

# Investment in “Real Time” and “High Definition”: A Big Data Approach

Ali.B. Barlas, Seda Güler, Alvaro Ortiz & Tomasa Rodrigo

**Modelling with Big Data & Machine Learning: Measuring Economic Instability**  
The Bank of England, The Federal Reserve Board and King's College London

November 2020

## Why this paper is relevant

- **The first “Real Time” Investment indicator** from a Bank’s Big Data
- **Validation** with national accounts and proxies is **successful**
- **Improves** the properties of a **Standard Nowcasting Model**
- It’s not only “Real Time” **but also “High Definition” (Sector & Geography)**
- Validation results are **promising for other Countries**

# The Covid-19 crisis has reinforced the role of Big Data tools for Economic Analysis

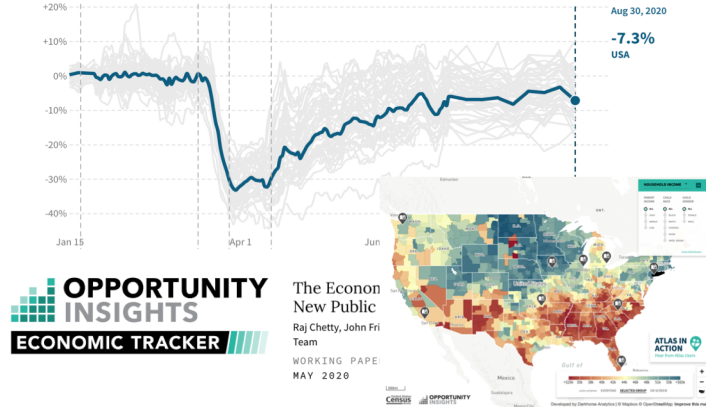
The high uncertainty triggered by the Covid-19 crisis has stressed the need to monitor the evolution of the economy in “real time”. These efforts have been materialized in several ways:

- **Focusing on timely, alternative indicators:** soft data surveys (particularly the Purchasing Manager Indexes, PMIs) and other high frequency indicators like electricity production or chain store sales released on daily or weekly basis.
- **Developing higher frequency models:** Some CBs have relied on this High Frequency indicators to develop weekly activity tracker models such as the FED'WEI (Lewis, 2020) and the Bundesbank WAI (Eraslan, S. and T. Götz, 2020).
- **Developing New Big Data Indicators\*:** Focusing on daily aggregate information of banking transactions to track consumption, employment , turnover, mobility... ... [link](#).

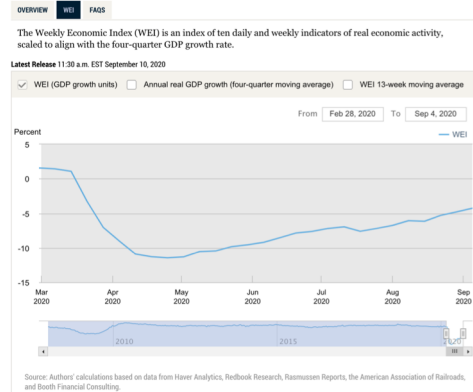
\* Some of the Recent literature on Big Data analysis Andersen, Hansen, Johannesen, & Sheridan (2020a), Andersen, Hansen, Johannesen, & Sheridan (2020b), Alexander & Karger (2020), Baker, Farrokhnia, Meyer, Pagel, & Yannelis (2020a), Baker, Farrokhnia, Meyer, Pagel, & Yannelis (2020b), Bounie, Camara, & Galbraith (2020), Chetty, Friedman, Hendren, & Stepner (2020), Chronopoulos, Lukas, & Wilson (2020), Cox, Ganong, Noel, Vavra, Wong, Farrell, & Greig (2020), Surico, Kanzig, & Hacıoglu (2020).

# Efforts on Big Data Information have been accelerated due to the importance to know an updated state of the economy...

In the United States, as of August 30 2020, total spending by all consumers decreased by **7.3%** compared to January 2020.

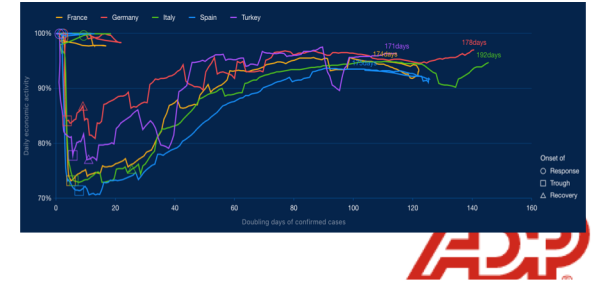


## Weekly Economic Index (WEI)



## TRACKING AND NAVIGATING THE PANDEMIC ECONOMY

Michael Spence



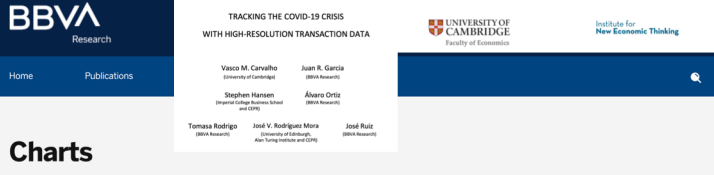
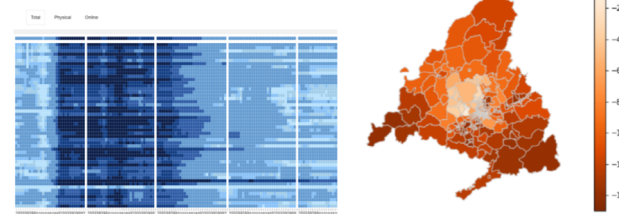
WORKING PAPER · NO. 2020-58

## The U.S. Labor Market During the Beginning of the Pandemic Recession

Tomaz Cajner, Leland D. Crane, Ryan A. Decker, John Grigsby, Adrian Hamins-Puertolas, Erik Hurst, Christopher Kurz, and Ahu Yildirmaz  
JULY 2020

Income per Capita

## SPANISH TOTAL CONSUMPTION BY PROVINCE



## Charts

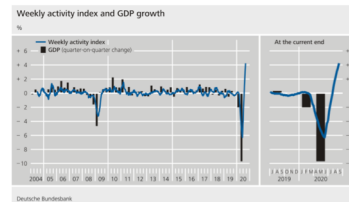
In this section you will find the base data used for



## Weekly activity index for the German economy

> DE

The weekly activity index (WAI) is an index designed to measure real economic activity in Germany in a timely manner. The index is based on daily, weekly, monthly and quarterly indicators for the German economy.



