BBVA
Research

# Monitoring social, economic and geopolitical events through Big Data

**48th World Continuous Auditing and Reporting Symposium**

**Tomasa Rodrigo**, September 2020

Creating Opportunities

# Index

Creating Opportunities

# 01

# Opportunities in the digital era. Big Data at BBVA Research

# Traditional data could not answer some relevant questions

Geopolitical events

Political events and social reaction

Natural disasters and epidemics

The use of **Big Data** and **Data science** techniques allows  us to quantify these trends

# New framework in the digital era…

- New availability of data
- Better and faster infrastructure
- New answers to old questions

- Combination of historical data with real time data
- Advanced data science techniques and algorithms
- Higher computational abilities to face more data granularity

# …which needs the development of new competences to take advantage of it

Making the right questions

Developing the data management and programming capabilities to work with large-scale datasets

Deepening the statistical and econometric skills to analyze and deal with high-dimensional data

Interpreting the results: summarize, describe and analyze the information

New data may end up changing the way in which  economists approach empirical questions and the tools they use to answer them

# Internal and external diffusion of what we do at BBVA Research
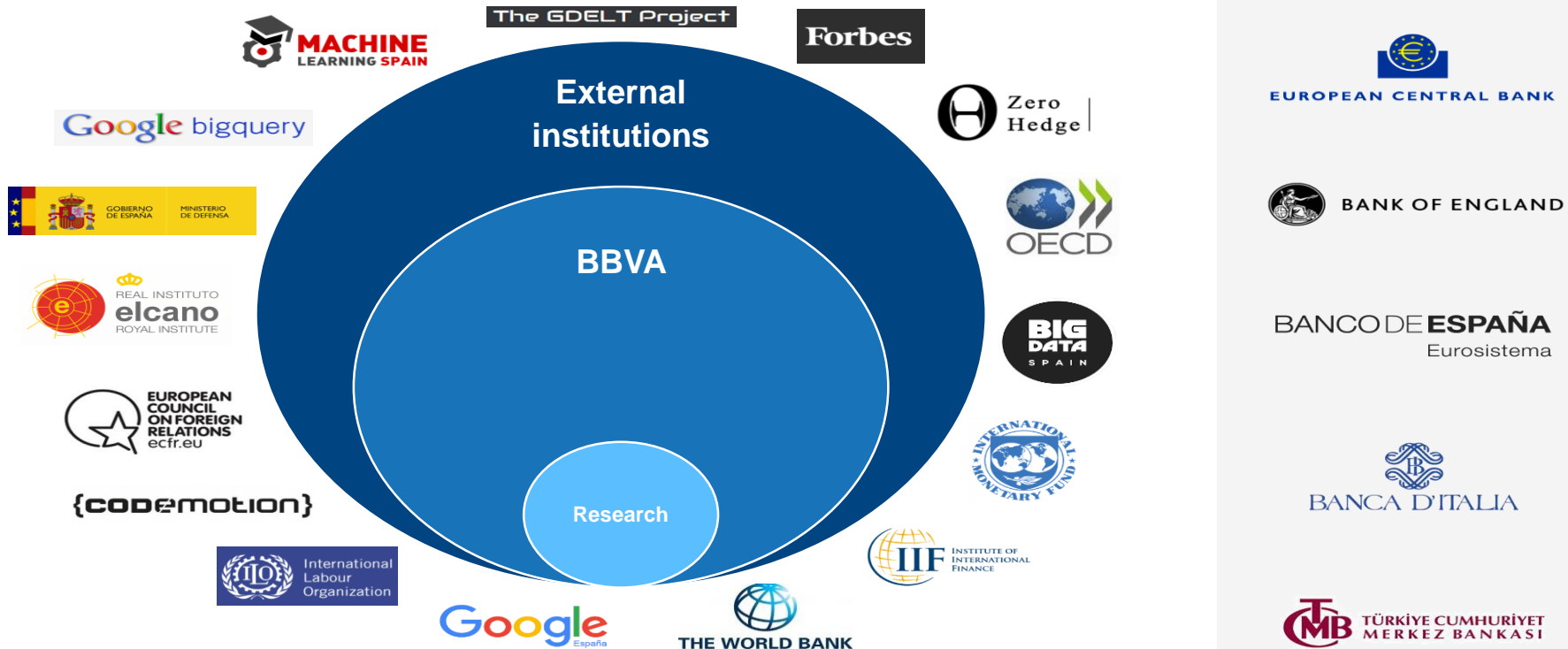
# Text as a key source of information to enrich analysis

**80**% of the total amount of web pages on the internet is given by **textual or unstructured data**

## Text mining

to extract meaning from strings of letters

The potential use of textual information and text sources **improves the understanding of the social, economic and financial systems**

It helps us to **understand what drives monetary policy** decisions

20%
of used data

80%
of available data

# What is GDELT?

## **G**lobal **D**atabase on **E**vents **L**ocation and **T**one

Open database of human society from every corner of the globe dating back to 1979 …

… including over 300 events around the world and more than 30000 themes…

… georeferenced across the entire planet…

…and collecting emotions using some of the most sophisticated algorithms

Going further into the database

# GDELT database and sentiment analysis

$$\text{Average tone} = \frac{\sum Positive\ words\ -\ \sum Negative\ words}{\sum Total\ words}$$
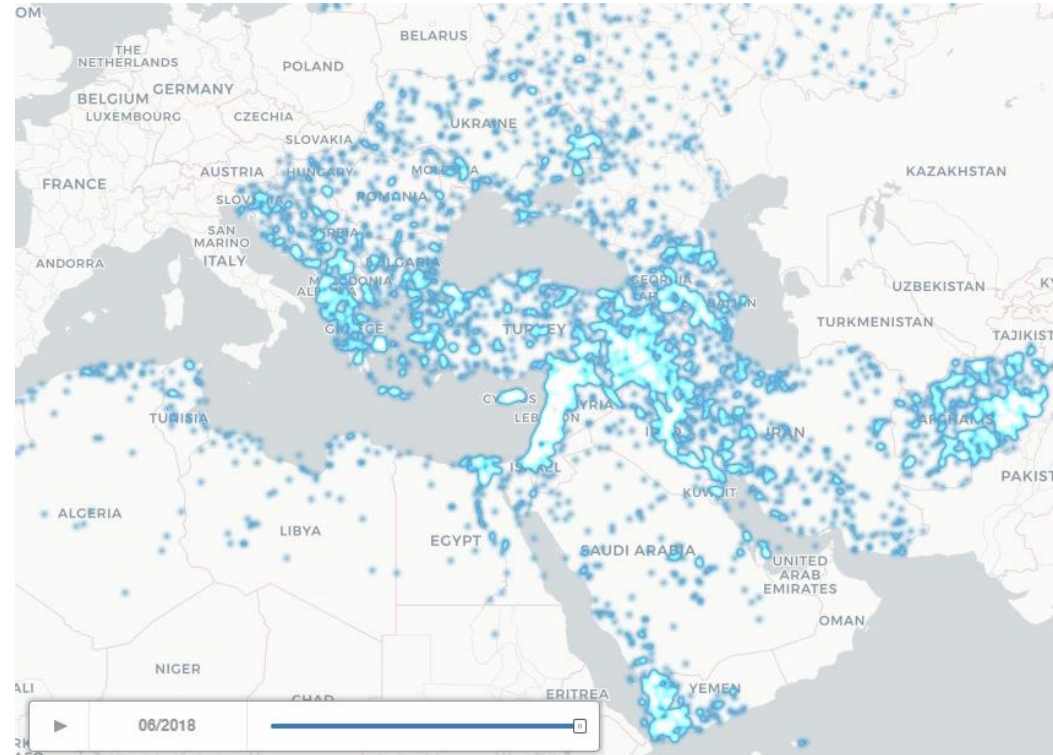
**+**

