisis through the lens of 1.4 billion transactions

Vasco Carvalho, Juan Ramón García, Stephen Hansen, Álvaro Ortiz, Tomasa Rodrigo, José V. Rodríguez Mora, Pep Ruiz 27 April 2020

*Individual transaction-level data can be used to map the changes in consumer behaviour in response to the COVID-19 lockdown measures. This column exploits data from BBVA to analyse the nature of the impact in Spain. The findings indicate that there has been a large decline in overall expenditure and that anticipatory spending (stockpiling) has taken place, whilst non-essentials have more or less collapsed. The authors find no evidence that differential exposure to the pandemic has affected regional expenditure dynamics, but within cities, areas with more prevalence of the pandemic have suffered a bigger economic collapse overall.*

<https://voxeu.org/article/tracking-covid-19-crisis-through-transactions>

### ROYAL SOCIETY OPEN SCIENCE

Research articles

#### Tracking the COVID-19 crisis with high-resolution transaction data

Vasco M. Carvalho, Juan R. García, Stephen Hansen, Álvaro Ortiz, Tomasa Rodrigo, José V. Rodríguez Mora and Pep Ruiz

Published: 11 August 2021 <https://doi.org/10.1098/rsos.210218>

**NBER** | NATIONAL BUREAU of  
ECONOMIC RESEARCH

Summer Institute 2020

## Economic Fluctuations and Growth, Summer 2020

<https://conference.nber.org/conference/papers/f143494.pdfspan>

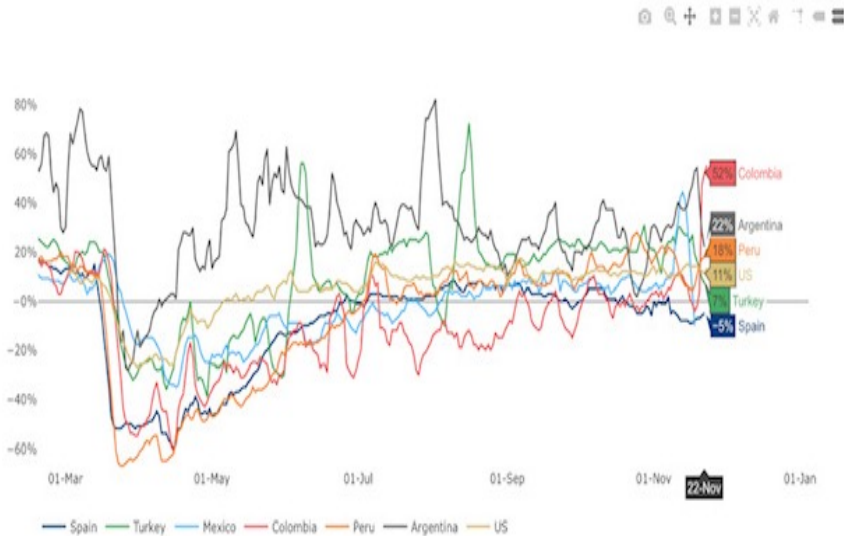
# We rapidly find interesting things on Covid Effects: Magnitude of Shock, the rapid recovery and the order of Sectoral Adjustment

National Accounts in a World  
of Naturally Occurring Data



## BIG DATA CONSUMPTION INDICES: TOTAL CARD CONSUMPTION BY COUNTRY

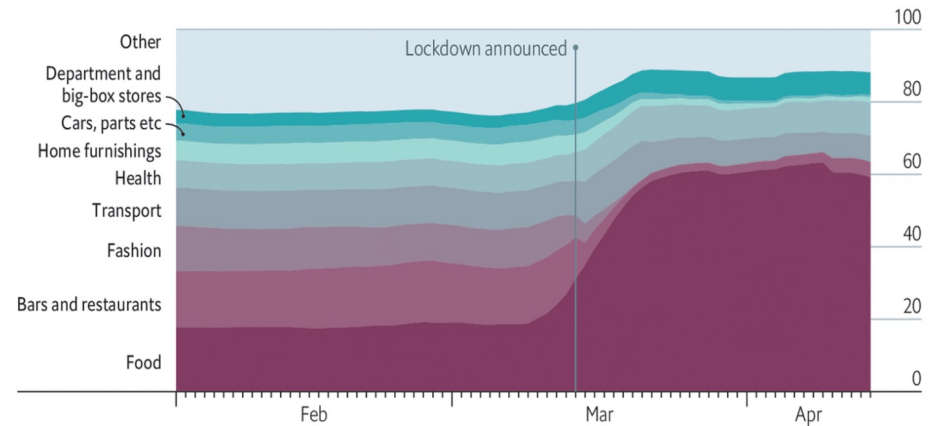
(% year over year, 7D cumulative)



US is proxied by US Sunbelt. Spending at BBVA POS by BBVA and non-BBVA customers plus spending by BBVA customers at non-BBVA POS. This data considers final expenditures, so not intermediate consumption is included. Therefore, the correspondence with national accounts household consumption is not fully equivalent. Reference for the methodology can be found in the publication "Tracking the COVID-19 Crisis with High-Resolution Transaction Data" and the weekly publications "The COVID-19 impact on Consumption in Real Time and High Definition".

## Locked down, loading up

Spain, selected retail sales, % of total\*



Source: "Tracking the COVID-19 Crisis with High-Resolution Transaction Data," by Carvalho et al., Cambridge Working Papers in Economics (2020)

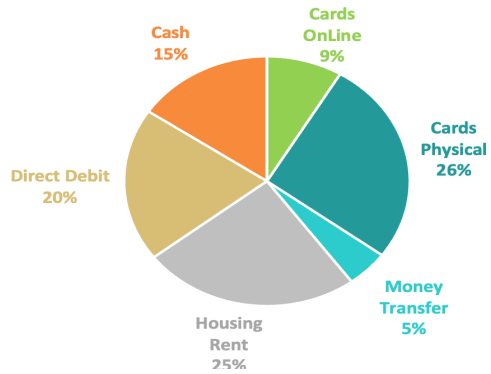
\*Seven-day moving average

# ... But we realized that "Consumption" its not only about Cards



→ **100%**  
Total Consumption Payments

**Spain: Distribution of Payments for Consumption**  
(Cumulative last 12 Months, % of Total)

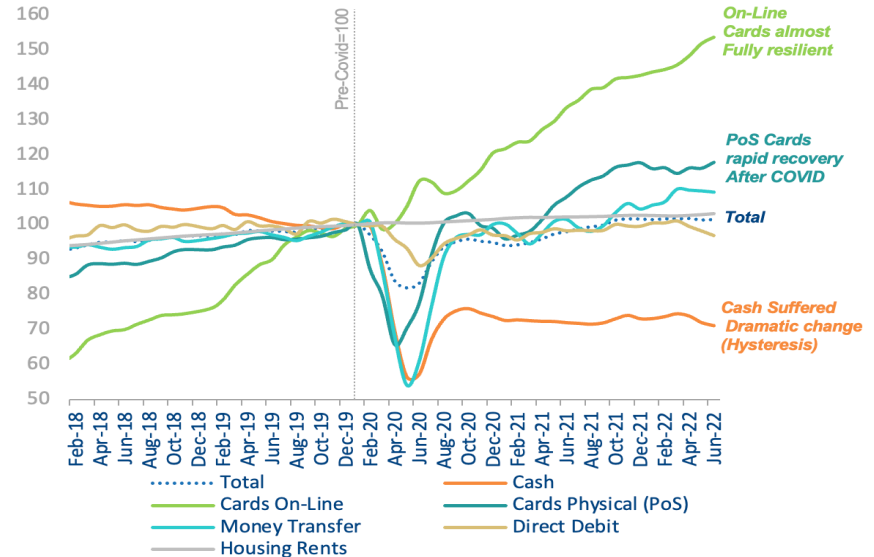


**Innovative Data in Household Finance:  
Opportunities and Challenges, Fall 2022**

National Accounts in a World  
of Naturally Occurring Data



**Spain: Evolution of Payments for Consumption**  
(SAAdj. Cum 3M January 2020=100)



Source: Buda et Al (2022) National Accounts in a World of Naturally Occurring Data: A Proof of Concept for Consumption. CEPR Discussion Paper

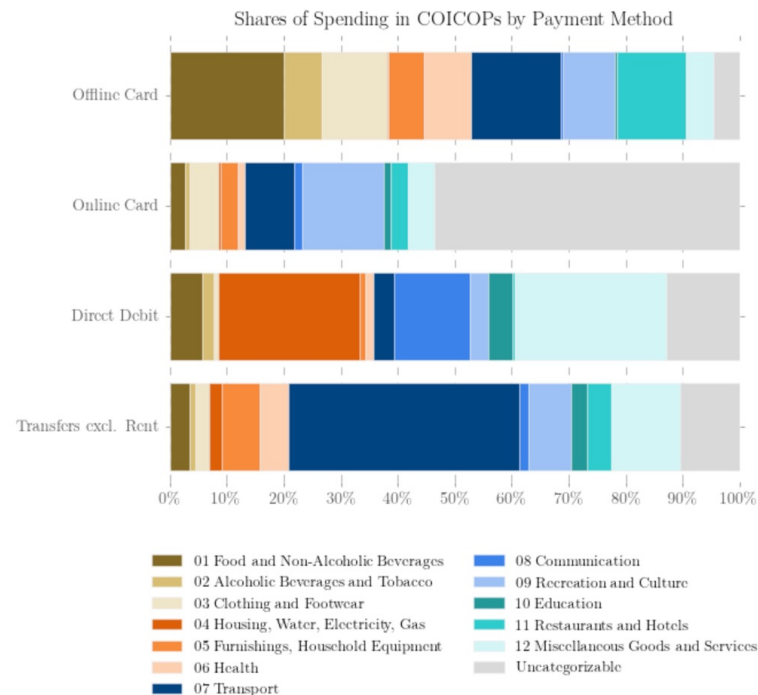
Source: BBVA Research & Buda et Al (2022).

# In fact, Consumption by means of payments is quite heterogeneous..

## 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..)

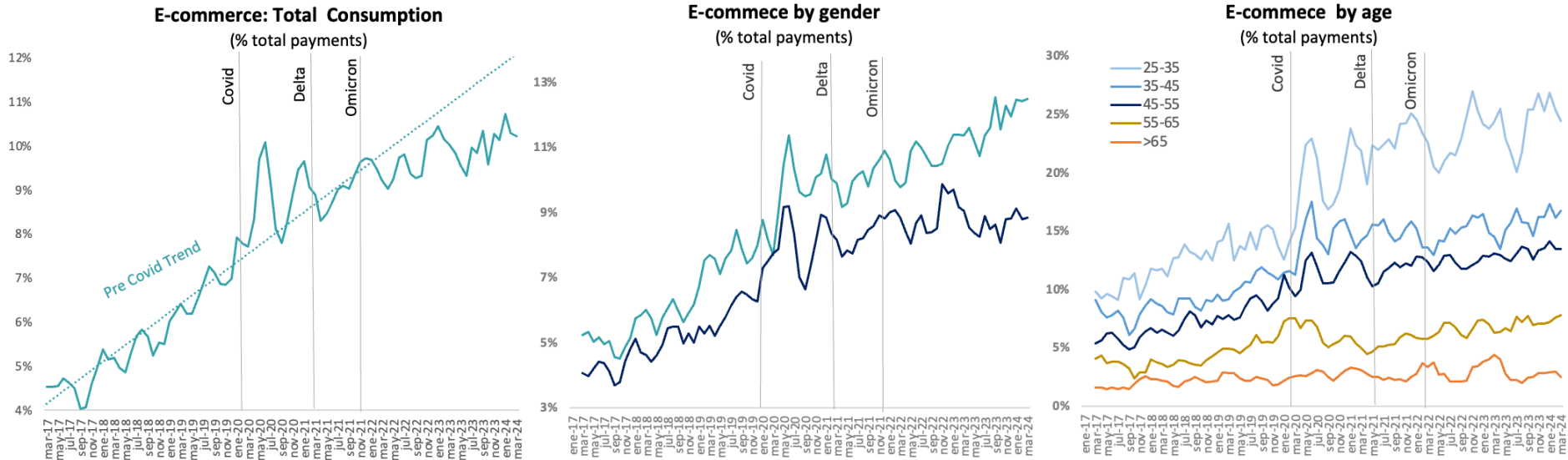
## Spain: Share of Spending in COICOPs by Method



# Some payments are developing faster than others... and heterogeneity is even present inside the payments: Female and Young Lead