Updated as of 14 September 2020



# We continue to enhance our “Real Time & High Definition” extending our Big Data indexes to Investment (GFCF)

The investment spending is done mostly by companies and, to a lesser extent, by individuals

We track investment payments through



individual to firm transactions



firm to firm transactions

Firms are classified by their **NACE codes** to identify their **business activity** (in line with the European statistical classification of sectors)

We approximate investment demand in one type of asset taking into account the aggregate flows or transactions done from any firm or individual to the sector which produce the fixed assets

**Total Investment**

**Machinery Investment\***

**Construction Investment**

\*Machinery & Equipment, Media & ICT, Agriculture & Animals, Forestry, Durable Goods, Retail Trade, Textile & Clothing, Transportation and Shipping.

# The data: The Big Data Investment Data Representativeness

## INVESTMENT DATA 2019: BBVA vs COMPANY ACCOUNTS (CENTRAL BANK & TURKSTAT)

	BBVA Big Data			Turkey CBRT		
	Total	Machinery	Construction	Total	Machinery	Construction
<b>Transactions</b> (000s)	<b>24.6</b>	<b>22.3</b>	<b>2.3</b>			
<b>Amount</b> (US\$ bn)	<b>308</b>	<b>280</b>	<b>28</b>	<b>440</b>	<b>257</b>	<b>183</b>
<b>Firms</b> (000s)	<b>179.7</b>	<b>156.5</b>	<b>23.2</b>	<b>730.2</b>	<b>614.4</b>	<b>115.8</b>
<b>Firms</b> (% CBRT)	<b>24.6%</b>	<b>25.5%</b>	<b>19.8%</b>			

# A caveat: we develop Real Time & High Definition indicators ... we do not fully replicate the national accounts

## National Accounts

## Indicators

**Gross Fixed Capital Investments**  
(ESA, Yearly)

**Gross Fixed Capital Investments**  
(QNA, *Quarterly*)

**Net Disposal of Fixed assets**

Estimations Supply, Demand, Suveys...

**Balanced**  
( $P=E=I$ , One GDP figure Real Vs N)

**Balanced**  
( $P=E=I$ , One GDP figure Real Vs N, SA and NSA)

**Stability**  
(short samples)

**Frequent Revisions**  
(short samples)

**Sample Representativity**

**Seasonality**

**Investment Proxies**  
(Proxies, Monthly)

**Big Data Index**  
(Proxies, Daily)

Gross Fixed Capital Formation – <i>AN_FS</i> *	Sources/Indicators
Dwellings	<ul style="list-style-type: none"> <li>Building activity statistics (e.g. value/volume of work done by builders) relating to dwellings</li> <li>Capital outlays by purchasers of capital goods (improvements to dwellings, public construction)</li> <li>Number of units sold (brokers' commissions on sale of new dwellings)</li> <li>Index of construction output or turnover</li> <li>Number of building permits issued, with adjustments for delay/realization</li> <li>Production or sale of building products, such as concrete</li> <li>Labour inputs in physical terms and labour cost</li> </ul>
Other buildings and structures	<ul style="list-style-type: none"> <li>Building activity statistics (e.g. value/volume of work done by builders) relating to non-residential building</li> <li>Civil and other engineering construction activity statistics</li> <li>Capital outlays by purchasers of capital goods</li> <li>Index of construction output or turnover, other than dwellings</li> <li>Labour inputs in physical terms and labour cost</li> <li>Investment intentions</li> <li>Production or sale of building products, such as concrete</li> </ul>
Transport equipment	<ul style="list-style-type: none"> <li>Capital outlays by purchasers of capital goods</li> <li>Product (i.e. commodity) flow approach (using manufacturing output, export and import data by product)</li> <li>Estimated commercial share of dealers' sales, new motor vehicle registrations</li> </ul>
Other machinery and equipment, of which:	<ul style="list-style-type: none"> <li>Capital outlays by purchasers of capital goods (corporate sector)</li> <li>Product flow approach</li> <li>Average purchased by farms/unincorporated businesses multiplied by estimated number of farms/unincorporated businesses (machinery and equipment)</li> </ul>
Weapons systems	<ul style="list-style-type: none"> <li>Government finance statistics</li> </ul>
Cultivated assets	<ul style="list-style-type: none"> <li>Extension of annual models used to derive estimates of the production of cultivated assets</li> </ul>
Intellectual property products, of which:	<ul style="list-style-type: none"> <li>Value/volume of work done by capital goods producers</li> <li>Product flow approach</li> <li>Metres drilled (oil and gas exploration well drilling)</li> <li>Labour inputs in physical terms and labour cost</li> <li>Turnover from VAT statistics or business surveys (for computer software)</li> </ul>

\* Breakdown of fixed assets

**Cross Validation**



**Gross Fixed Capital Investments**  
(QNA, *Quarterly*)



**Investment Proxies**  
(Proxies, Monthly)

# Big Data Investment in Real time and High Definition: The case of Turkey

## Big Data Investment in detail: The case of Turkey

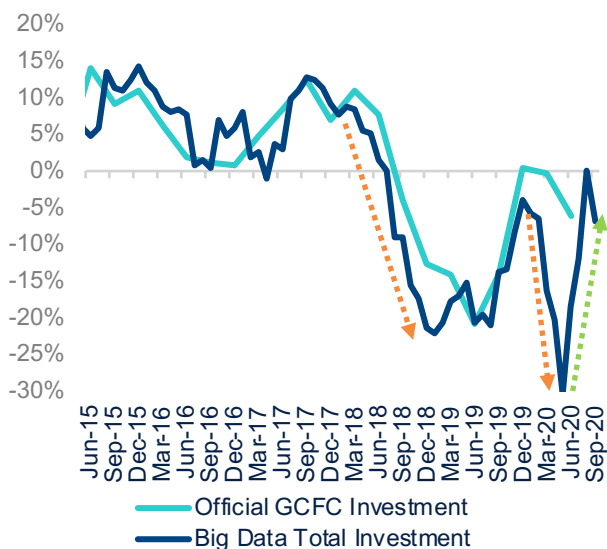
- The volatility of investment in Turkey and the delay in the publication of official data (standard in EMs) make it an ideal case for analysis
- We compute investment indicators in real time (daily) for 22 sectors and 81 provinces
- Validation results are positive in terms of the main Aggregates (Machinery & Transport and Construction) as well as the aggregate investment
- A deeper analysis show how the Big Data investment can improve the properties of an standard Nowcasting Model in terms of Accuracy, Anticipation and News contribution
- The high definition analysis show the characteristics of the latest investment shocks and the Geography response of Investment to the Covid-19 crisis