## Global Content Analysis Measures (GCAM)

**40** different **sentiment dictionaries**
to capture
more than **2300 emotions and dimensions**
with a value score

# Tracking Geopolitics on real time
## Is useful to identify the main hot spots and potential spillovers

**CONFLICT INTENSITY MAP  2017-18**
(NUMBER OF CONFLICTS/ TOTAL EVENTS)



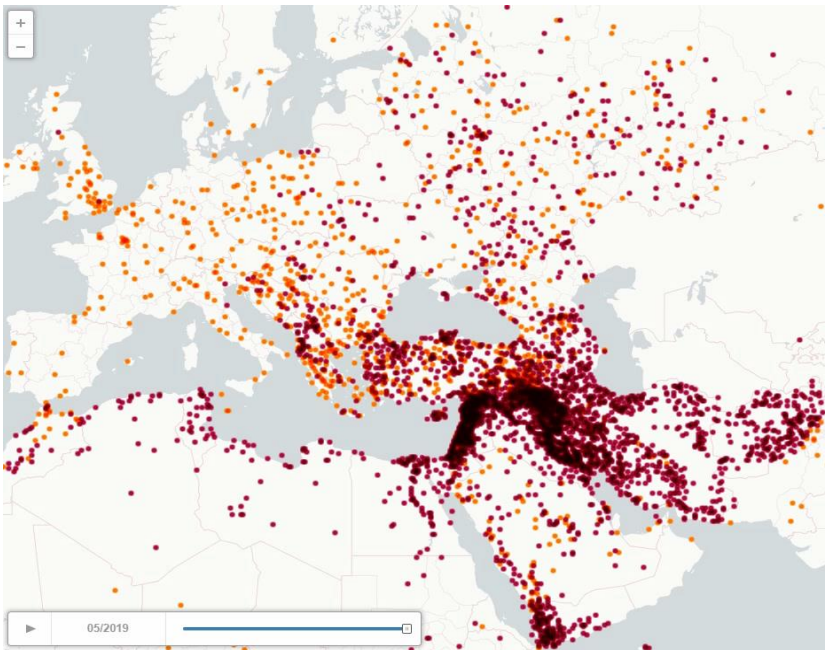More detailed information can be found here
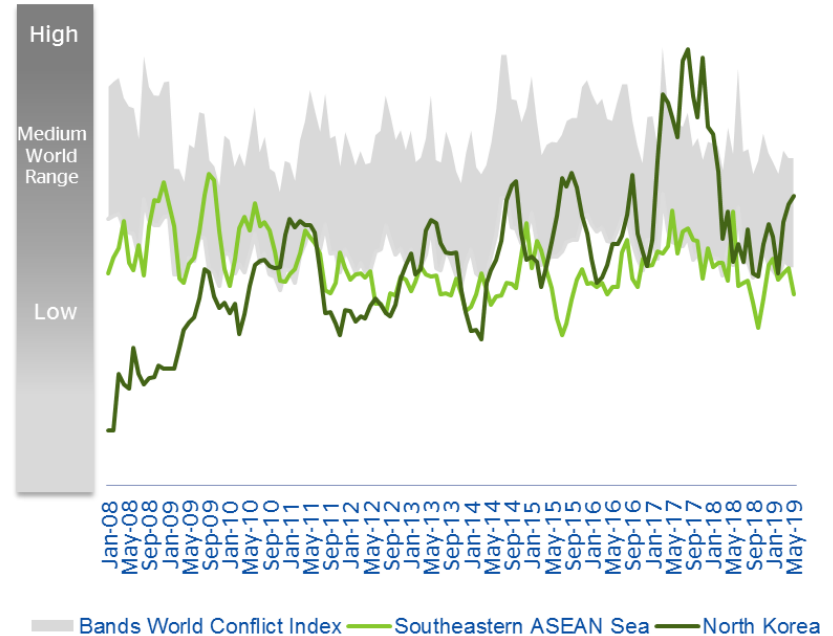
Source: www.gdelt.org & BBVA Research

# From an historical perspective…

### BBVA RESEARCH WORLD CONFLICT INTENSITY INDEX 1979-2018

### WORLD PROTEST INTENSITY MAP 1979- 2017



Source: www.gdelt.org & BBVA Research

# …to the current "hot" spots…

## BBVA RESEARCH REFUGEES FLOWS MAP IN 2015-19
NUMBER OF MEDIA CITATIONS ABOUT REFUGEES' INFLOWS AND OUTFLOWS



05/2019

Source: www.gdelt.org & BBVA Research

## BBVA RESEARCH ASIA CONFLICT INTENSITY INDEX 2008-19



High

Medium World Range

Low

Jan-08 May-08 Sep-08 May-09 Sep-09 May-10 Sep-10 May-11 Sep-11 Jan-12 Sep-12 Jan-13 Sep-13 Jan-14 May-14 May-15 May-15 Jan-16 May-16 Jan-17 May-17 Sep-17 Jan-18 Sep-18 Jan-19 May-19

Bands World Conflict Index    Southeastern ASEAN Sea    North Korea

# … in "high definition"

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



Source: www.gdelt.org & BBVA Research
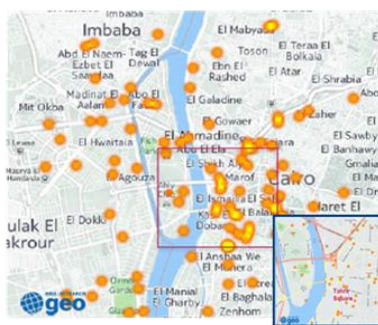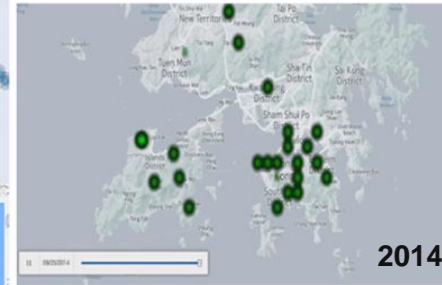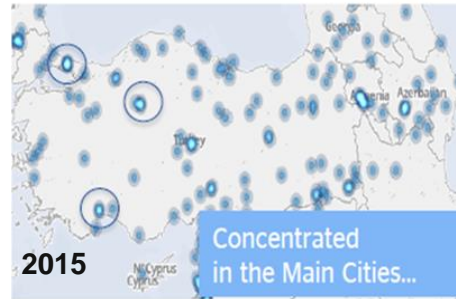
# Exploring spillover effects: from the China's links with the rest of the world…