## Spain: E-commerce Aggregate , Gender and Age (% of Total Payments)

The Evolution of E-commerce  
in Spain 



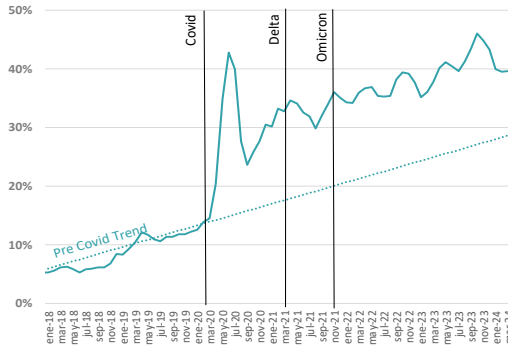
Source: Mishra, Ortiz, Spilimbergo, Rodrigo & Vazquez (2024) "One Click Does Not Fit All". CEPR DP

# ... indeed "One-Click" does not fit all..

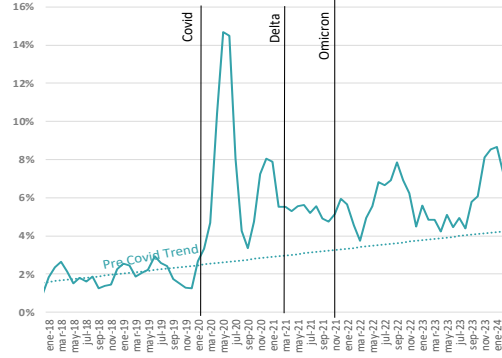
## Spain: E-commerce by Consumer Category (% of Total Payments)

The Evolution of E-commerce  
in Spain 

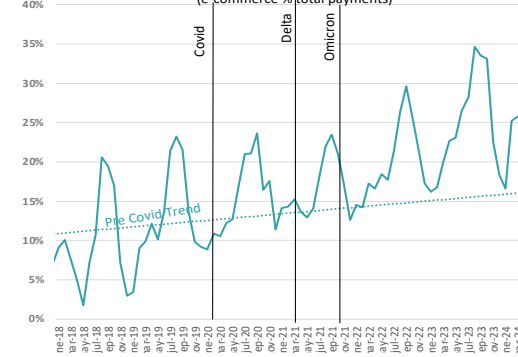
**Clothing & Footwear**  
(e-commerce % total payments)



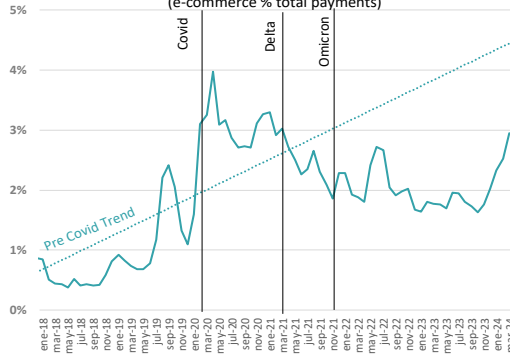
**Furniture & H. Equipment Maintenance (White Goods)**  
(e-commerce % total payments)



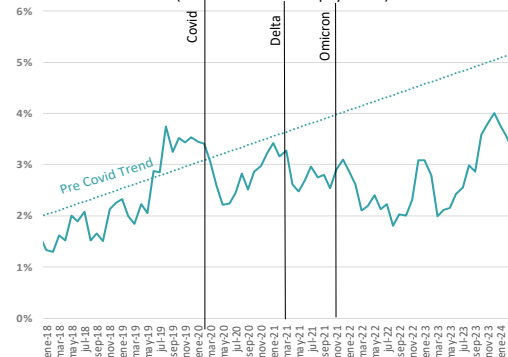
**Education**  
(e-commerce % total payments)



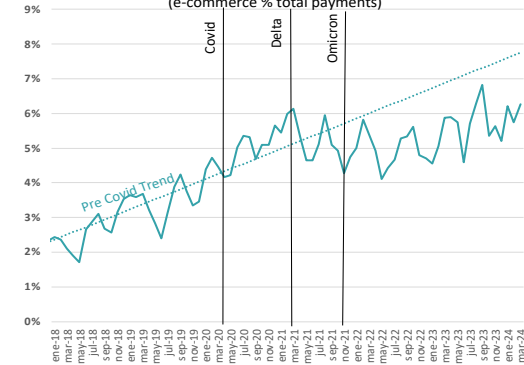
**Food & Non Alcoh. Beverage**  
(e-commerce % total payments)



**Communications**  
(e-commerce % total payments)



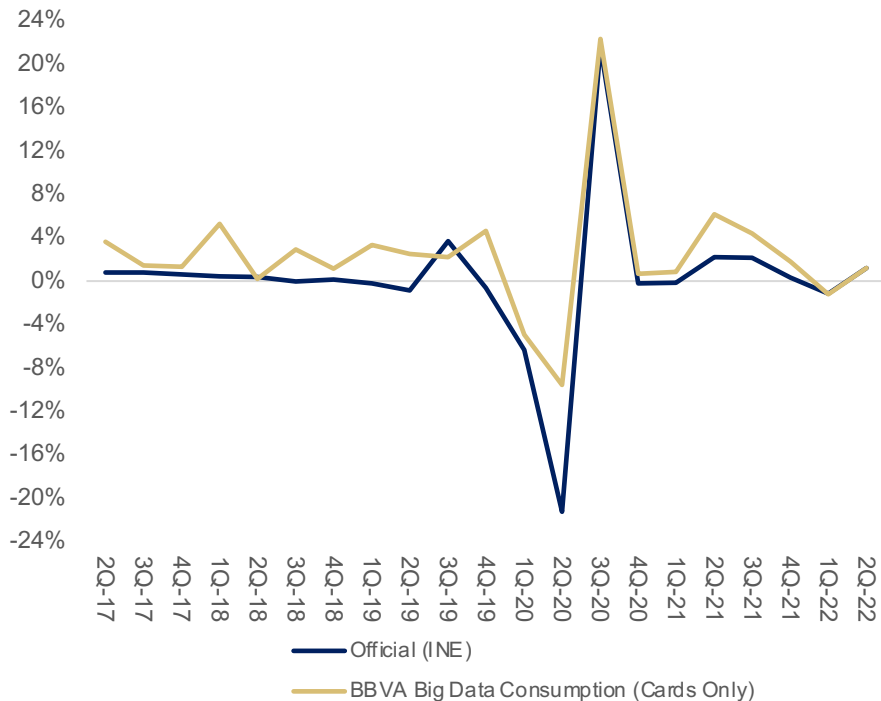
**Other (Finance & Insurance)**  
(e-commerce % total payments)



# Beyond this Big Data can be used to Nowcast Official Statistics: Cards

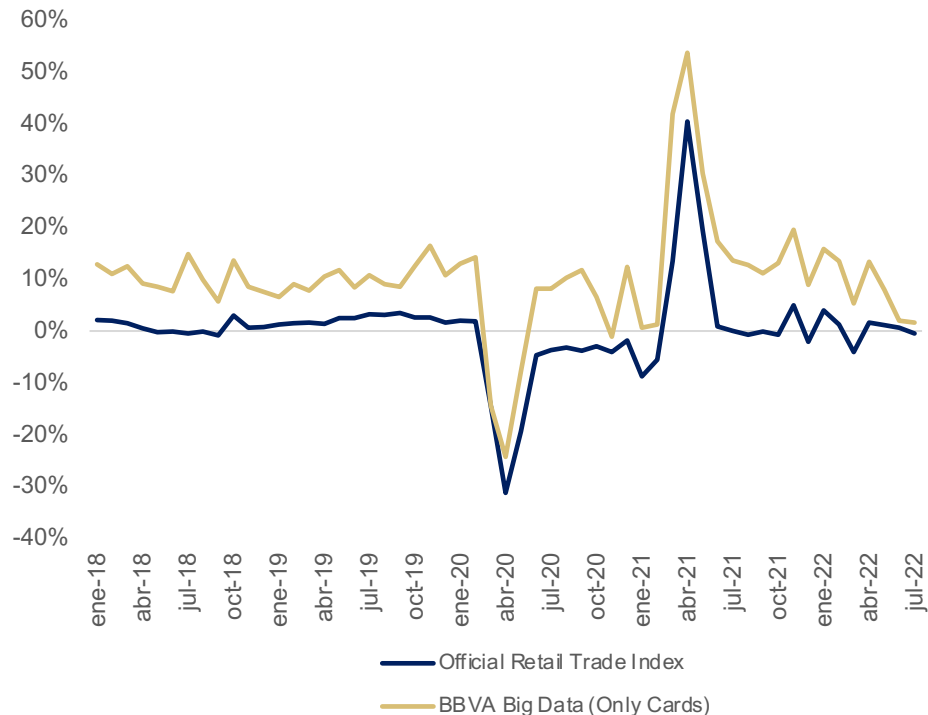
## Spain: Real Household Consumption vs BBVA BigData

(Sadj %QoQ Growth Rate CPI deflated)



## Spain: Retail Sales Index vs BBVA Big Data

(Sadj %mom Growth Rate. Big Data CPI deflated)

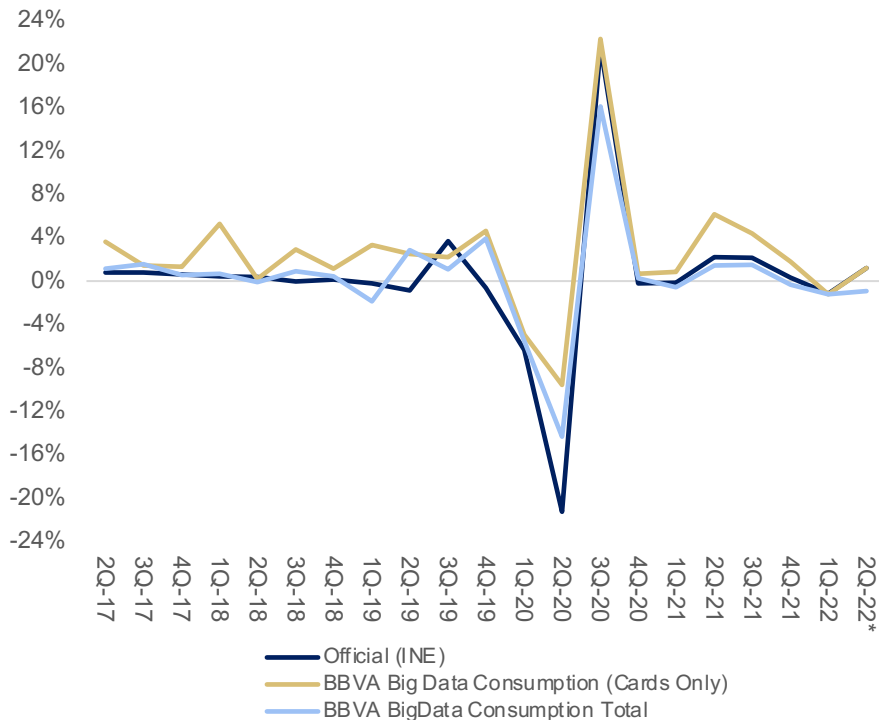




# Beyond this Big Data can be used to Nowcast Official Statistics: and the rest

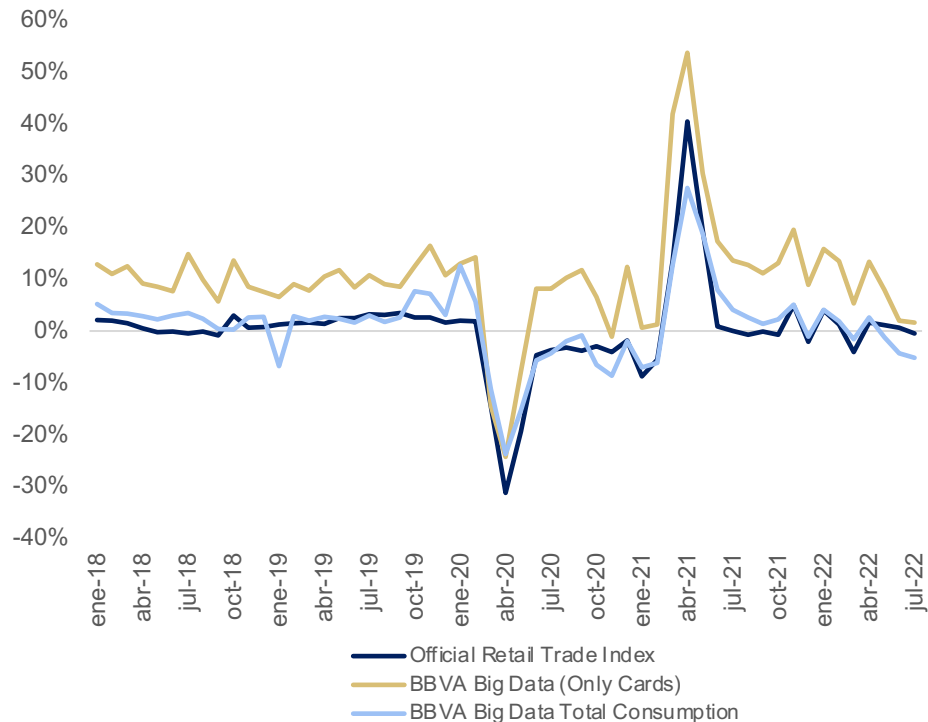
## Spain: Real Household Consumption vs BBVA BigData