# Validation I: The Big Data investment index shows a high correlation an co-movement with the oficial data

## TURKEY: GBBBVA BIG DATA INVESTMENT INDICES

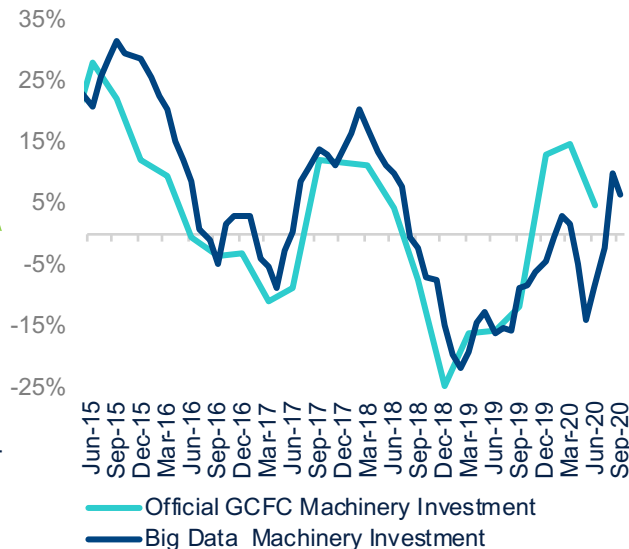
(28-day cum. YoY real )

### Total Investment



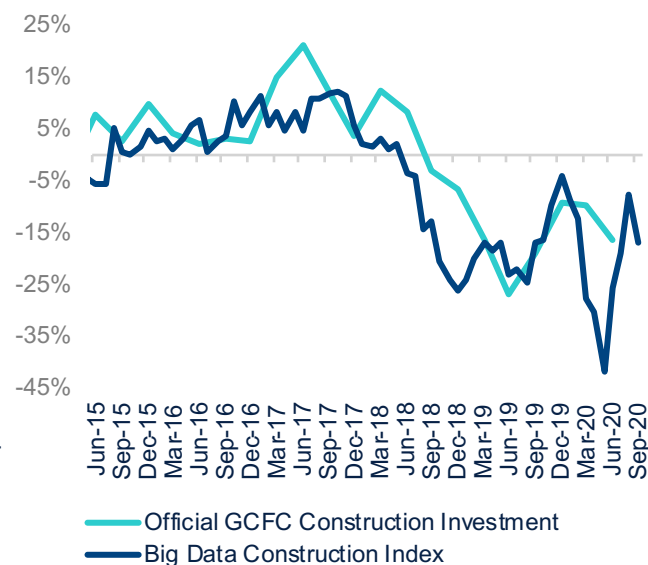
**Correlation coefficient: 0.88**

### Machinery & Equipment



**Correlation coefficient: 0.84**

### Construction



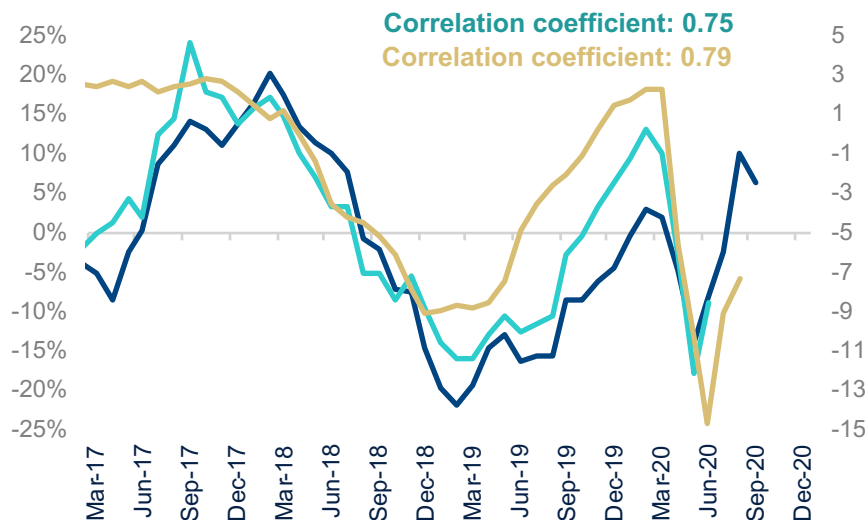
**Correlation coefficient: 0.77**

# Validation II: The sincrony of Big Data Investment with the Investment Cycle is validated by high correlation coefficients with HF proxies

## BBVA BIG DATA INVESTMENT & HIGH FREQUENCY PROXIES

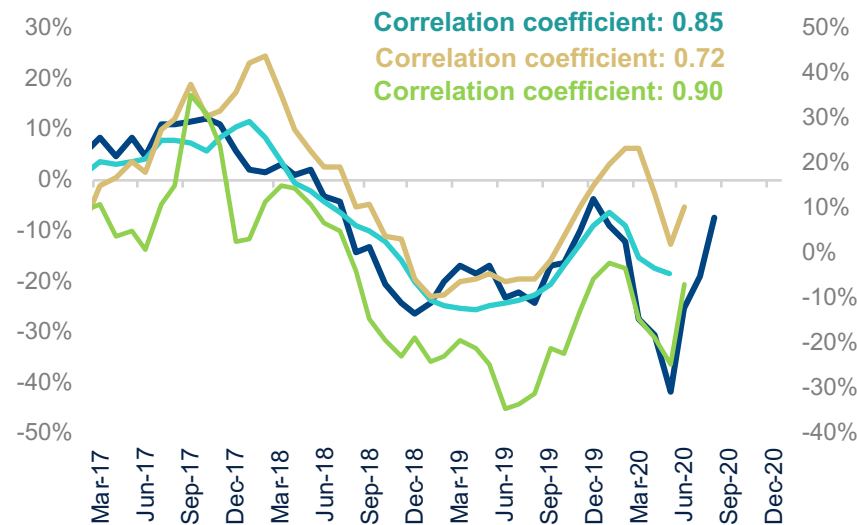
(28-day cum. YoY real )

### Maquinery & Equipment



- Garanti BBVA Big Data Total M&T Investment
- Industrial Production Machinery
- Capacity Utilization Rate Investment Goods - yoy, rhs