## CHINESE SLOWDOWN: THE COUNTRY NETWORK AND THE MEDIA PERCEPTION
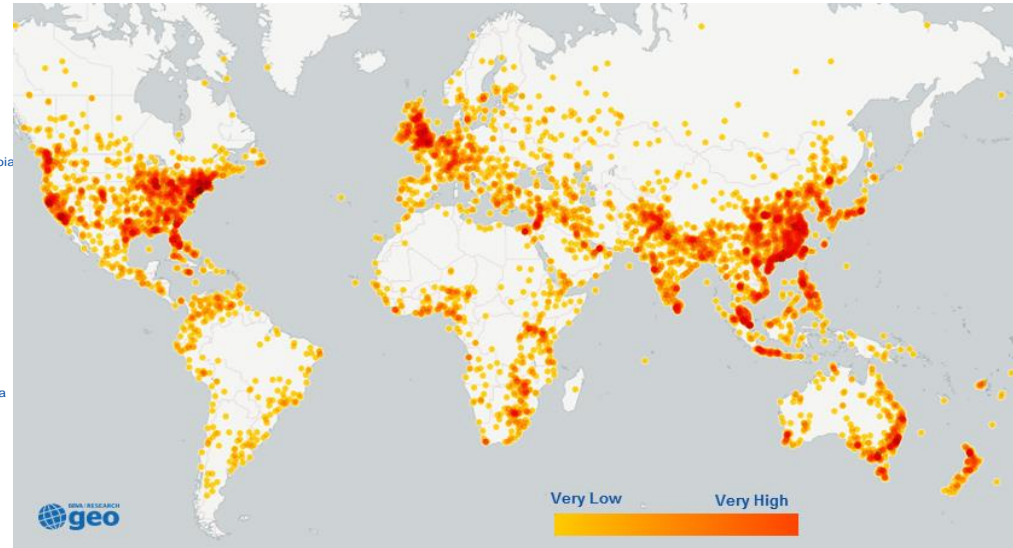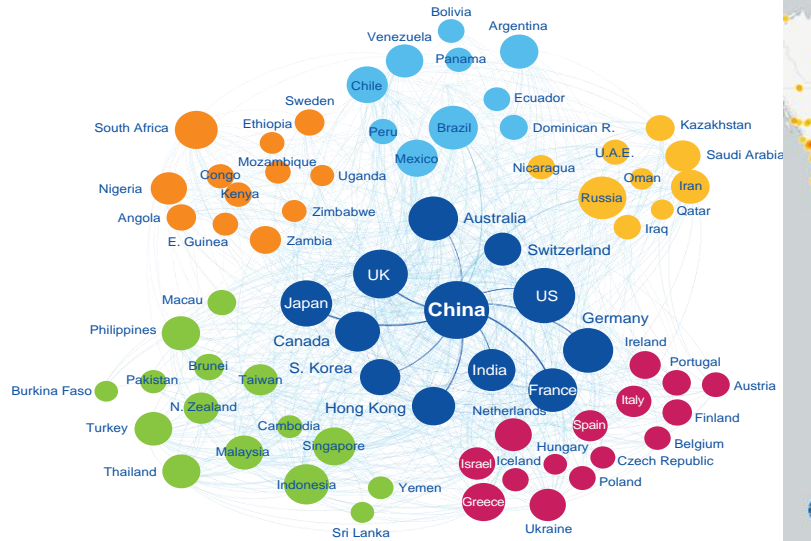


- ● Financial & Trade lines impaired (Main trade partners)
- ● Impact on commodity Exporters from Latam
- ● Conversation on Oil demand
- ● Activity Impact in Europe
- ● Activity Impact in EM Asia
- ● Impact on commodity Exporters from Africa

Source: www.gdelt.org & BBVA Research

# … to the world trade evolution from an historical perspective

## BBVA RESEARCH WORLD TRADE SUPPORT INDEX
(TONE & COVERAGE VERBAL COOPERATION AT WTO)



Verbal Cooperation (3 months mov.avg)
Material Cooperation (3 months mov.avg)

## BBVA RESEARCH TRADE SUPPORT INDEX CHANGES 2008-17
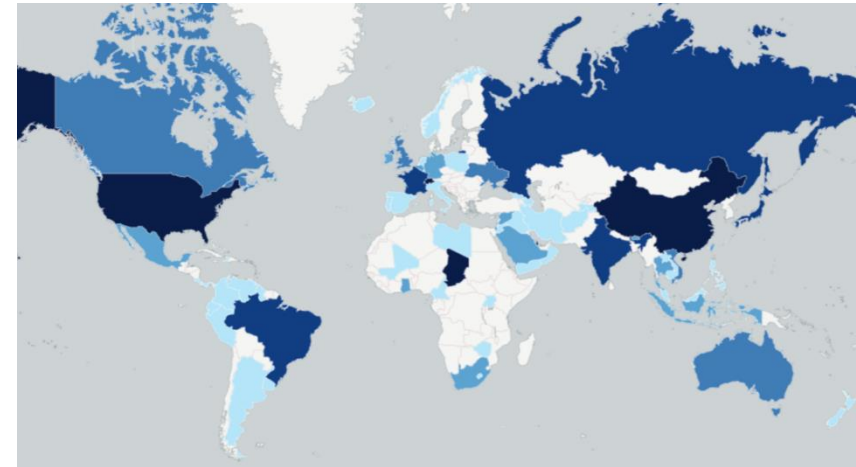


### THE GLOBAL INCIDENCE OF PROTECTIONISM 2008-2015
(GLOBAL TRADE ALERT)
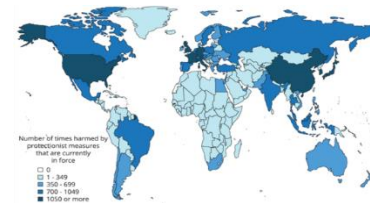
Source: www.gdelt.org & BBVA Research

More detailed information can be found here

# The FinTech Universe through the Eyes of the Media and Social Networks

**MEDIA SENTIMENT EVOLUTION OF FINTECH ALL OVER THE WORLD 2015-2018**
(RED COLOURS INDICATE NEGATIVE TONE, GREEN COLOURS POSITIVE TONE AND YELLOW ONES, NEUTRAL)

**MOST USED TERMS IN THE FINTECH UNIVERSE IN TWITTER**





More detailed information can be found here

# Internal databases: working with aggregated and anonymized BBVA Data to have indicators in real time and high definition
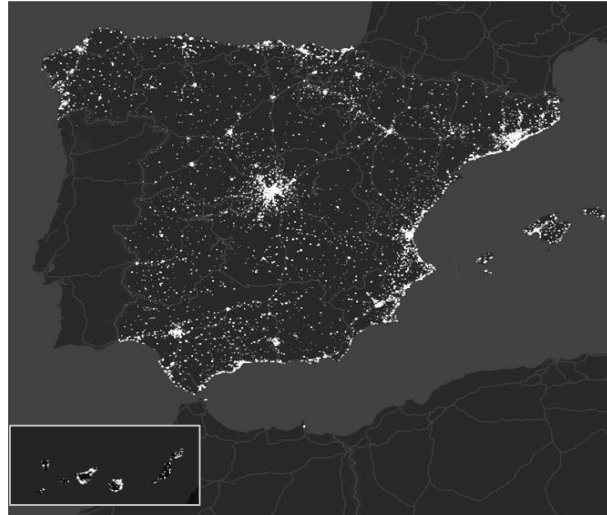