(Sadj %QoQ Growth Rate CPI deflated)



## Spain: Retail Sales Index vs BBVA Big Data

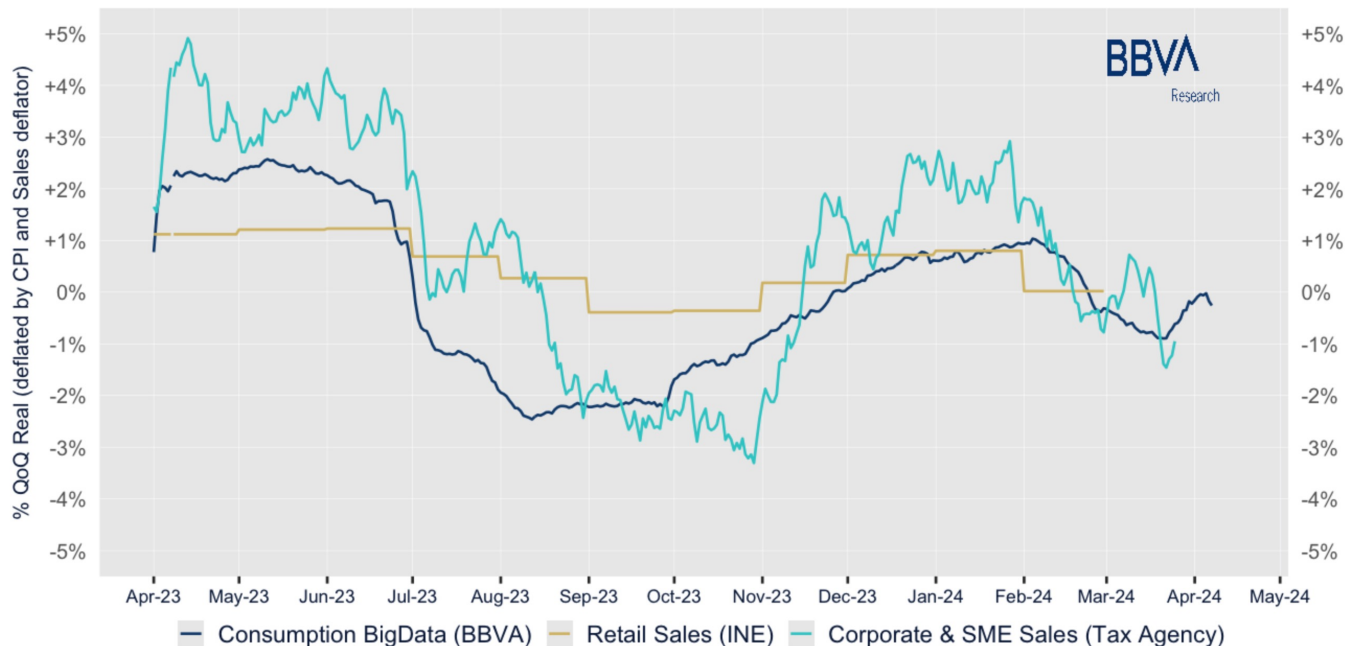
(Sadj %mom Growth Rate. Big Data CPI deflated)



# This allows us to track National Consumption in Real Time & High Definition on daily basis by tracking directly the indicators ...

## Consumption (BBVA), Retail sales (INE) & Corporate & SME Sales (AEAT)

(Real deflated by CPI & Sales deflator. Cum 91D % Var. QoQ Sadj)



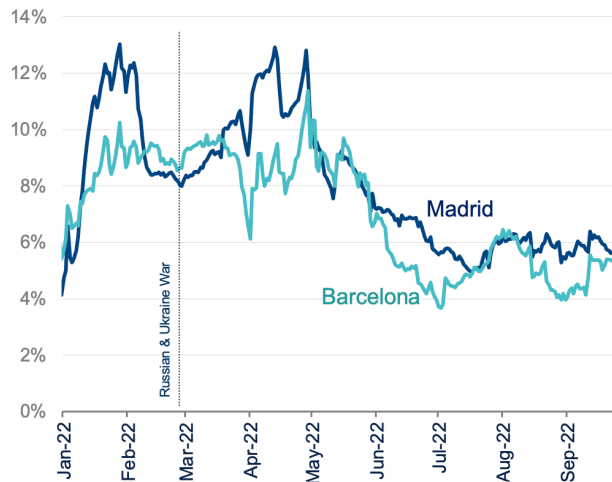
Source: BBVA Research, INE, AEAT

# ... and Zooming Out: National, Regional, Provincial, Urban & Zip Code..

National Accounts in a World  
of Naturally Occurring Data

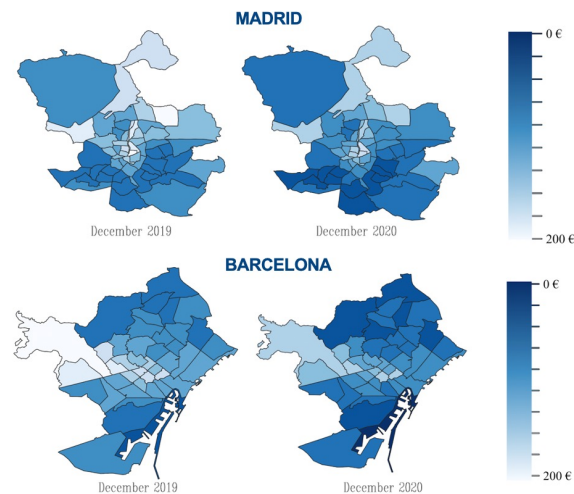


**Urban Big Data Consumption: Barcelona & Madrid**  
(Moving Average 28D. YoY Nominal)

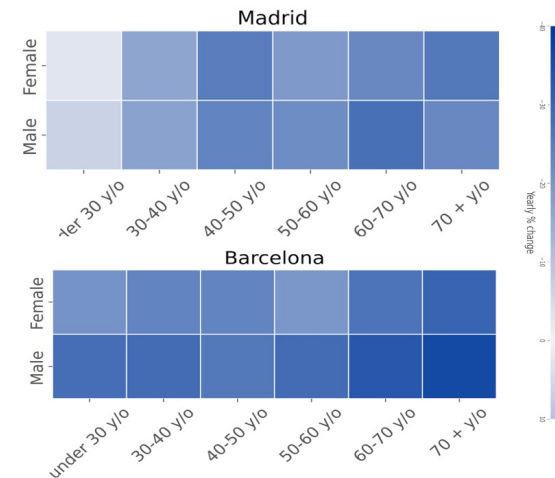


Source: Buda et al (2022) BBVA Research

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



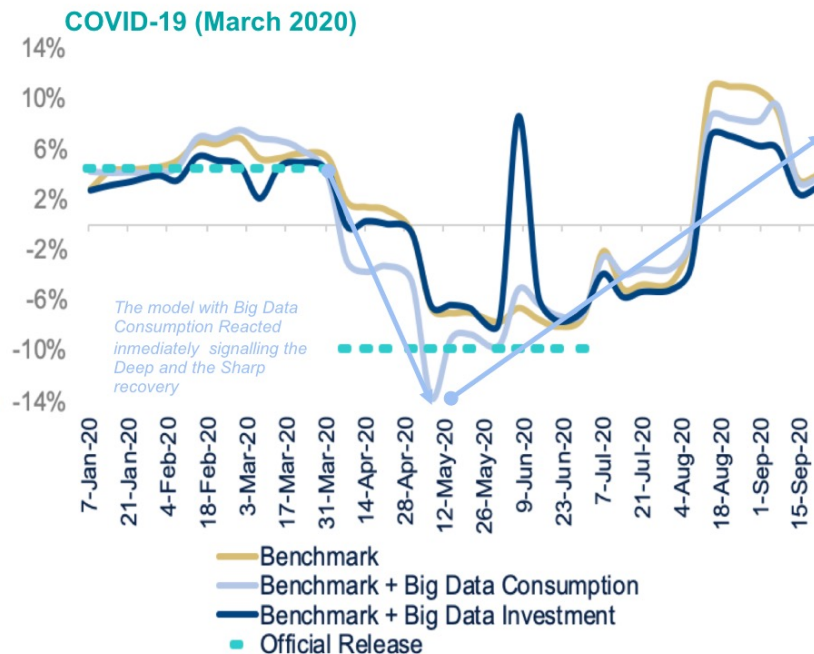
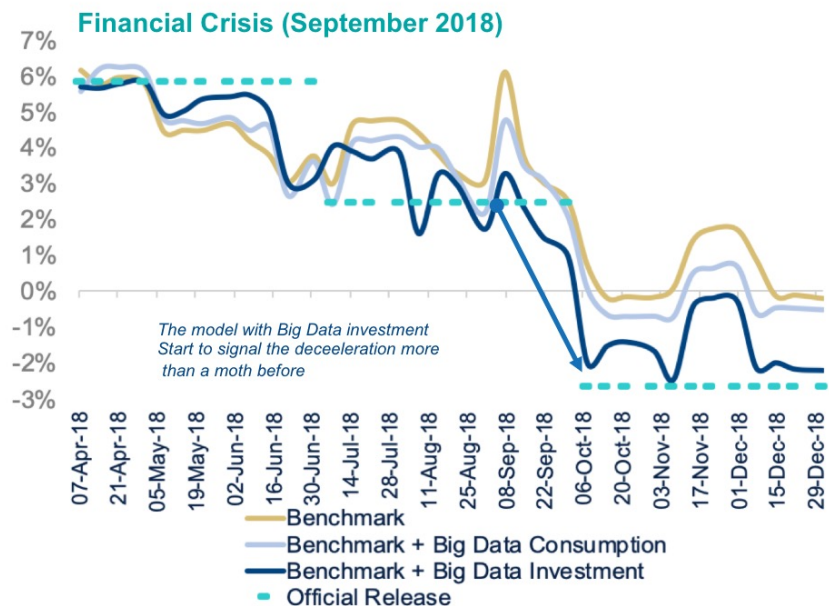
**Change in Consumption Restaurants & Hotels in 2020**  
(December 2020 vs December 2019, % YoY)



# Having “timely information” can be very useful for households, corporates and policy makers... as we experienced in Turkey

## TURKEY: NOWCASTING FINANCIAL CRISIS (SEPT 2018) & COVID CRISIS (MAR 2020)

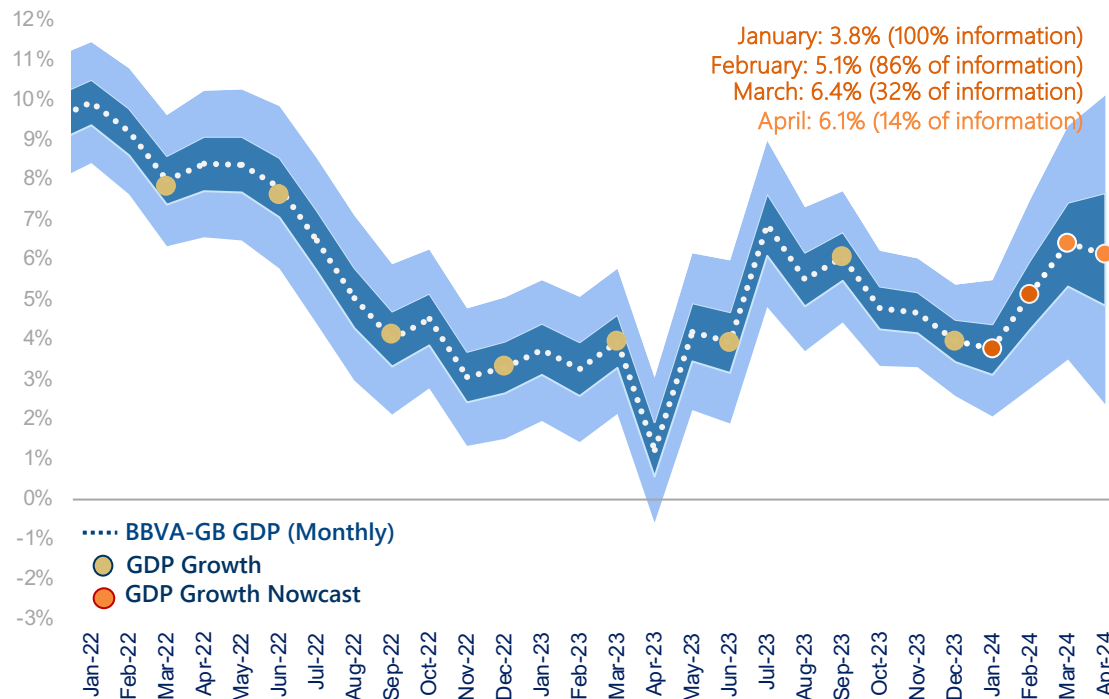
(quasi real time nowcasting with and without Big Data Indexes vs Benchmark)\*\*