### Construction



- Garanti BBVA Big Data Construction Index 3m
- Employment in Construction
- Non-Metallic Mineral (Cement)
- Turnover Building - rhs

# Big Data & Nowcasting Model (DFM): The framework

## A Dynamic factor Model (DFM)

$$y_t = \Lambda f_t + \epsilon_t ,$$

$$f_t = A_1 f_{t-1} + A_2 f_{t-2} + \cdots + A_p f_{t-p} + u_t ,$$

$$u_t \sim \text{i.i.d. } \mathcal{N}(0, Q)$$

## Expectation Maximization (EM) Algorithm

$$L(\theta, \theta(j)) = \mathbb{E}_{\theta(j)} [l(Y, F; \theta) | \Omega_T] ;$$

$$\theta(j+1) = \arg \max_{\theta} L(\theta, \theta(j)) .$$

the conditional moments of the latent factors,  
 $\mathbb{E}_{\theta(j)} [f_t | \Omega_T]$ ,  $\mathbb{E}_{\theta(j)} [f_t f'_t | \Omega_T]$ ,  $\mathbb{E}_{\theta(j)} [f_{t-1} f'_{t-1} | \Omega_T]$   
 and  $\mathbb{E}_{\theta(j)} [f_t f'_{t-1} | \Omega_T]$

obtained through the Kalman smoother  
 for the state space representation:

$$y_t = \Lambda(j) f_t + \epsilon_t , \quad \epsilon_t \sim \text{i.i.d. } \mathcal{N}(0, R(j))$$

$$f_t = A(j) f_{t-1} + u_t , \quad u_t \sim \text{i.i.d. } \mathcal{N}(0, Q(j))$$

## Outcomes

*Nowcasting Accuracy*

*Nowcasting Anticipation*

*News Contribution*



# Big Data & Nowcasting Model (DFM): Variables & Releases

## TURKEY: VARIABLES IN MONTHLY GDP DFM

	2020								
	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sept
GDP									
Industrial Production									
Non-metal Mineral Production									
Auto Sales									
Number of Employed									
Number of Unemployed									
Auto Imports									
Auto Exports									
Electricity Production									
Manufacturing PMI									
Real Sector Confidence									
Total Loans growth 13-week									
Bog Data Total Consumption Indicator									
Bog Data Total Investment Indicator									

Hard Data (M & D)

Soft (M)

Fin (W)

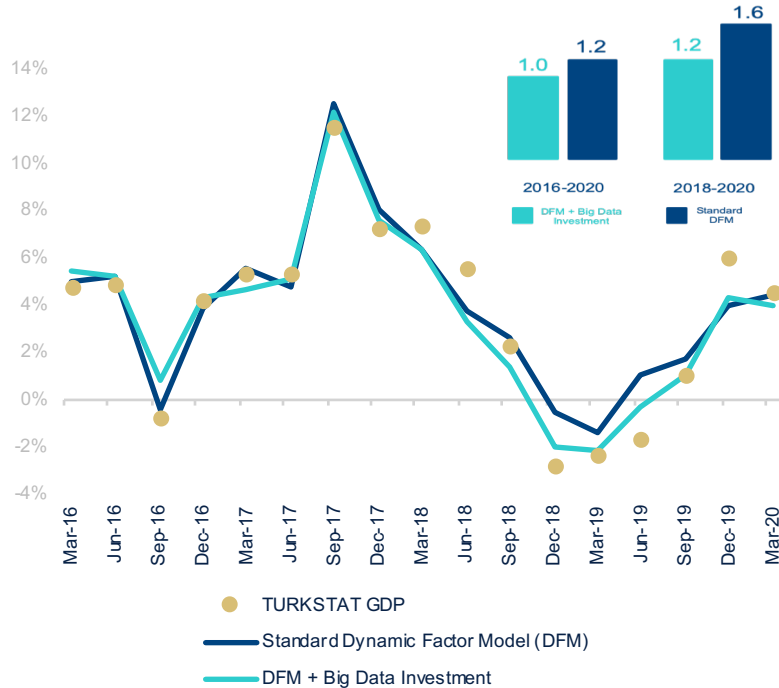
Big Data (D)

## TURKEY: VARIABLES IN MONTHLY INVESTMENT DFM

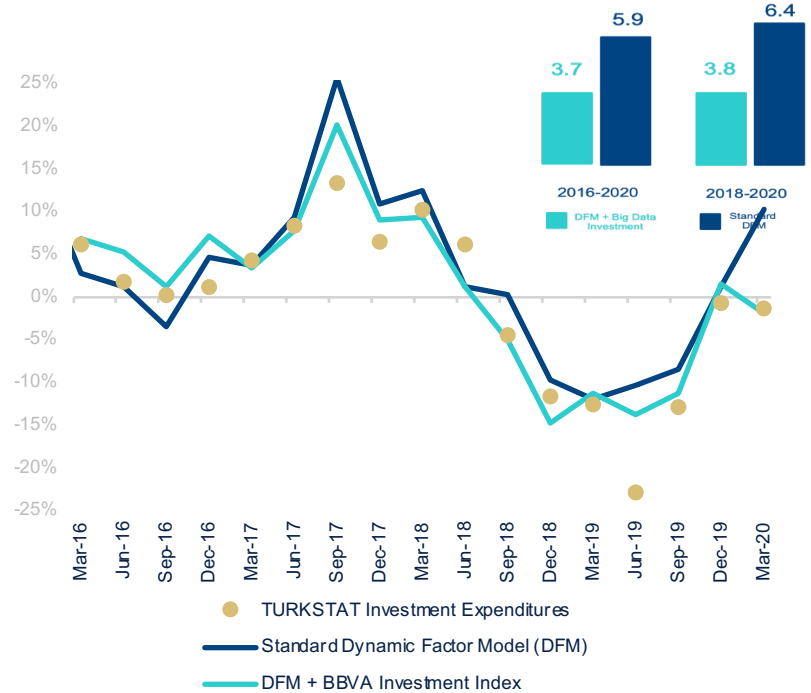
	2020								
	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep
GFCF									
Capital Goods Production									
Non-metal M. Production									
Capital Goods Imports									
Commercial Vehicle Sales									
Real Sector Confidence									
Corporate Lonas									
Big Data Investment									