710M card transactions from 1M PoS, made by 53M people, representing €43.000M

4 billion card transactions made by 14M people

1.500M card transactions from 1,1M PoS, made by 88M people, representing €41.000M

**+ US, Colombia, Perú, Argentina**

# Positive growth in USA, Turkey, Mexico, Peru and Argentina. Spain registered negative growth rates at the end of August, while Colombia continued stagnant

## BBVA RESEARCH BIG DATA CONSUMPTION INDICES
(Total consumption by card, % YoY, 7D cumulative)



**Spain** — **Turkey** — **Mexico** — **USA\*** — **Colombia** — **Peru** — **Argentina (right)**
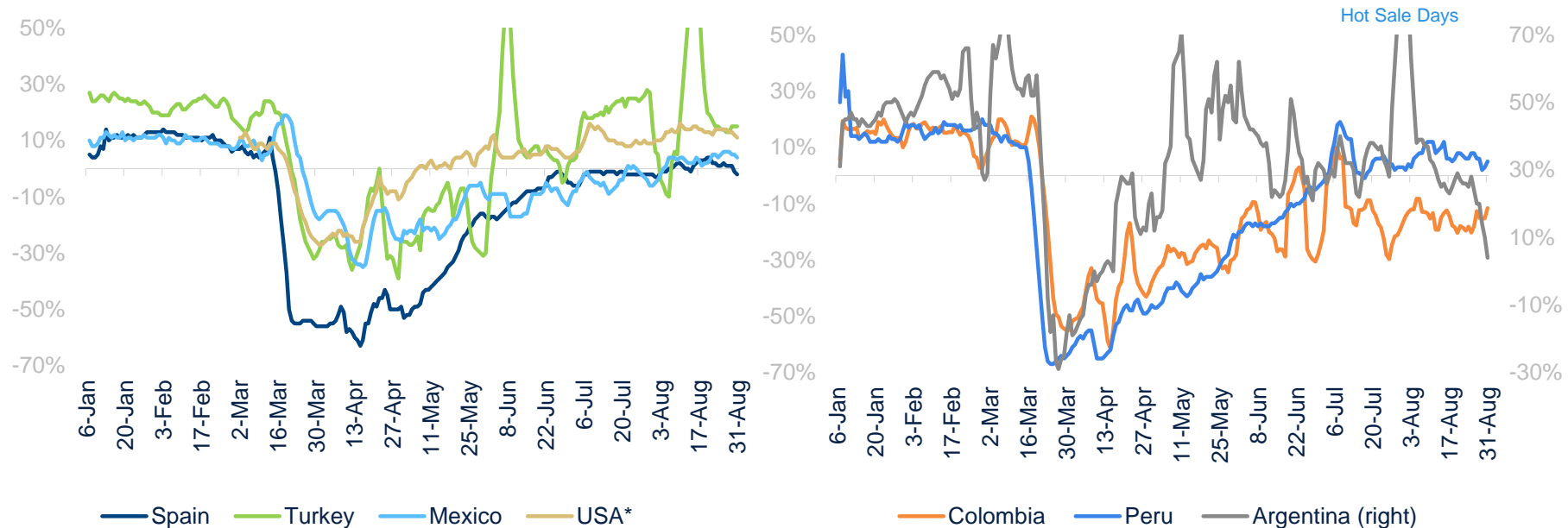
Interactive charts with weekly updated information can be found <u>here</u> and more detailed information can be found <u>here</u>

Source: BBVA Research. * Proxied by USA 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. More detail about the data could be found in the following <u>link</u>.

# Consumption Diffusion Heat Map: August led some recovery by sector of activity in Peru, USA, Turkey and Spain

**BBVA RESEARCH BIG DATA CONSUMPTION HEAT MAP** (YoY, 7D cumulative)



|  | MAR | APR | MAY | JUN | JUL | AUG |
|---|---|---|---|---|---|---|
| **Total consumption** | | | | | | |
| Spain | | | | | | |
| Turkey | | | | | | |
| Mexico | | | | | | |
| Colombia | | | | | | |
| Peru | | | | | | |
| Argentina | | | | | | |
| USA* | | | | | | |

Lower Growth — Higher Growth
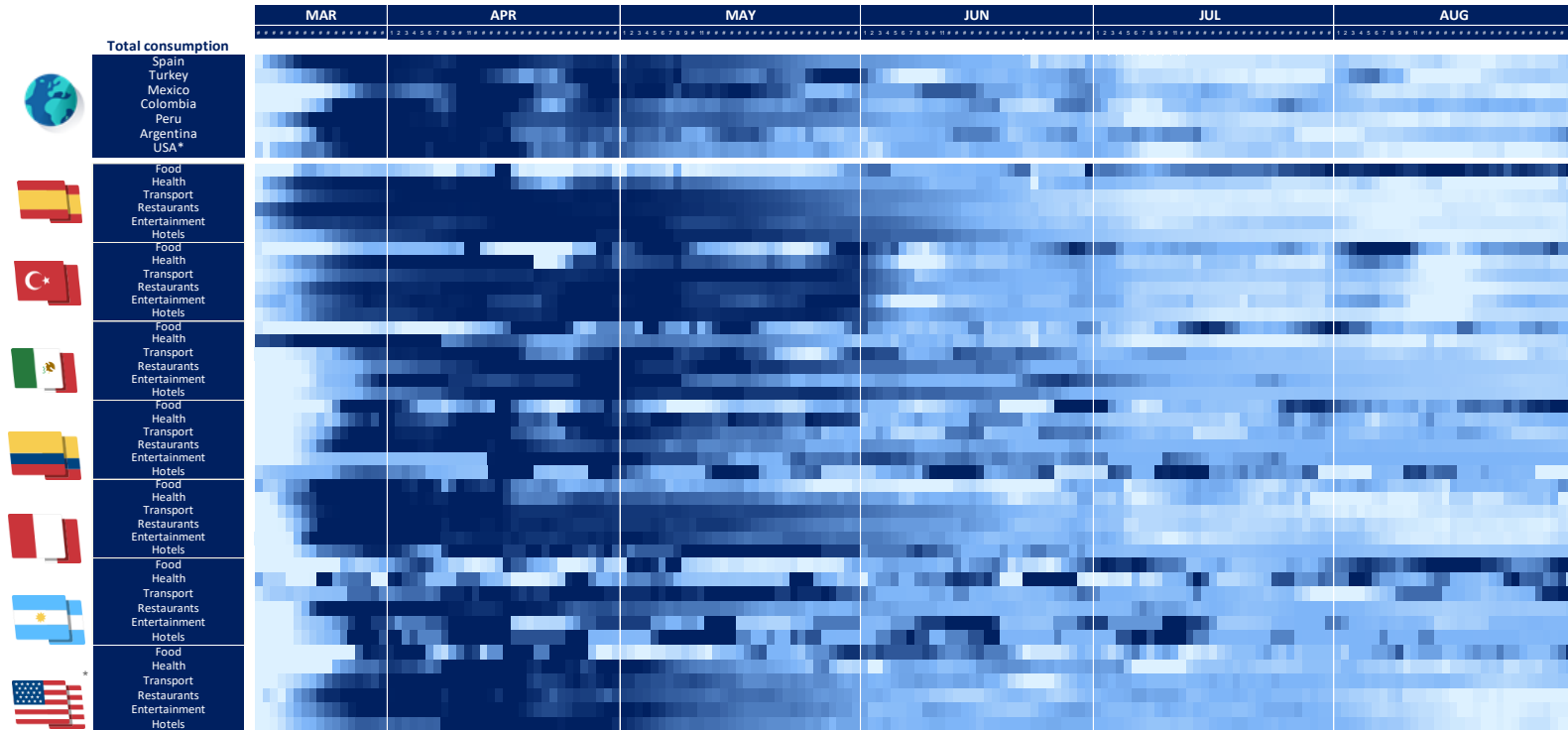
*Proxied by USA Sunbelt. Source: BBVA Research.