Source: Own Elaboration

# The recent data ..signal still strong Domestic Demand (helped by some base effects) and this entails some challenges

## Turkey: Nowcasting Model Including Big-Data (% YoY)

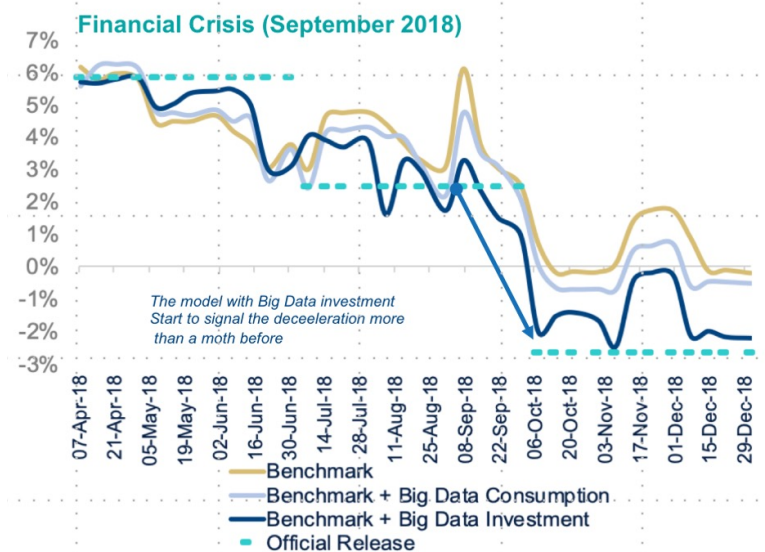
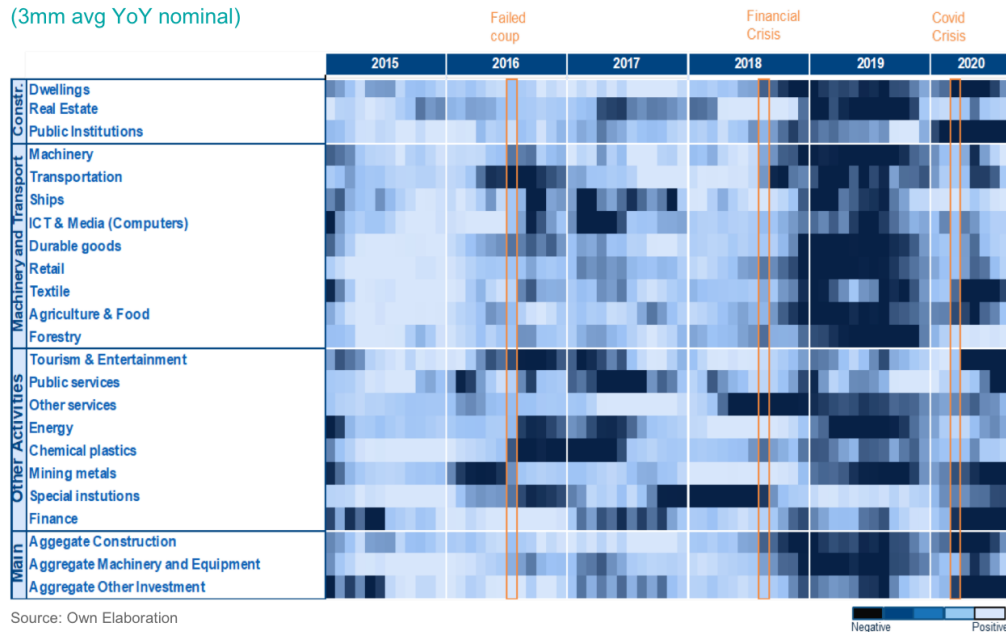


Source: Garati-BBVA Research

# Big Data information was very useful in Turkey to identify the different nature of Investment Shocks to anticipate GDP movements

## TURKEY:GB-BBVA BIG DATA INVESTMENT HEAT MAP

(3mm avg YoY nominal)

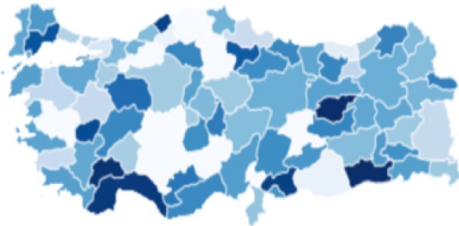


Source: Barlas, Guller, Ortiz & Rodrigo. Investment in Real Time and High Definition

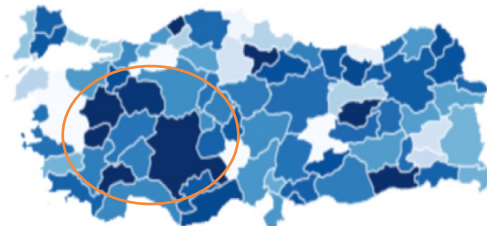
# ... and the Geographical impacts of Global Value Chains

## TURKEY:GB-BBVA Big DATA INVESTMENTS GEO-MAPS RECOVERY AFTER THE COVID-19

Change in Machinery Investment, February 2020 Average (YoY)

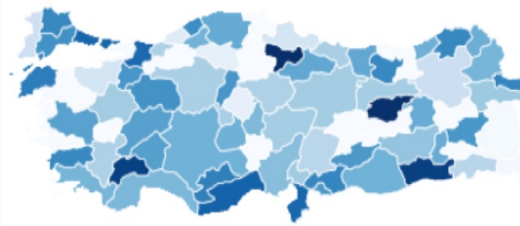


Change in Machinery Investment, April-May 2020 Average (YoY)



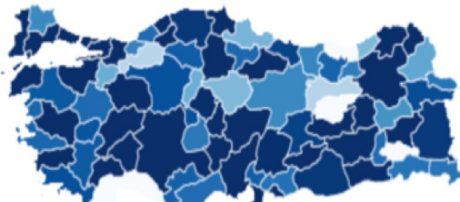
Covid trigger GVA problems In Textile & Cars  
Export platform Provinces (Center Anatolia)...

Change in Machinery Investment, July-August 2020 (YoY)



... but rapidly recovers once  
GVA problems ease

Change in Construction Investment, February 2020 Average (YoY)



Change in Construction Investment, April-May 2020 Average (YoY)



Legacies with 2018- 2019 Financial Crisis with permanent effects in Construction

\*White stands for YoY rates >50% nominal & Dark Blue <-30% nominal declines

\*White stands for YoY rates >50% nominal & Dark Blue <-30% nominal declines

Source: Barlas, Guller, Ortiz & Rodrigo. Investment in Real Time and High Definition



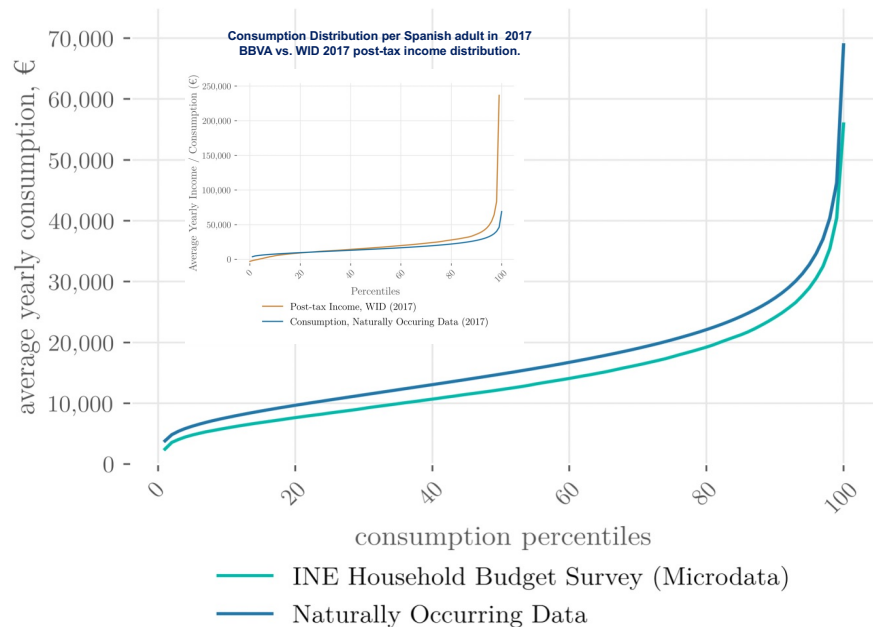
# Big Data for Policy Analysis: Designing Smart Policies

---



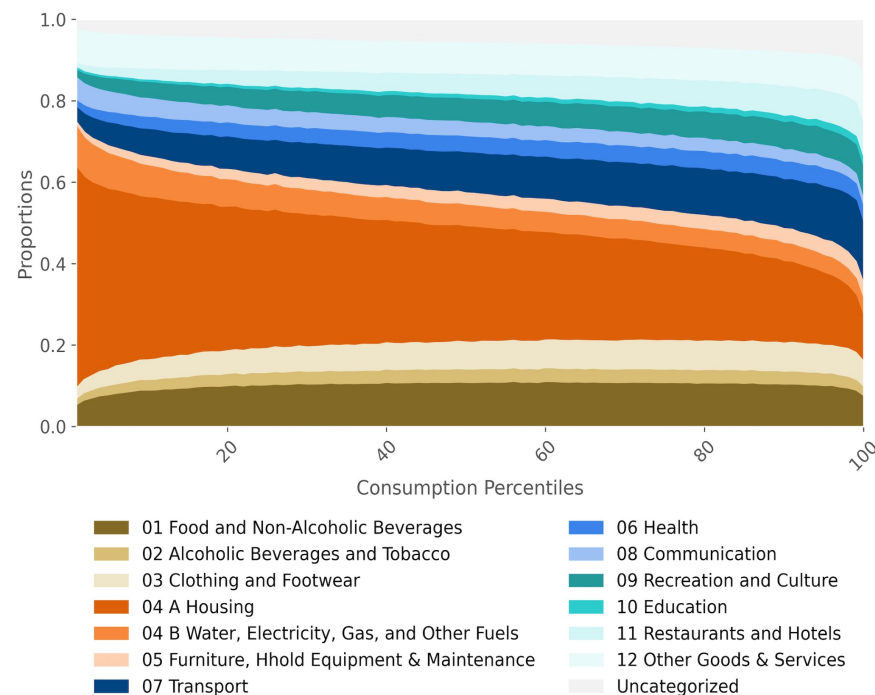
# The potential for “High Definition” economics is higher than geographical

## Spain: Consumption distribution by percentile of Consumption (2017)



Source: Buda et Al (2022) National Accounts in a World of Naturally Occurring Data: A Proof of Concept for Consumption. CEPR Discussion Paper

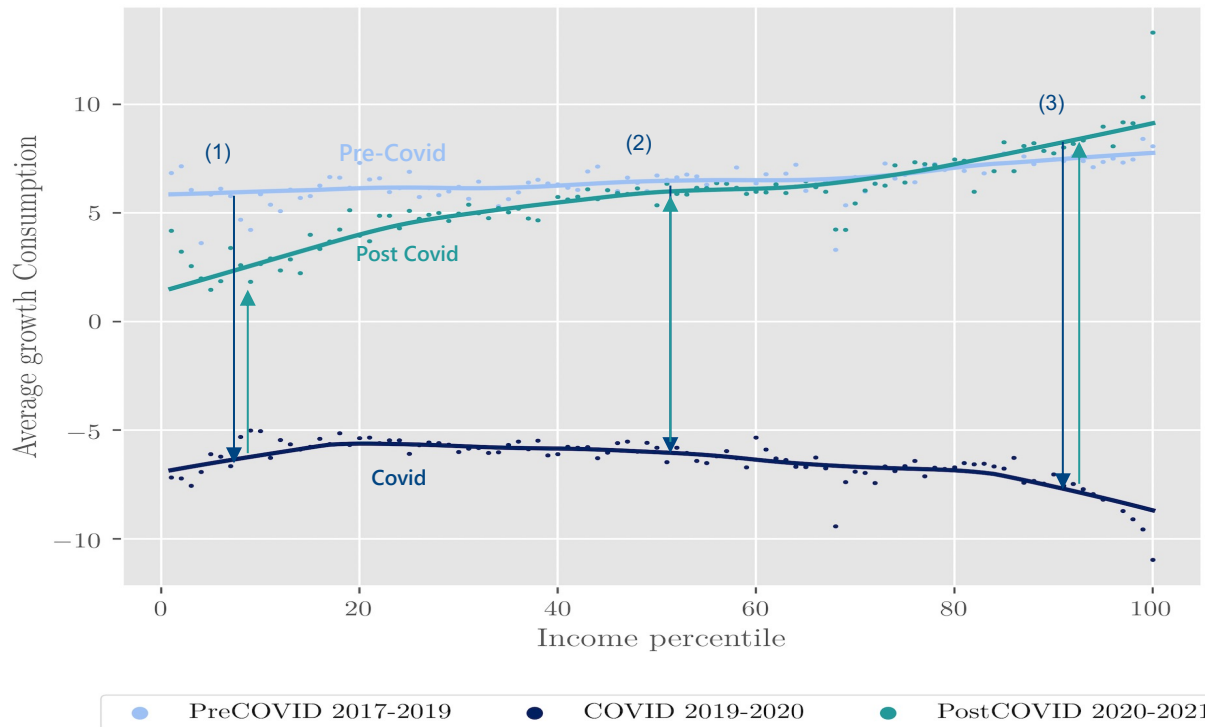
## Spain: Consumption distribution disaggregated by categories pre Covid (2017)



Source: Buda et Al (2022).