# Big Data & Nowcasting Model (DFM): Out-of-Sample errors

## TURKEY: OUT-OF-SAMPLE ERROR GDP MODEL



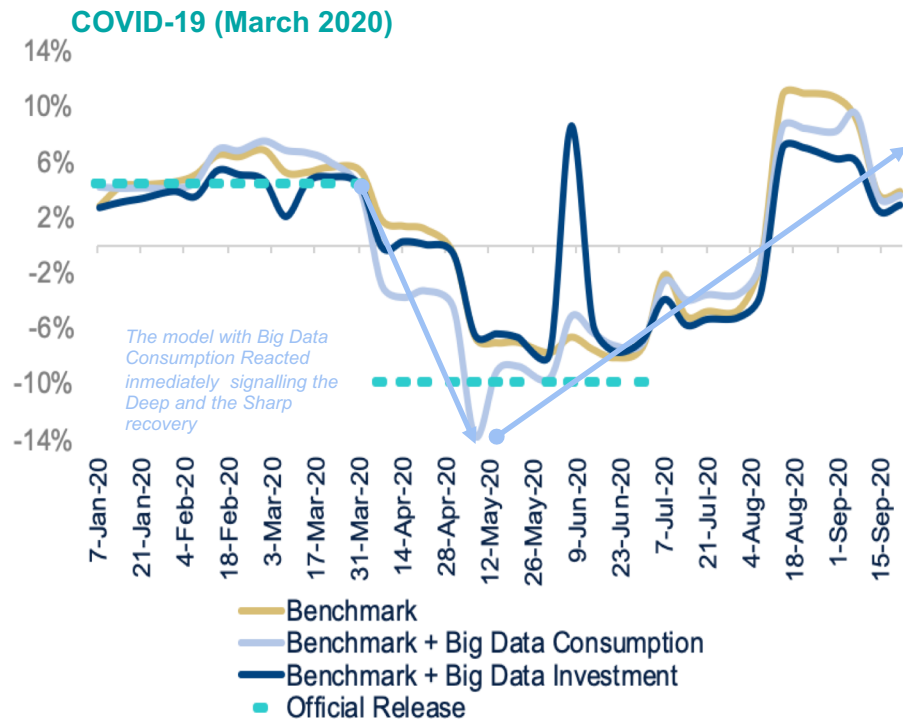
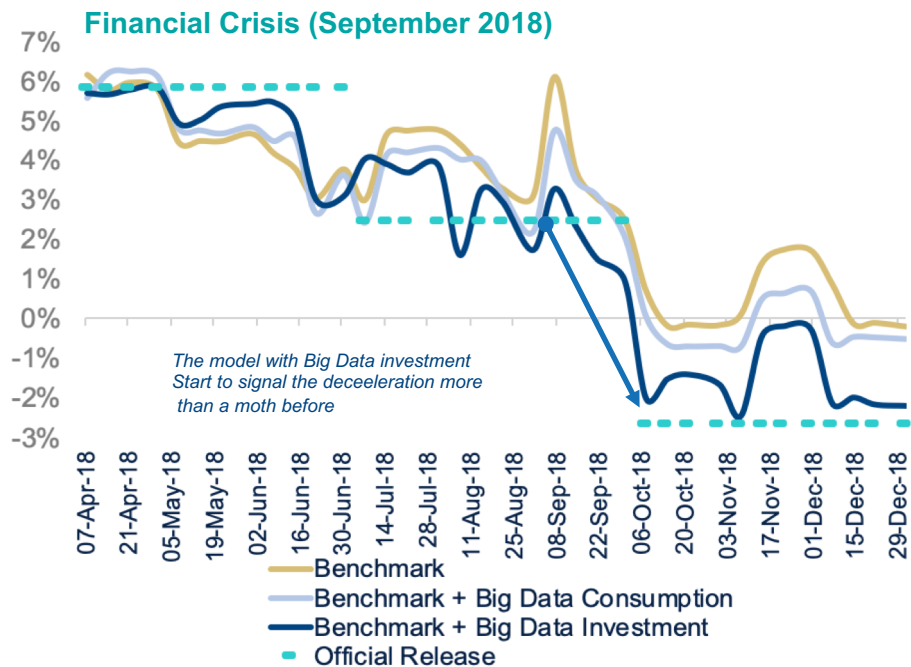
## TURKEY: OUT-OF-SAMPLE ERROR INVESTMENT MODEL



# Big Data & Nowcasting Model (DFM): Anticipation

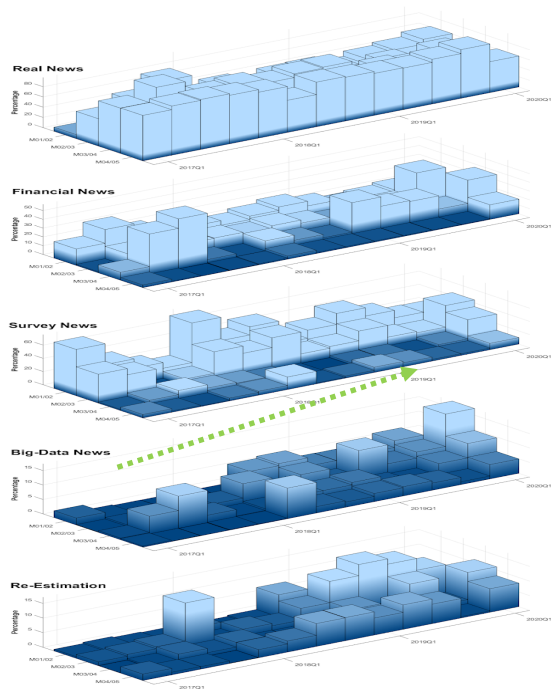
## TURKEY: NOWCASTING FINANCIAL CRISIS (SEPT 2018) & COVID CRISIS (MAR 2020)