Series are coloured according to the history of each series since March 20 until the end of the series. The darker blue represents the lower 90% percentile and the lighter one represents the upper 10% percentile.

# Technology matters: the gap between e-commerce growth and physical purchases remained. Online sales are increasing in all the countries, expect in Spain given the importance of travel expenditures

**BBVA RESEARCH BIG DATA CONSUMPTION INDICES: PHYSICAL VS ONLINE**
(Total consumption by Point of Sales. % YoY, 7D cumulative. Pie chart reflects the proportion of online and physical purchases in 2019)
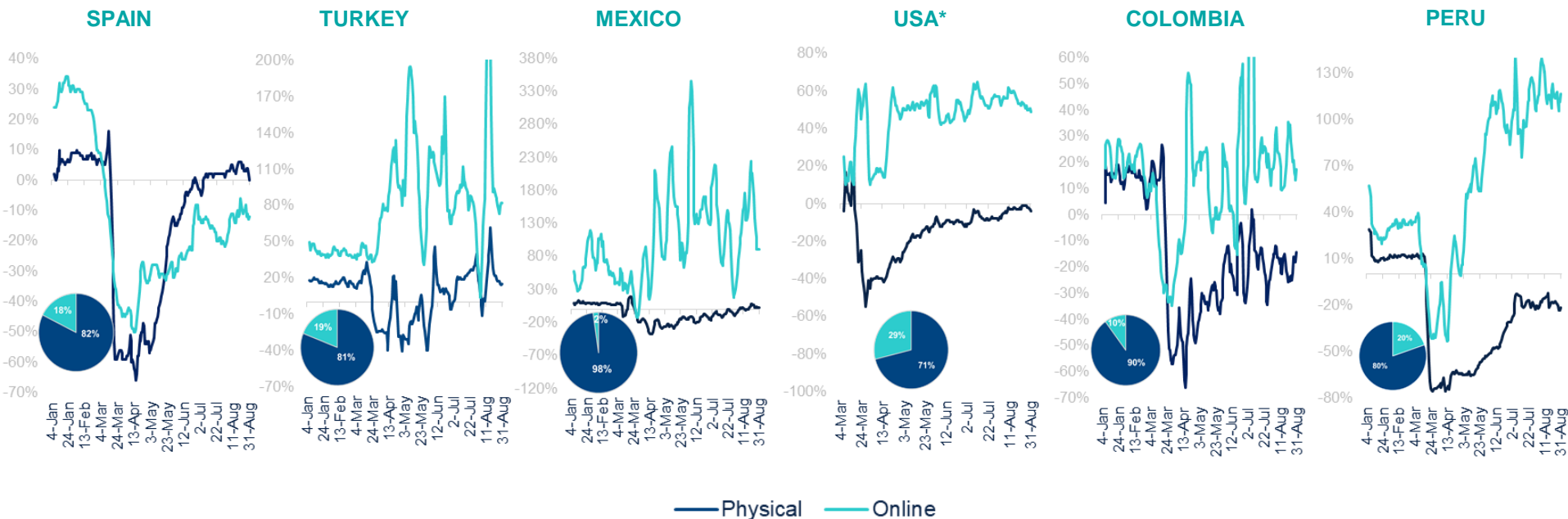


Source: BBVA Research. * USA = Proxied by USA Sunbelt.

# No big differences in card consumption patterns during August compared with July

**BBVA RESEARCH BIG DATA CONSUMPTION INDICES: TOTAL CONSUMPTION EVOLUTION IN AUGUST AND SECTORIAL DISTRIBUTION\* DURING THE MONTH** (Bar chart: % YoY by week and month. Pie chart: % YoY by sector in August)



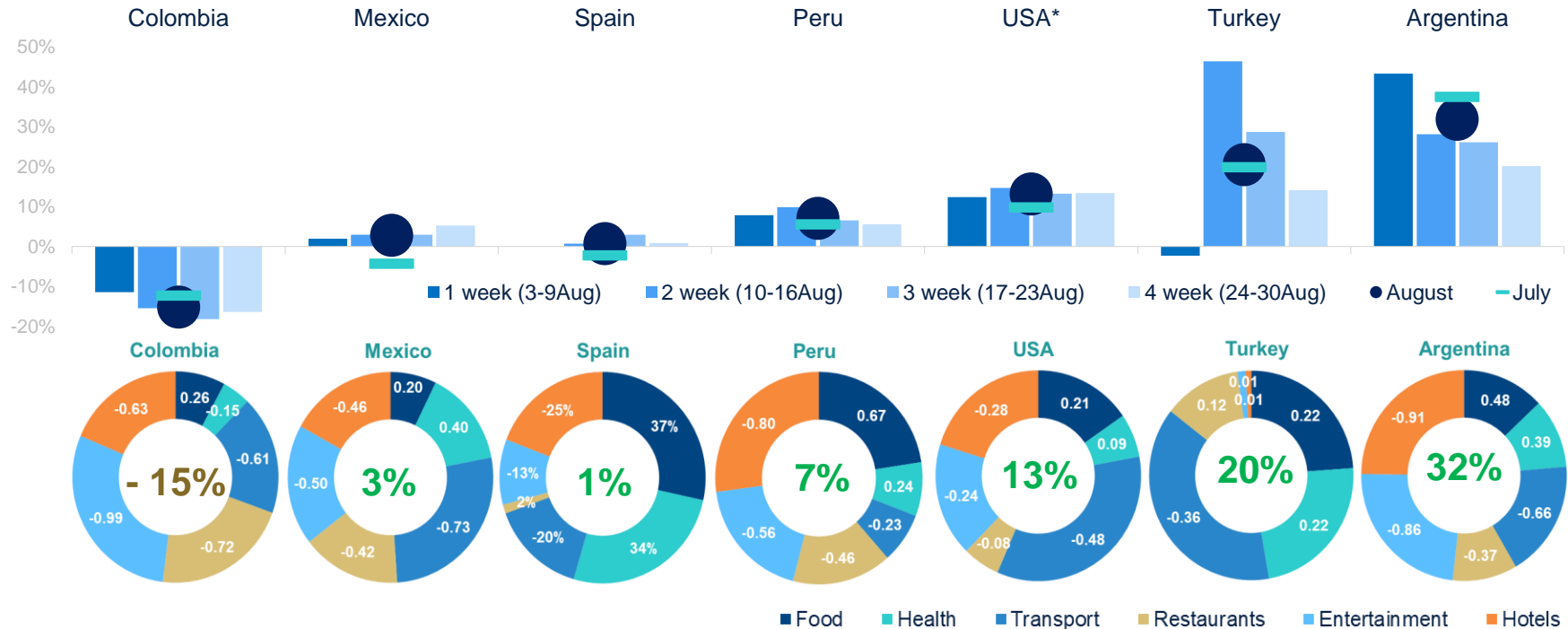Legend (bar chart): 1 week (3-9Aug) · 2 week (10-16Aug) · 3 week (17-23Aug) · 4 week (24-30Aug) · ● August · — July