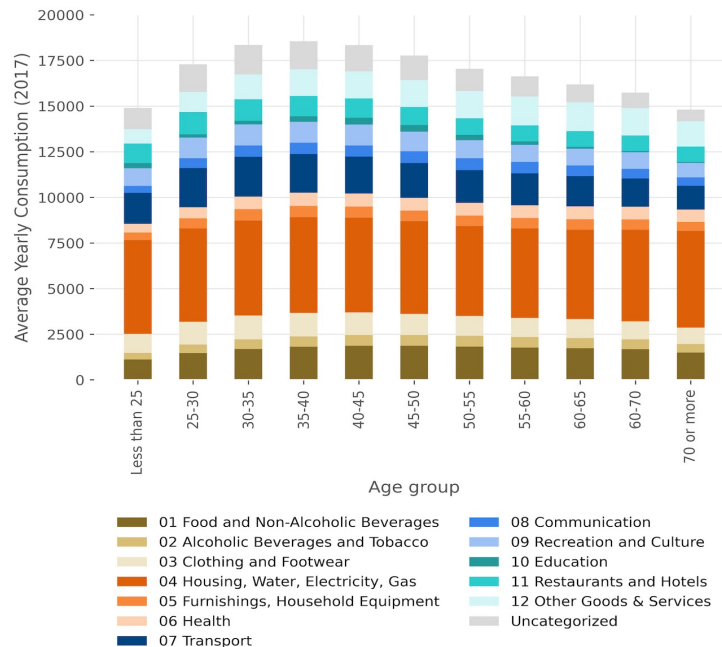
# Dealing with heterogeneity is very relevant: One size "Does Not" fit all

## Spain: Consumption growth by Income Percentile



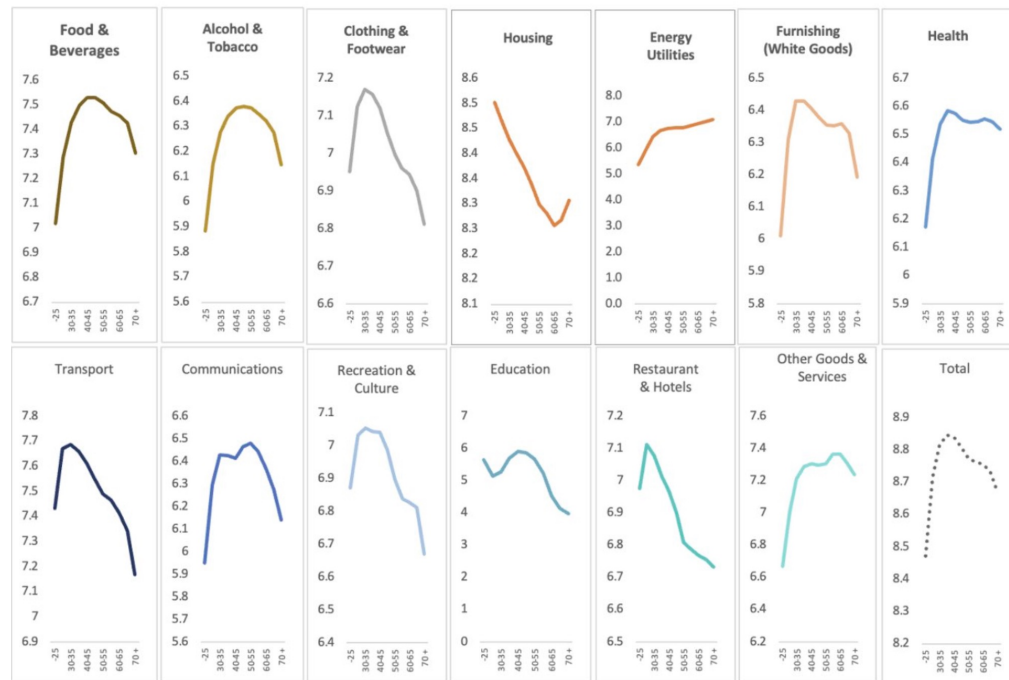
# Consumption Distributional Accounts: Age Matters, specially for the old Europe..

## Spain: Consumption per Age (2017)



## Spain: Consumption by Category and Age (2017)

(log scale Source: Buda et al (2022) National Accounts in a World of Naturally Occurring Data..)

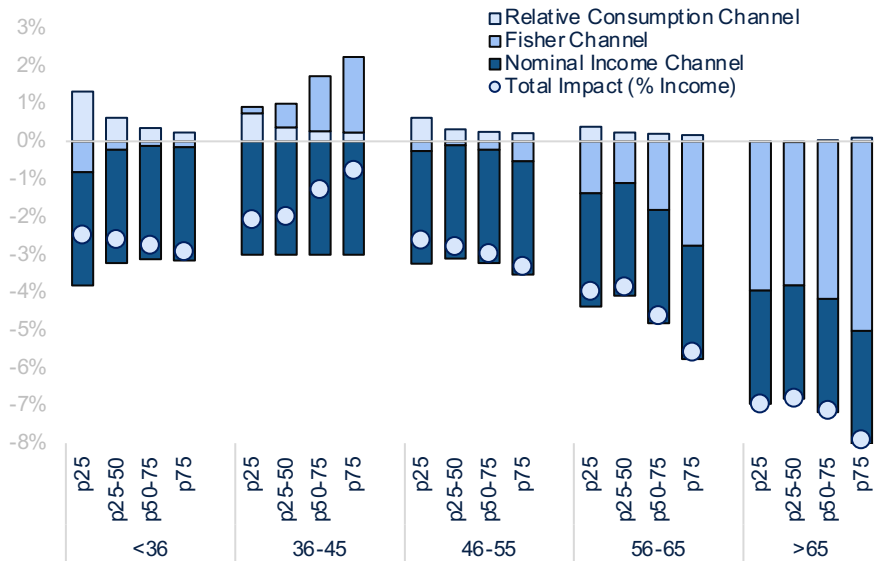


# The Impact of the inflation shock could be different depending of your income conditions, age and debt..

BBVA  
Research  
The Heterogeneous Impact of  
Inflation on Households  
BANCODE ESPAÑA  
Eurosistema

Spain: Effect of "Inflation Shock" on Household by Income & Age  
(Unexpected Inflation Shock by Age & Percentiles of Income. in % of Income)

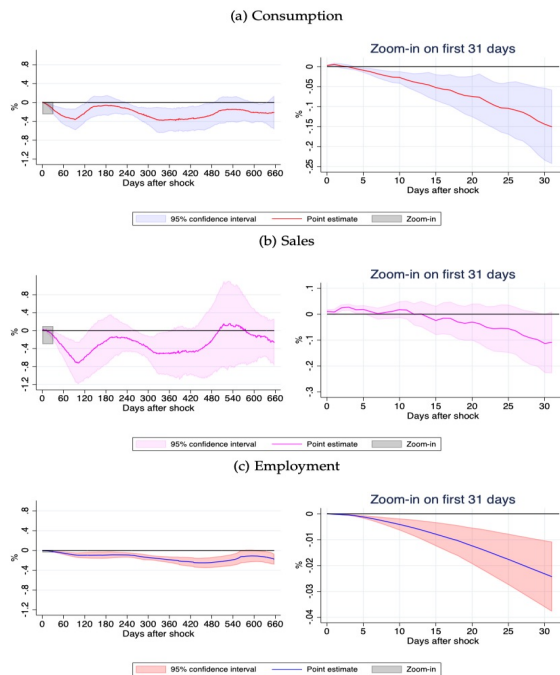
Indebted  
Mid-Age  
Balanced  
the impact  
(Fisher Effect)



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

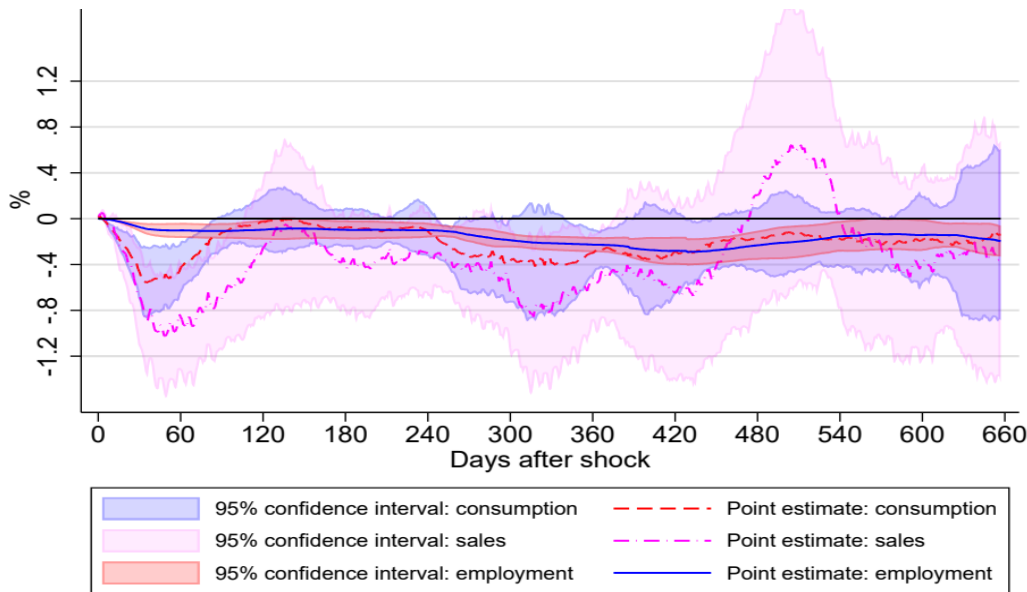
# and Big Data can be useful to analyse the transmission of Monetary policy: Lags can be shorter than you think

Figure 1: Daily response of economic activity variables to a monetary policy shock



## Spain: Effects Monetary Policy Shock on Consumption, Sales & Employment

(Daily Indicators. One Standard Deviation Monetary Policy Altavilla et al Shock)

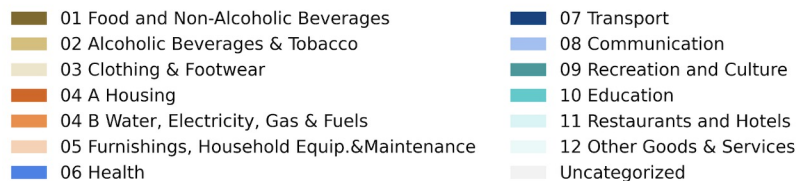
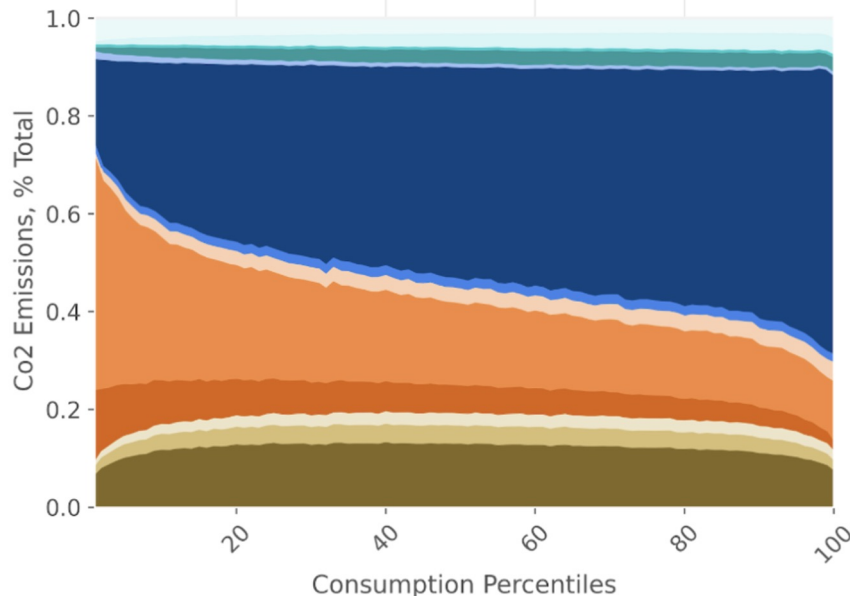


Notes: The monetary policy shock is one standard deviation. The responses are reported in levels. We obtain them by cumulating year-on-year changes. The 95% confidence intervals are computed from heteroskedasticity-robust standard errors.