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



# Big Data & Nowcasting Model (DFM): News & Prevalence Bias

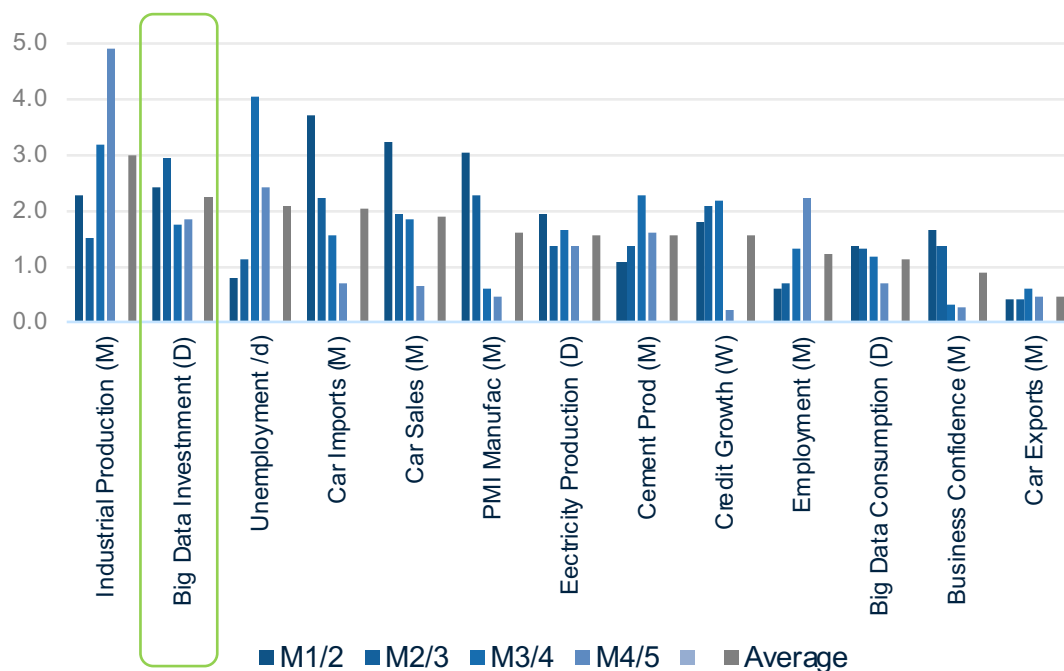
## DFM News Contributions (Unbalanced Sample)



Source: Own Elaboration

## DFM News Contributions correcting Bias

(Maintaining individually all the variables for the sample of Big data information)



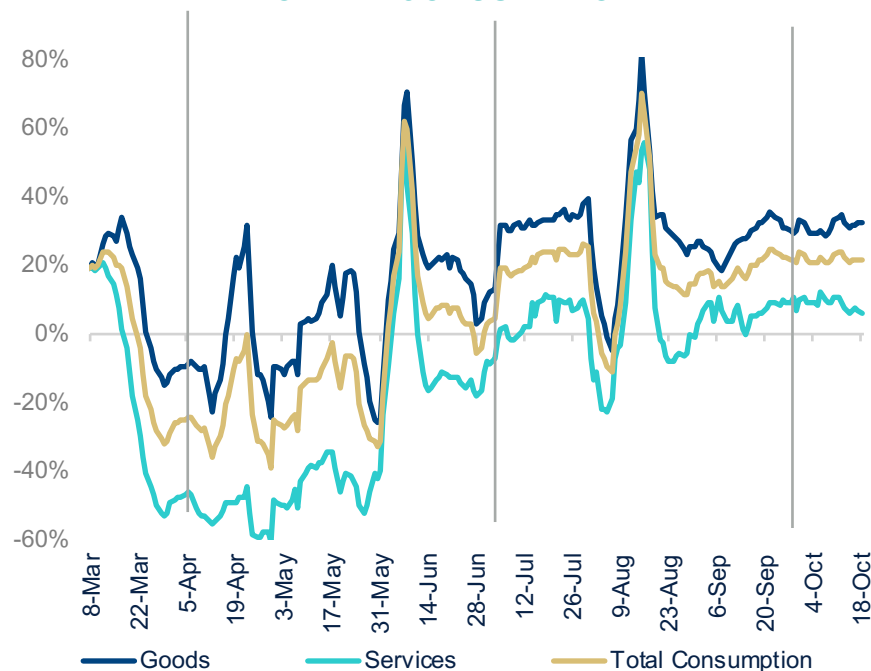
Source: Own Elaboration

# One of the key results is that we can examine Investment in Real Time

## TURKEY: BBVA BIG DATA CONSUMPTION & INVESTMENT

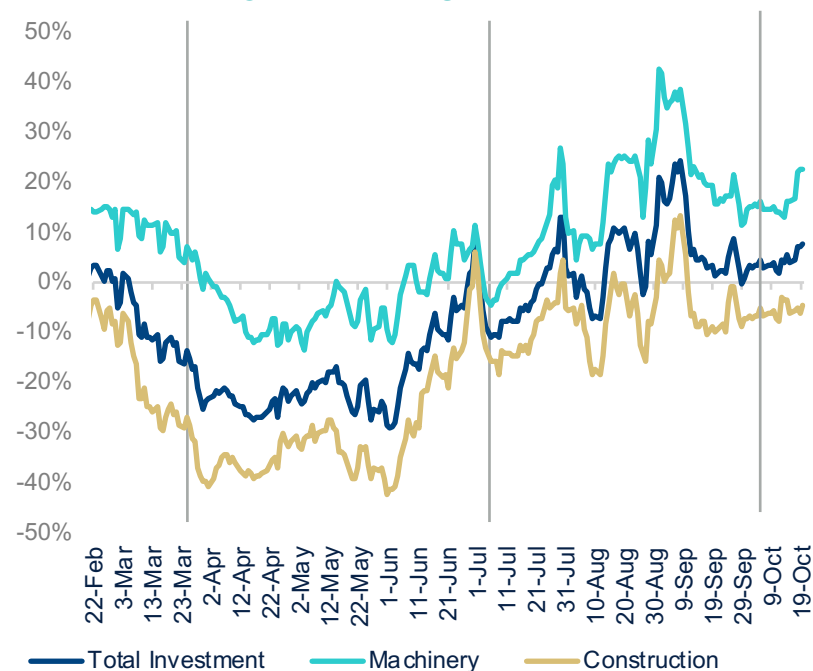
(7-day cum. YoY nominal in Cons., 28-day cum. YoY nominal in Invest.)

### BIG DATA CONSUMPTION



Source: Own Elaboration

### BIG DATA INVESTMENT



# The "High Definition" sectorial information can help us to track the evolution of sectors in detail and differentiate different shocks...