**Colombia:** -15%
- Food -0.63, 0.26, -0.15, -0.61
- -0.99, -0.72

**Mexico:** 3%
- -0.46, 0.20, 0.40
- -0.50, -0.42, -0.73

**Spain:** 1%
- -25%, -13%, 2%, 37%
- -20%, 34%

**Peru:** 7%
- -0.80, 0.67, 0.24, -0.23
- -0.56, -0.46

**USA:** 13%
- -0.28, 0.21, 0.09
- -0.24, -0.08, -0.48

**Turkey:** 20%
- 0.01, 0.01, 0.12, 0.22
- -0.36, 0.22

**Argentina:** 32%
- -0.91, 0.48, 0.39
- -0.86, -0.37, -0.66

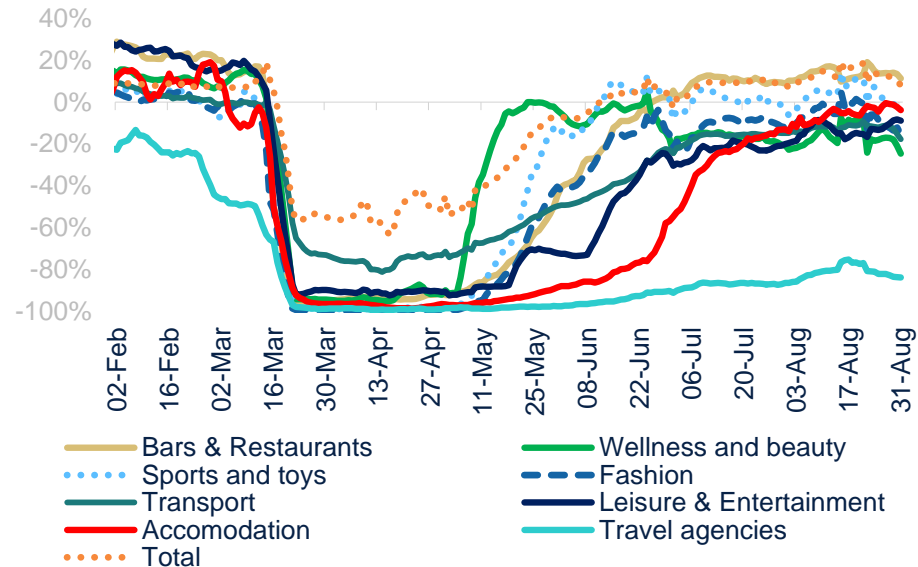Legend (pie chart): ■ Food ■ Health ■ Transport ■ Restaurants ■ Entertainment ■ Hotels

\*It should be take into account there are other categories of consumption not considered in the pie chart, so the total consumption is not equal to the average of the components of the chart
Source: BBVA Research. \* USA = Proxied by USA Sunbelt

# The Spanish recovery slowed among national consumers. Foreign expenditure is still stagnant

**SPAIN: AVERAGE EXPENDITURE BY SECTOR**
(% YoY, 7D cumulative)



NATIONAL

FOREIGN

Bars & Restaurants
Sports and toys
Transport
Accomodation
Total
Wellness and beauty
Fashion
Leisure & Entertainment
Travel agencies

Bars & Restaurants
Transport
Food
Total
Fashion
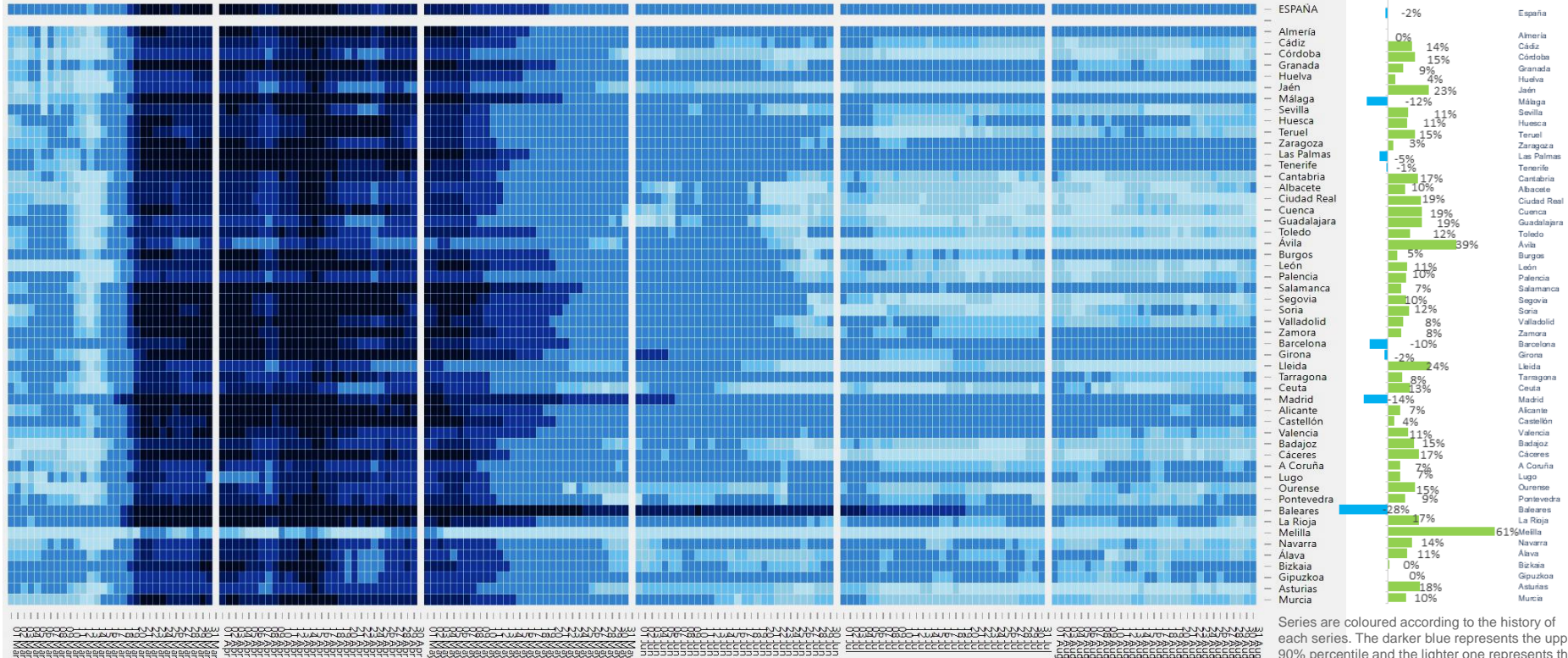Accomodation
Other services