Source: Buda et al (2023) Short and Variable Lags. CEPR Discussion Paper

# Sustainability is a very promising area: Here.. the Evil is in the details...

Distribution of CO2 Emissions by Spanish Households  
(as % of total expenditure 2017)



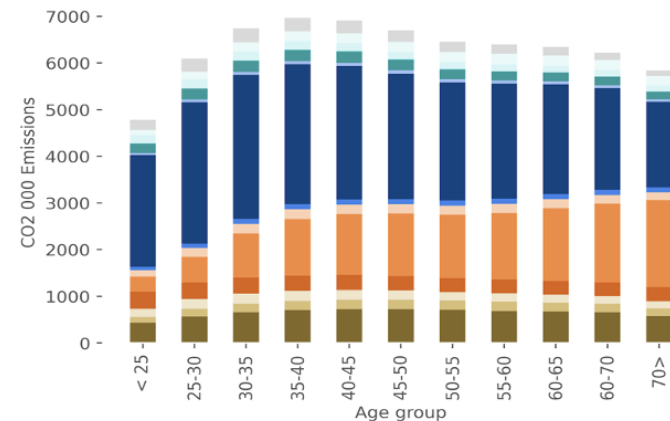
Some Policies have higher co for low income families

National Accounts in a World of Naturally Occurring Data



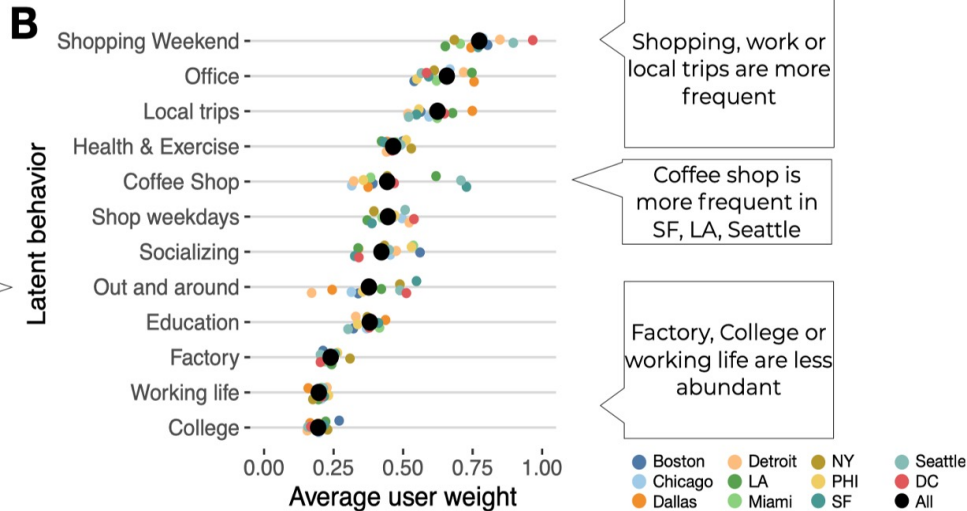
There is a Relevant Margin for Heterogeneous Policies

Distribution of CO2 Emissions by Age  
(as % of total expenditure 2017)



# ... and our Life-Styles can affect sustainability as well?

## US Cities Lifestyles: Latent Lifestyles

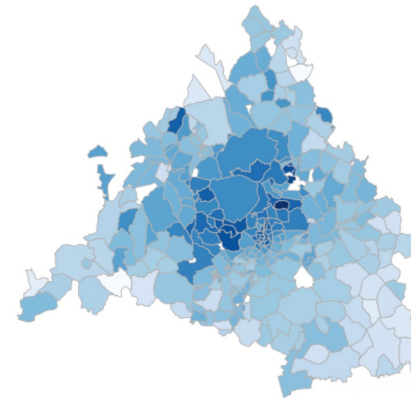


## Urban Life-Styles and Sustainability

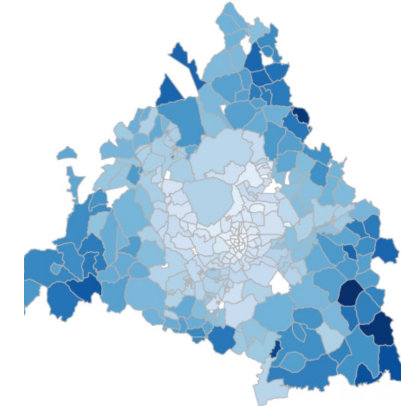


## Madrid Lifestyles: Inner vs Outer Lifestyles Madrid Region (Average Lifestyle intensity by zip-code)

### Go Bars & Restaurants



### Car Trips



Source: Esteban Moro



Source: BBVA Research & MIT Media Lab (forthcoming)



# Text as Data: Non-Structured Information for Social, Political and geopolitical Analysis

---





... and the development of **New Algorithms** allowed harnessing "Text as Data"... The history of LLM is short but has been very productive...

## Bag of Words Model

It is individual Words  
What Matters...

Boolean Searches  
Dictionary Assisted  
Sentiment Analysis

## Topic Models

...But structural  
Words dependence too  
(Global  
Co-occurrence)...

LDA  
Dynamic Topic Models  
(Factor Models of Text)

## Word Embeddings

... Local Context  
is a more Natural Guide  
(Synonymy Vector)...

Global Vectors  
Word2Vec

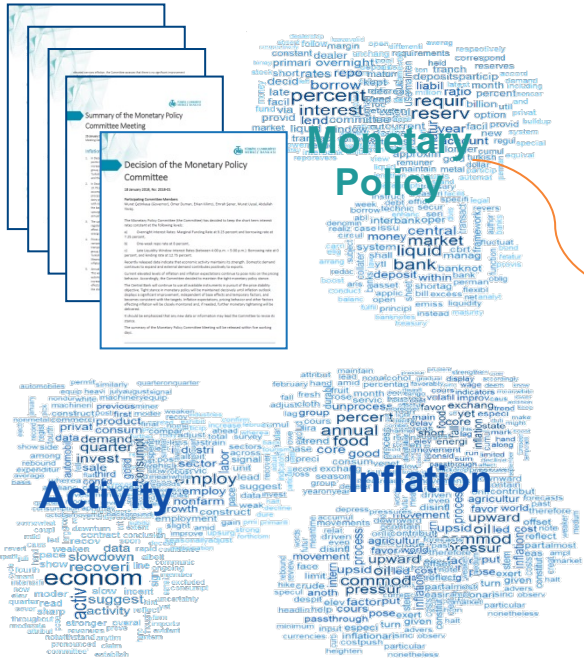
## Large Language Models (LLM)

... Addressing Polysemy  
requires "Attention"  
(weighing embeddings) for  
Prediction Tasks

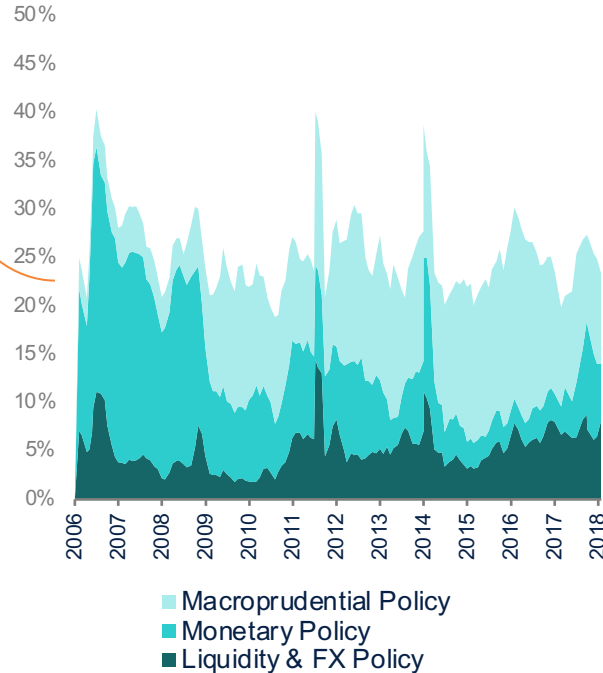
Transformers  
BERT  
LLaMa  
GPT

# Text mining techniques offer lots of opportunities ...You can go inside of a Central Bank Reports to learn what they say and how they talk...