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

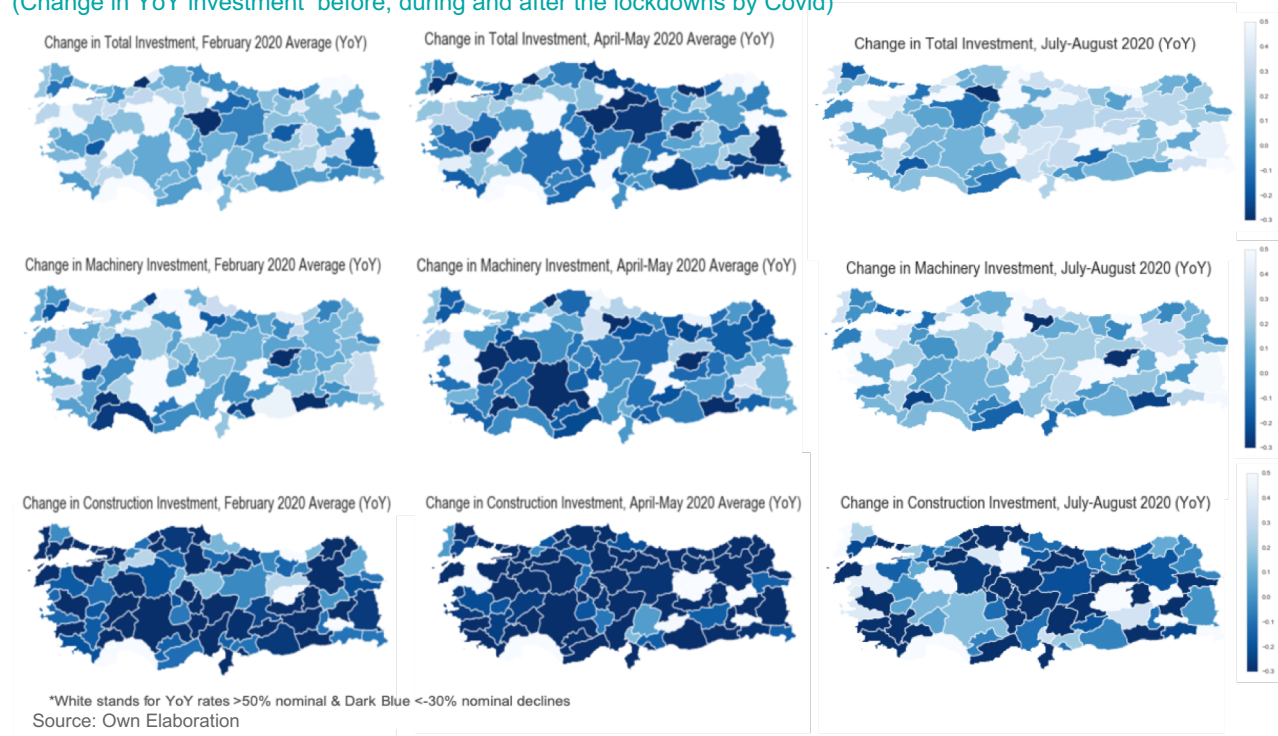
(3mm avg YoY nominal)



... while geo localization of the big data investment will help us to analyze shocks at regional levels to design targeted policies

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

(Change in YoY investment before, during and after the lockdowns by Covid)



# Investment in the rest of the countries: Validation

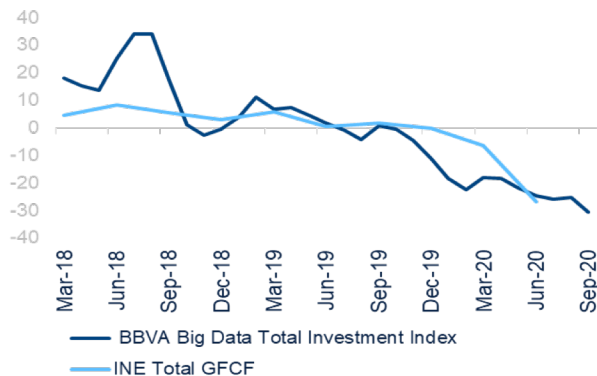


# Validation for Spain: Still working on raising Transaction Data

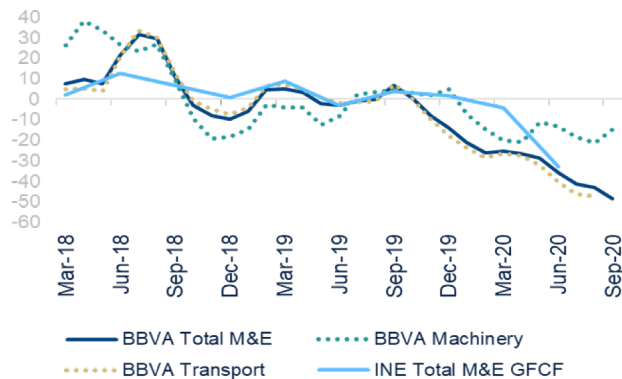
## SPAIN: BBVA BIG DATA INVESTMENT INDEX AND GROSS FIXED CAPITAL INVESTMENT COMPONENTS

(YoY Real deflated by Producer Prices, 3 months moving average)