More detailed information can be found here

Source: BBVA Research
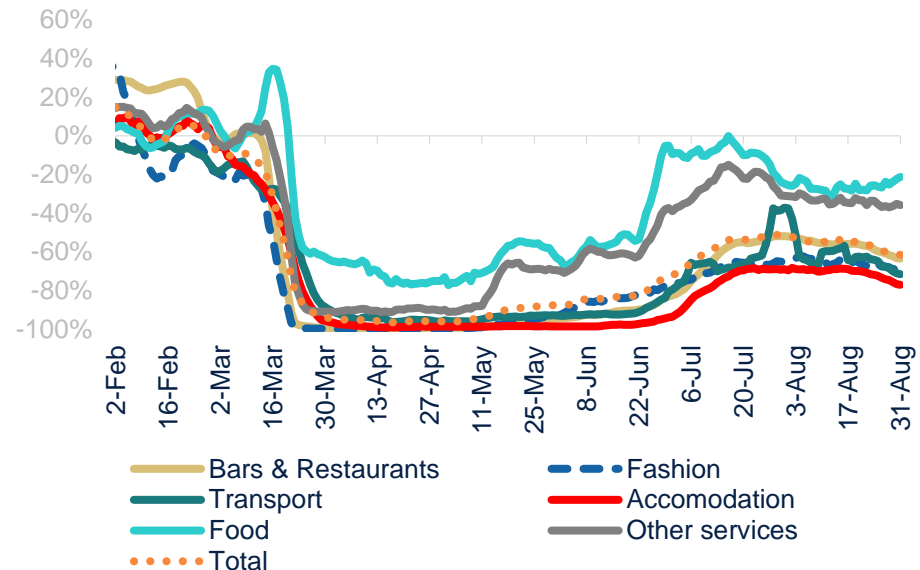
# 43/50 Spanish provinces are growing despite the stagnation of foreign expenditure

**SPAIN: TOTAL AVERAGE WEEKLY EXPENDITURE BY PROVINCE***
(Considering national and foreign cards and physical and online purchases . % YoY, 7D cumulative)

**25-31 AUG EXPENDITURE** (% YOY)



| Province | % YoY |
|---|---|
| España | -2% |
| Almería | 0% |
| Cádiz | 14% |
| Córdoba | 15% |
| Granada | 9% |
| Huelva | 4% |
| Jaén | 23% |
| Málaga | -12% |
| Sevilla | 11% |
| Huesca | 11% |
| Teruel | 15% |
| Zaragoza | 3% |
| Las Palmas | -5% |
| Tenerife | -1% |
| Cantabria | 17% |
| Albacete | 10% |
| Ciudad Real | 19% |
| Cuenca | 19% |
| Guadalajara | 19% |
| Toledo | 12% |
| Ávila | 39% |
| Burgos | 5% |
| León | 11% |
| Palencia | 10% |
| Salamanca | 7% |
| Segovia | 10% |
| Soria | 12% |
| Valladolid | 8% |
| Zamora | 8% |
| Barcelona | -10% |
| Girona | -2% |
| Lleida | 24% |
| Tarragona | 8% |
| Ceuta | 13% |
| Madrid | -14% |
| Alicante | 7% |
| Castellón | 4% |
| Valencia | 11% |
| Badajoz | 15% |
| Cáceres | 17% |
| A Coruña | 7% |
| Lugo | 7% |
| Ourense | 15% |
| Pontevedra | 9% |
| Baleares | 28% |
| La Rioja | 17% |
| Melilla | 61% |
| Navarra | 14% |
| Álava | 11% |
| Bizkaia | 0% |
| Gipuzkoa | 0% |
| Asturias | 18% |
| Murcia | 10% |

Series are coloured according to the history of each series. The darker blue represents the upper 90% percentile and the lighter one represents the lower 10% percentile. Expenditure by BBVA client and other Spanish cards registered in TPV BBVA

Source: BBVA Research

Lower Growth          Higher Growth

# The BBVA Big Data for investment shows a high correlation with the main investment aggregates, but giving us some anticipation to have the more updated view of the Business Cycle…

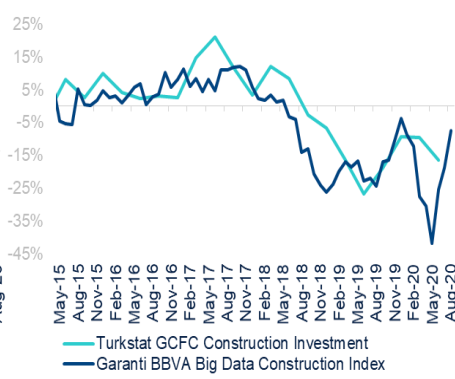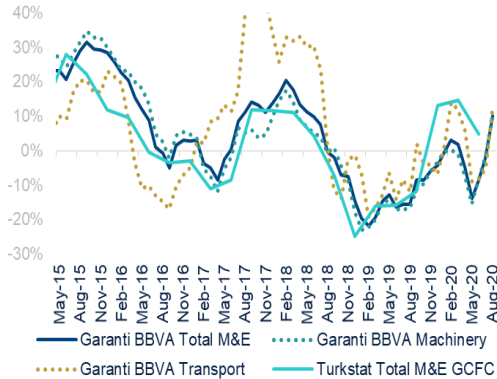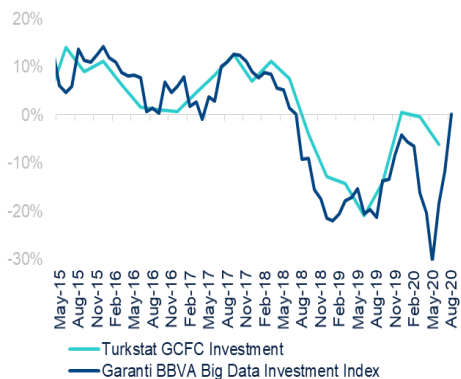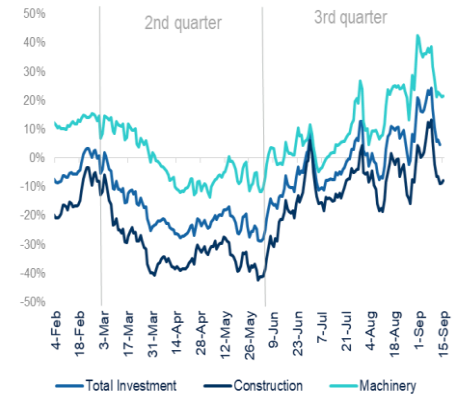**TURKEY: GBBBVA BIG DATA INVESTMENT INDICES**
(28-day cum. YoY real )