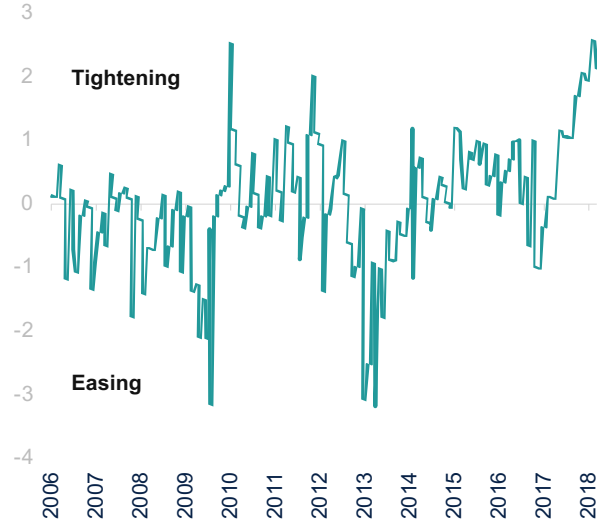
## CBRT Topics Evolution



## Topics in Detail Monetary Policy

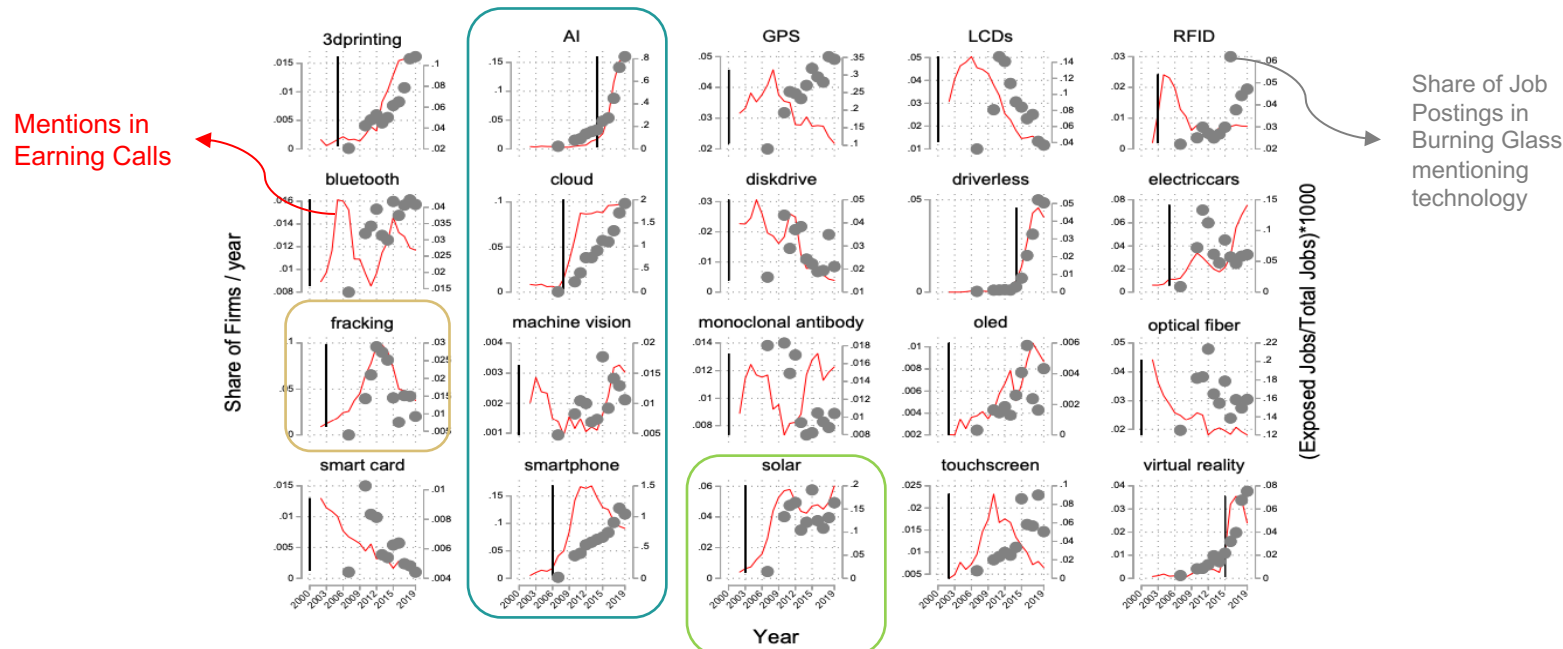


## Monetary Policy "Statements" Sentiment



# Examine Patent, Earning Calls and Job Posts information to explore who are the "Disruptive Technologies"...

## Exposure of Disruptive Bigrams in Earning Calls Job Offers (Time Evolution)



# Or read Corporate Transcripts .. and use LLMs to make some inference on Potential Investment

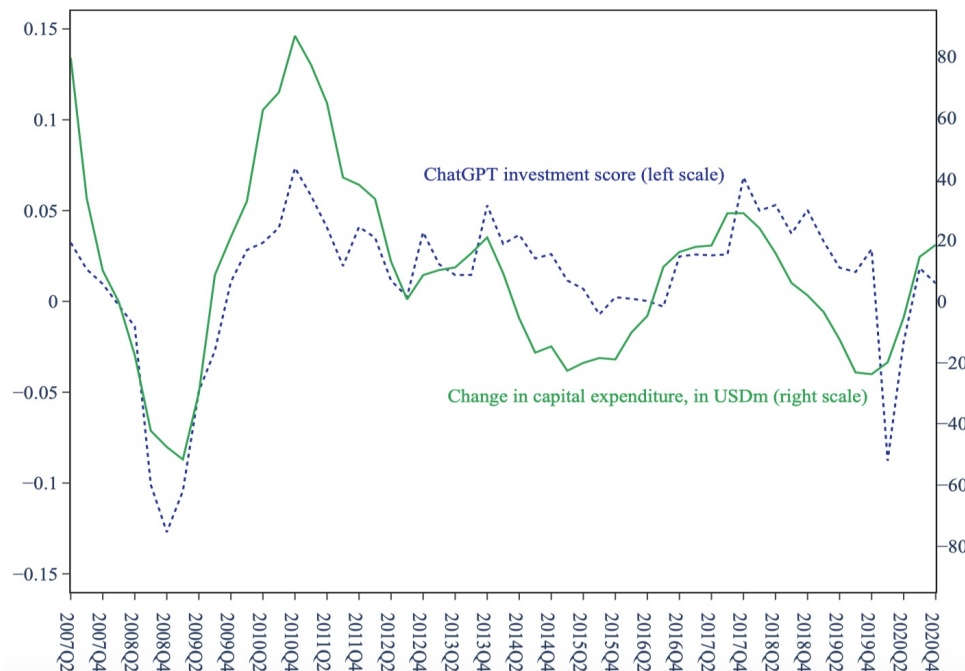
(b) Bigrams associated with low ChatGPT investment scores.



(a) Bigrams associated with high ChatGPT investment scores.



(c) Chat GPT Investment Score vs. Realized Investment



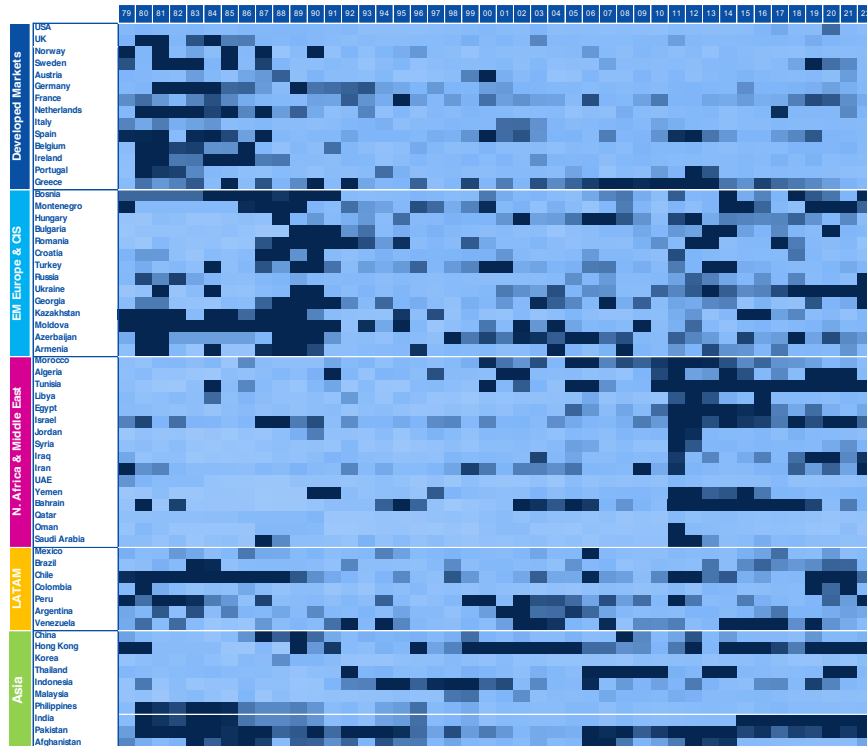
## Chat GPT Investment Score, Tobin's Q, and Investment

	(1)	(2)	(3)	(4)
	<i>Capital Expenditure<sub>t+2</sub></i>			
<i>ChatGPT Investment Score<sub>t</sub></i>	0.966*** (15.64)	0.795*** (13.24)	0.683*** (12.16)	0.638*** (11.37)
<i>Total q<sub>t</sub></i>		0.379*** (12.44)		0.177*** (6.53)
<i>Capital Expenditure<sub>t</sub></i>			0.115*** (9.98)	0.114*** (9.92)
<i>Total Cash Flow<sub>t</sub></i>			0.889** (3.00)	0.535 (1.83)
<i>Leverage<sub>t</sub></i>			-2.795*** (-16.94)	-2.535*** (-14.97)
<i>Size<sub>t</sub></i>			-0.006 (-0.14)	-0.008 (-0.19)
Firm FE	Y	Y	Y	Y
YearQtr FE	Y	Y	Y	Y
R-squared	0.694	0.697	0.707	0.708
N	74,586	74,586	74,586	74,586

This figure shows the time series of the average quarterly ChatGPT investment score and average future four-quarter change in capital expenditure. ChatGPT investment score is calculated based on conference call texts of the firm (described in Section 3.1). We calculate the change in capital expenditure as the difference between the average capital expenditure for the four quarters following the current quarter (t+1 to t+4) and the average capital expenditure for the four quarters prior to the current quarter (t-4 to t-1).

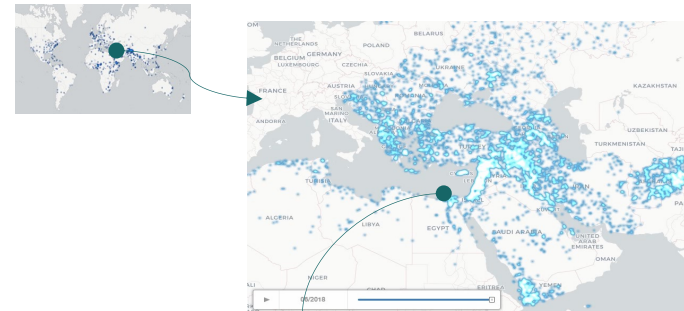
There is huge potential for social sciences where structured information was scarce. Now we can quantify now how the World feels in Detail..

## World Social Unrest Intensity Heat Map 1979- 2022

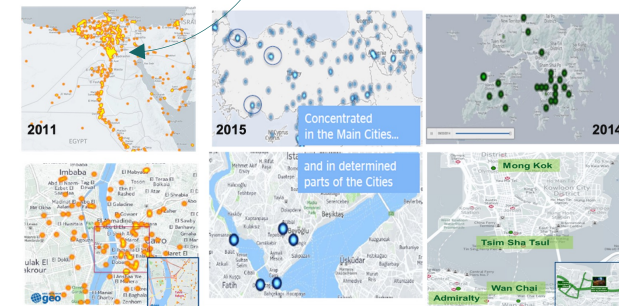


Source: [www.qdelt.org](http://www.qdelt.org) & BBVA Research

## Conflict intensity map 2017-18 (number of conflicts/ total events)



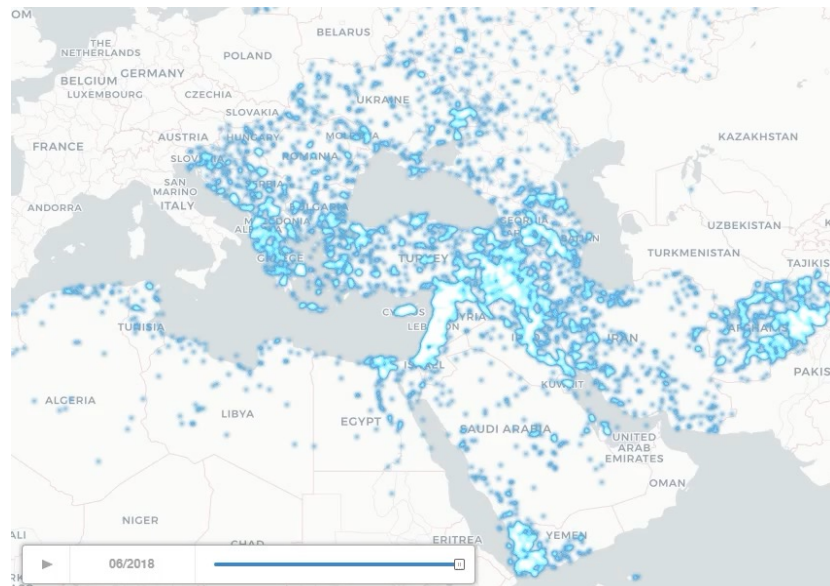
## SOCIAL UNREST EVENTS ACROSS THE WORLD: CAIRO, ISTANBUL & HONG KONG PROTEST EVENTS



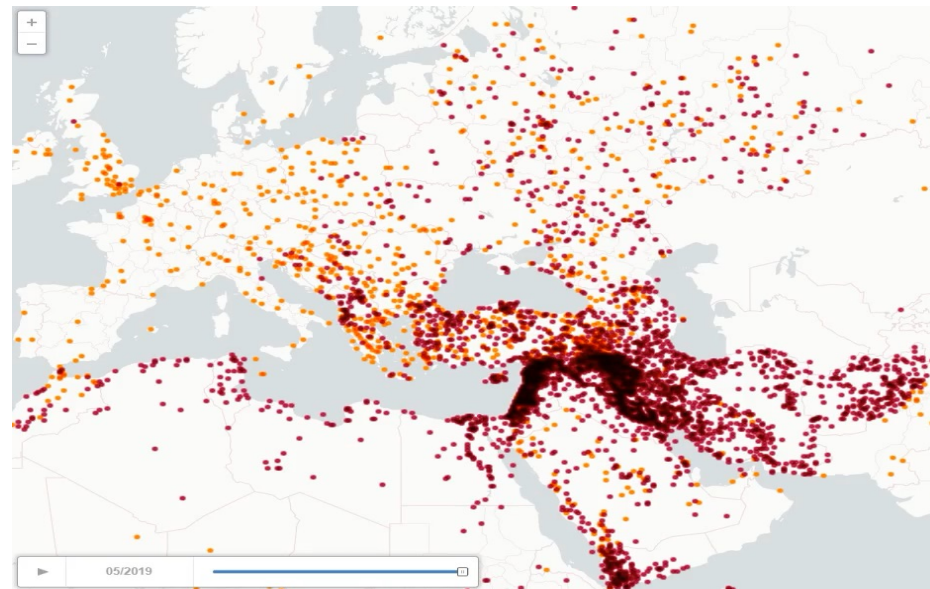
Source: [www.qdelt.org](http://www.qdelt.org) & BBVA Research

# The information is geo-localized, offering a high detailed analysis to track events and analyze contagion effects

**World Conflict Intensity Map: Inside the Middle East 2017-18**  
(number of conflicts/ total events)

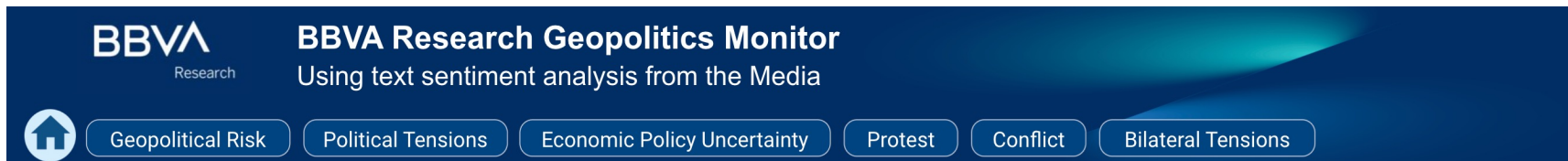


**BBVA Research: Syrian Crisis Refugees Flows Map 2015-19**  
(Number of Media Citations of Migrant inflows and outflows)



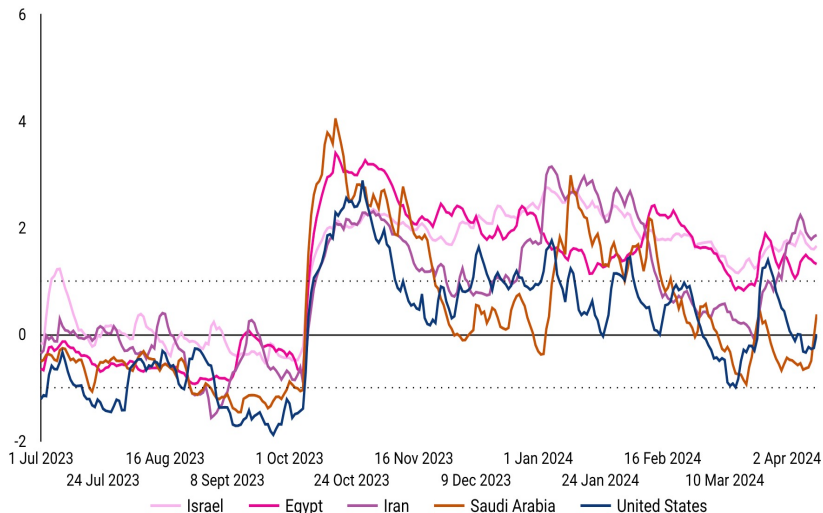
● Refugees Outflows ● Refugees Inflows

# We are using our BBVA Research Geopolitics Monitor: ... to track what is happening in Middle East and collateral effects on other countries...