**TURKEY: BBVA BIG DATA INVESTMENT**
(28-day cum. YoY nominal in Invest.)
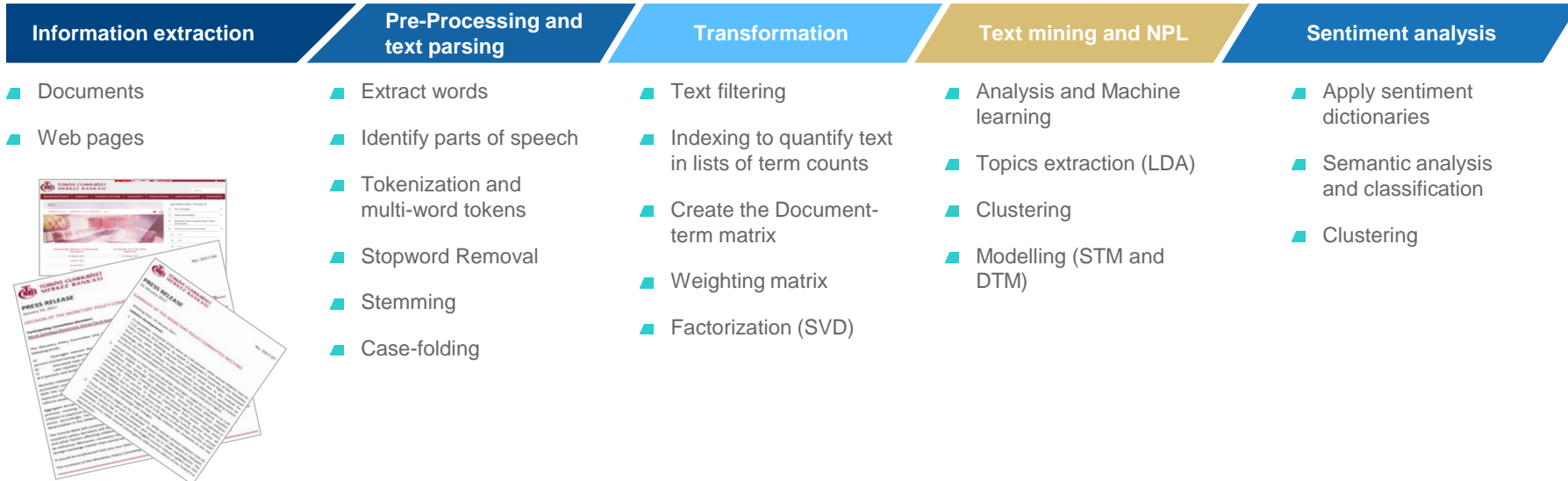
*Total Investment*

*Maquinery & Equipment*

*Construction*



The "High Definition" sectorial information can help us to track the evolution of sectors in detail and differentiate different shocks and to analyze regional economics to see how the provinces react to such shocks

More detailed information can be found here

Source: BBVA Research

# External databases: web scrapping and NPL techniques

Text mining makes information extraction from huge volumes of data easier and structures the information as important facts, key terms or persons.

| Information extraction | Pre-Processing and text parsing | Transformation | Text mining and NPL | Sentiment analysis |
|---|---|---|---|---|
| Documents | Extract words | Text filtering | Analysis and Machine learning | Apply sentiment dictionaries |
| Web pages | Identify parts of speech | Indexing to quantify text in lists of term counts | Topics extraction (LDA) | Semantic analysis and classification |
| | Tokenization and multi-word tokens | Create the Document-term matrix | Clustering | Clustering |
| | Stopword Removal | Weighting matrix | Modelling (STM and DTM) | |
| | Stemming | Factorization (SVD) | | |
| | Case-folding | | | |

(More information can be found in the annex

# Analyzing the Central Bank of Turkey, European Central Bank and Federal Reserve monetary reports to understand monetary policy decisions



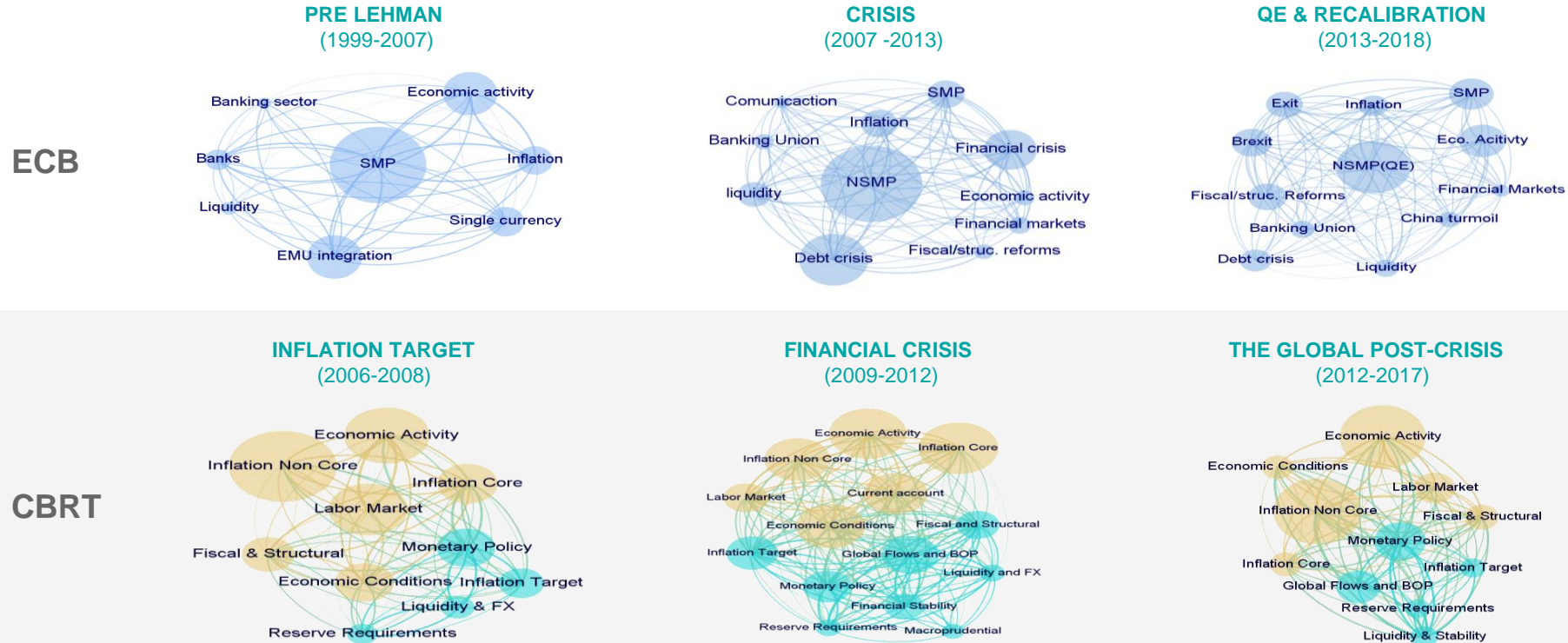More detailed information can be found here

# Understanding the Interconnectedness between Developed and Emerging Central Banks

## MONETARY POLICY IN DEVELOPED ECONOMIES AND RESPONSE IN THE EMERGING MARKETS (NETWORKS)

# You can find us at:

www.bbvaresearch.com

BBVA Research
Tomasa Rodrigo

@BBVAResearch
@TomasaRodrigo

# Thank you!