## BBVA Research Geopolitical Risk Index: Middle East and US

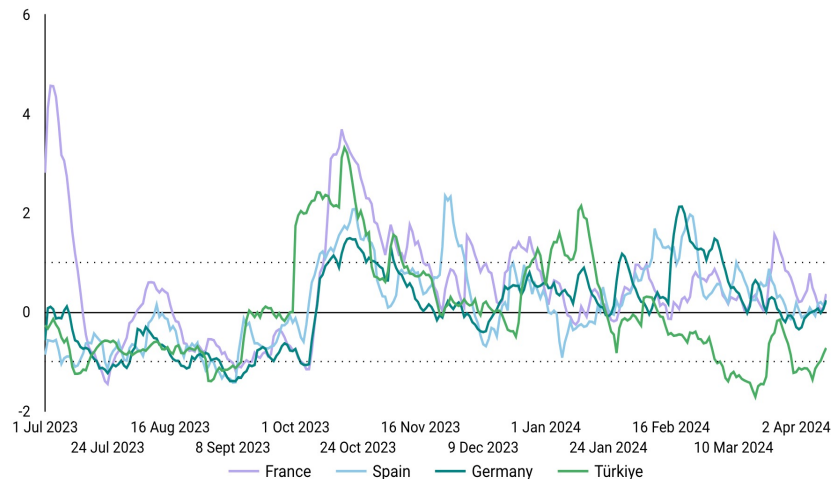
(28-day weighted moving average, normalized by its own country history. Positive (negative) values indicate greater (lower) risk relative to the mean of period 2021-nowadays)



Source: Big Data Analysis Unit (BBVA Research)

## BBVA Research Protest Intensity Index: Europe

(28-day weighted moving average, normalized by its own country history. Positive (negative) values indicate greater (lower) risk relative to the mean of period 2021-nowadays)





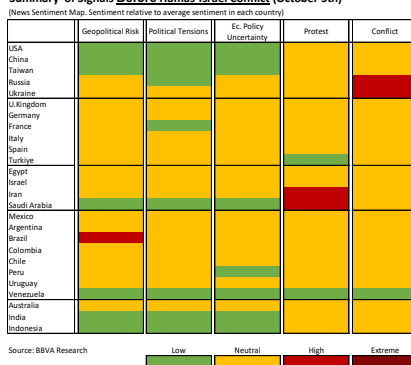
# The Geopolitical risk and Conflict sentiment is high ... while the Economic policy uncertainty remains safe so far ...

## BBVA Research Geopolitics Monitor main "Signal" changes after the Hamas-Israel Conflict

(Signals in Geopolitical Risk, Political Tensions, Economic Policy Uncertainties, Protest and Conflict according to Standard Dev)

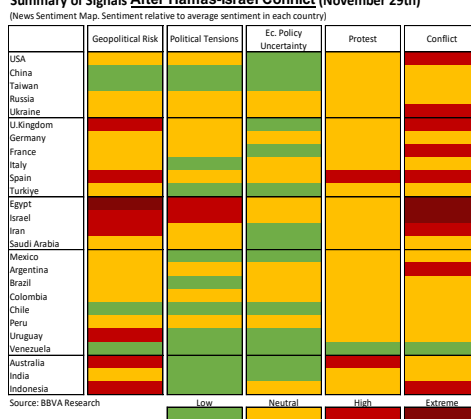
### Before the conflict

Summary of Signals Before Hamas-Israel Conflict (October 5th)



### One Month After

Summary of Signals After Hamas-Israel Conflict (November 29th)



### April 15th

Summary of Signals After Hamas-Israel Conflict (April 15th, 2024)

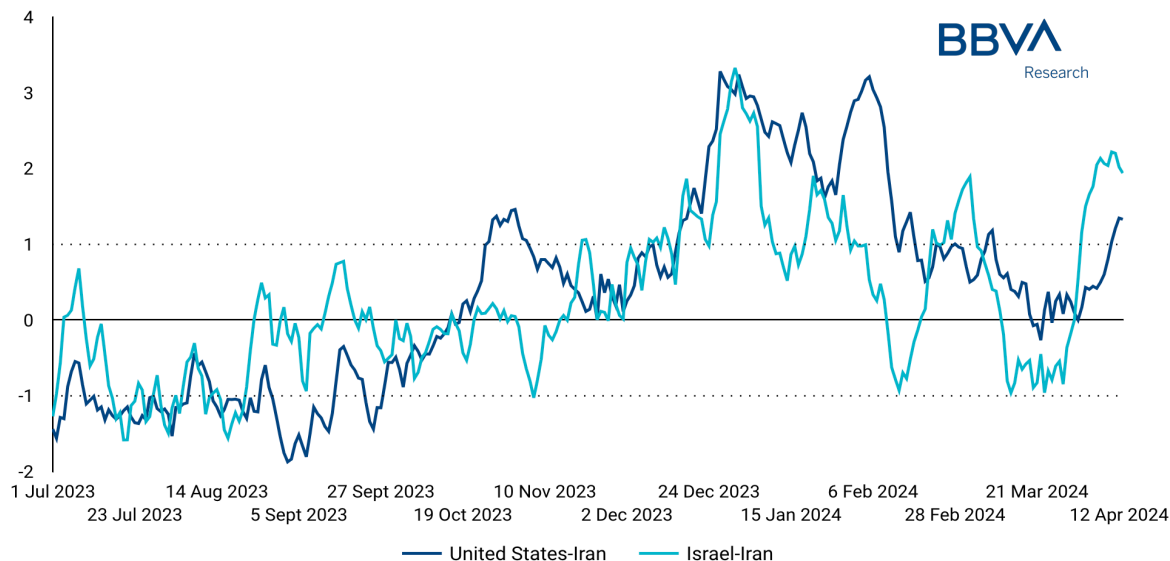


Contagion restricted to Geopolitical & Conflict sentiments... Economic Policy Resilience.. so far

# .. But we need to monitor the situation very carefully... as the extension of the Regional Conflict ... pose some risks

## Bilateral Tensions index

(28-day weighted moving average, normalized by its own country history. Positive (negative) values indicate greater (lower) uncertainty relative to the mean of period 2019-nowadays)

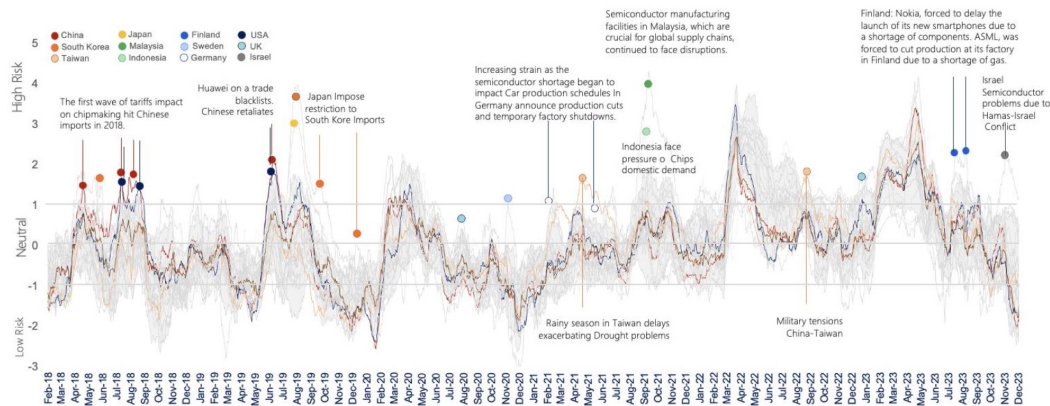


Should you require any additional relations, please do not hesitate to contact [data-science-community-research.group@bbva.com](mailto:data-science-community-research.group@bbva.com). This index collects all the events classified in the CAMEO taxonomy as verbal conflict or material conflict where the actor that makes the action is first one that appears in the legend and the actor that receives that action is the second country or region in the legend. The final indicator is a weighted product made of the tone (i.e. associated sentiment) and relative coverage (i.e. number of news related to each pair of countries/total number of published news) from GDELT. The indices are constructed using daily data, normalized (minus average and divided by standard deviation) and then smoothed by converting them into a 28-day weighted moving average to reduce noise and enhance the accuracy of the indices. For further information, please refer to this link.

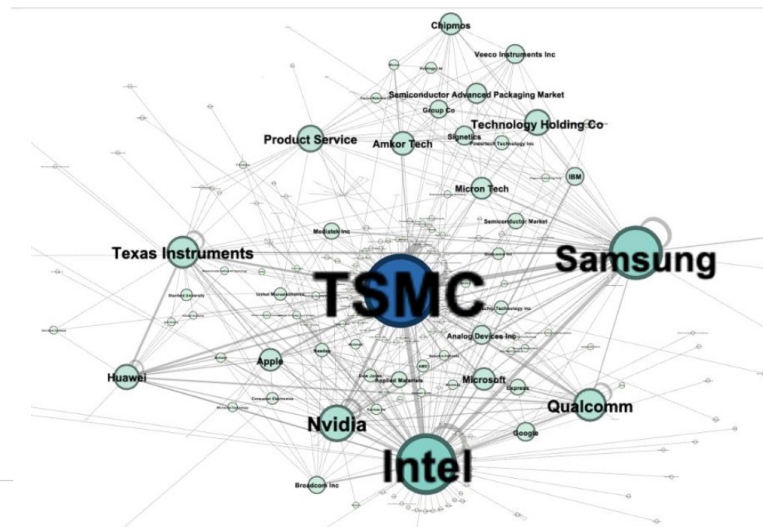
Data Last Updated: 15/04/2024 09:03:12 | [Privacy Policy](#)

But not all the sentiment is bad.. The Bottlenecks Semiconductors have eased ... Which has been transmitted to some tech players..

## Semiconductors Networks Sentiment & TSMC Networks News on Big Techs 2022



Source: BBVA Research and [www.qdelt.org](http://www.qdelt.org)



# Beyond all this examples...the new LLM Generative Models can help Economist dramatically in our daily lives ...

## Large Language Models Capabilities in Economics

### Ideation and Feedback (generating, selecting, developing of ideas):

- Brainstorming ideas
- Offering feedback
- Providing counterarguments
- Evaluating ideas

### Writing:

- Synthesizing text
- Providing examples
- Evaluating text
- Editing text (for mistakes, style, clarity, simplicity, ...)
- Generating catchy titles & headlines
- Generating tweets / promotional materials

### Math (LaTeX):

- Setting up models
- Deriving equations
- Explaining models

### Background Research:

- Researching the literature
- Summarizing text
- Formatting references
- Translating text
- Explaining concepts

### Coding:





- Writing code
- Explaining code
- Translating code
- Debugging code

### Data Analysis:

- Creating figures
- Extracting data from text
- Reformatting data
- Classifying and scoring categories
- Extracting sentiment
- Simulating human subjects

3 – green	capability is very well-developed & quite robust, highly useful – recommended
2 – yellow	capability works in limited circumstances, occasionally useful – good to experiment
1 – red	capability is emerging, useful with serious limitations – experiment at your own risk

## Some Examples

- 
**Classifying Text:**  
Identifying Remote Work
- 
**Data Analysis:**  
Extracting Sentiment from Central Banks
- 
**Data Analysis:**  
Extracting Sentiment from Earning Calls
- 
**Background Research & Data Analysis:**  
Summarizing + Identify Text + Figures



# Is Big Data & AI relevant for Economists?

---

## BigData & AI for Economists. Speculative & Personal Conclusions!!

- Big Data & AI are powerful tools for Economists
- They provide very useful Real Time & HD Information for Analysis
- ... also relevant for the design of Smart Policies
- Text Mining & LLMs are also a “+” with the potential to change lots of things in Social Sciences including Economics...
- But a solid background is also key to fully exploit Big Data & AI
- Economist are well prepared to exploit Data Science Techniques in many Disciplines..So take advantage of it !!

Thanks

# Thanks...

## You can follow our works at:

 @alvaroortiz1968



<https://www.linkedin.com/in/alvaro-ortiz-vidal-abarca/>

# Big Data & AI for Economics

## Economic in Real Time & High Definition for Smarter Policies

---

Dr. Alvaro Ortiz  
Head of Big Data & AI Economic Analysis  
BBVA Research

Seminar at Bilkent University. April 2024