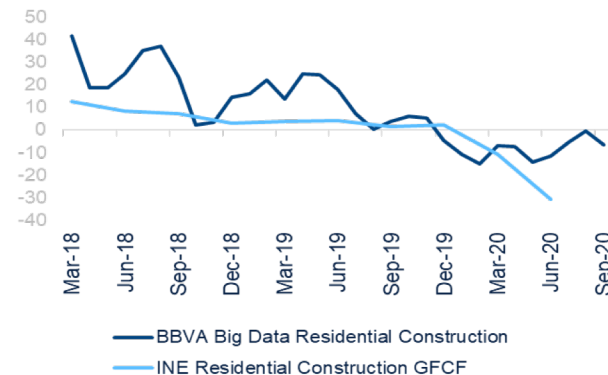
Correlation coefficient: 0.78



Correlation coefficient: 0.58



Correlation coefficient: 0.73

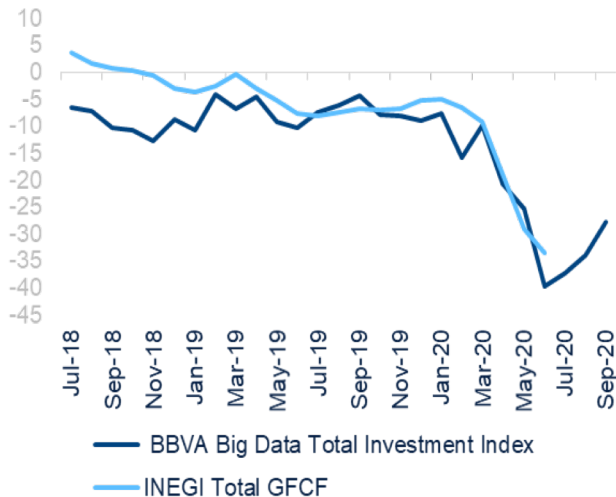


# Validation for Mexico: Positive results in Correlation & Comovement

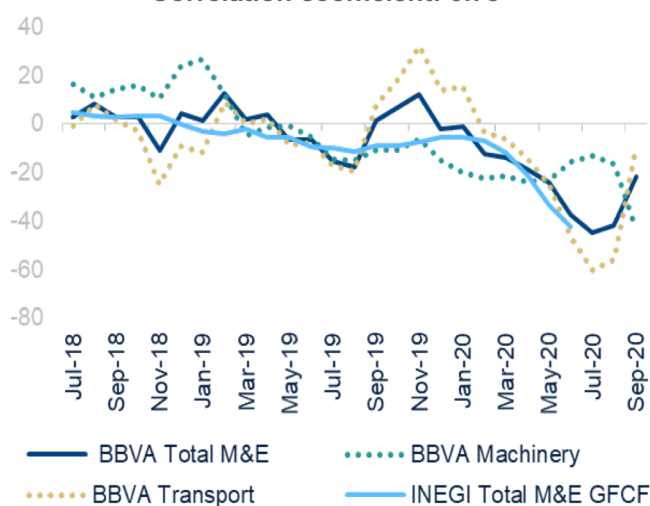
## MEXICO: BBVA BIG DATA INVESTMENT INDEX AND GROSS FIXED CAPITAL INVESTMENT COMPONENTS

(YoY Real deflated by Producer Prices, 3 months moving average)

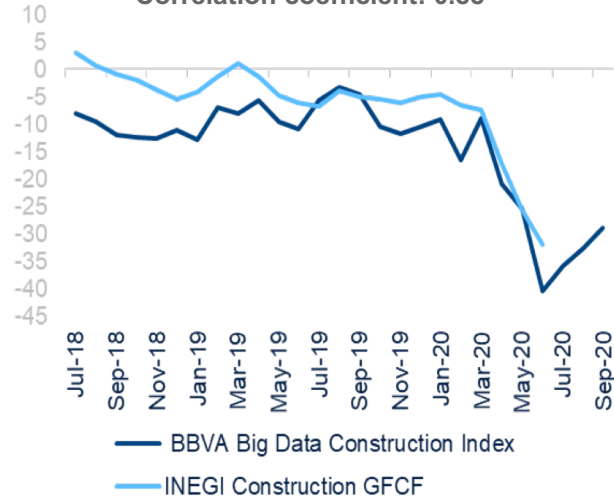
Correlation coefficient: 0.86



Correlation coefficient: 0.78



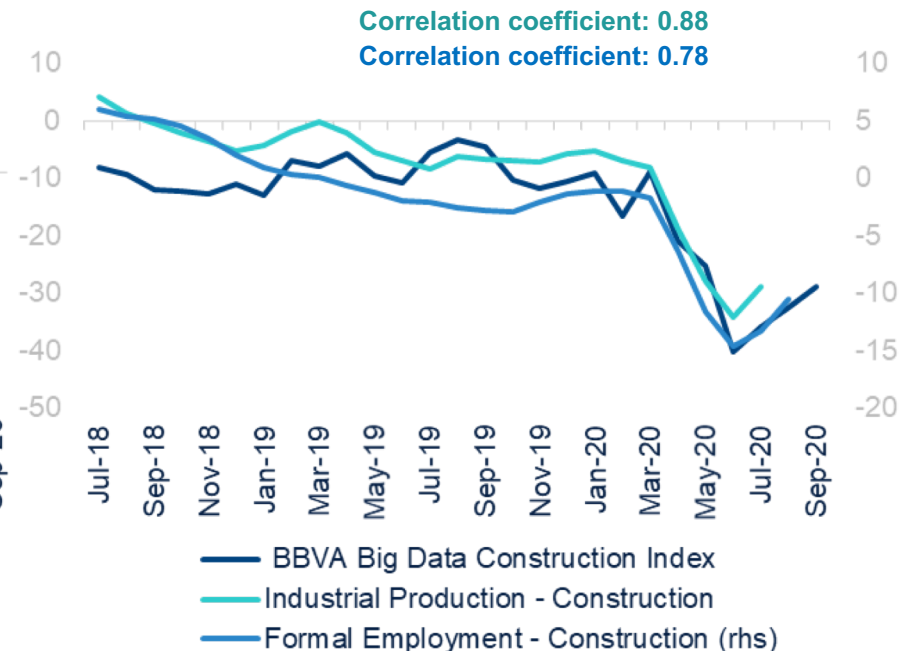
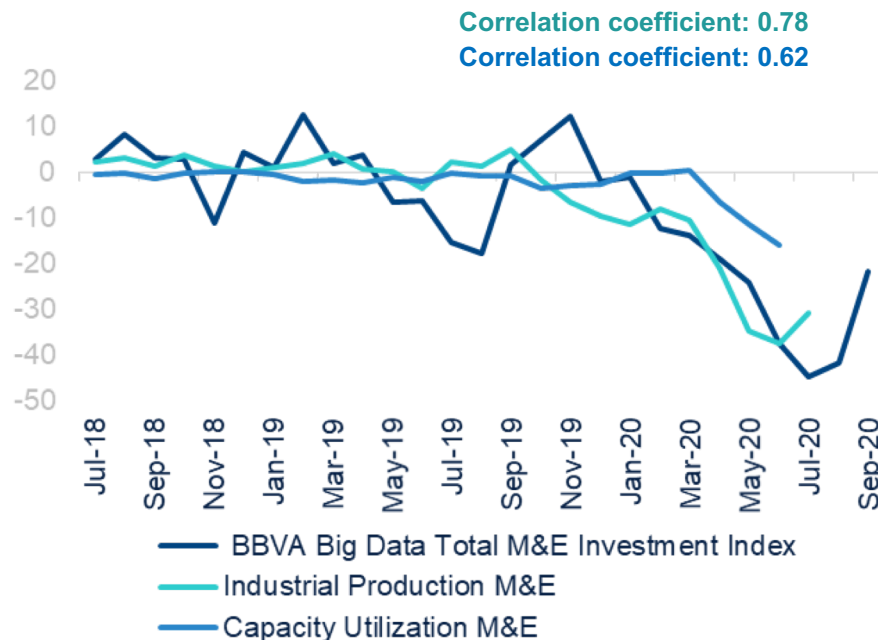
Correlation coefficient: 0.88



# Validation for Mexico: Positive results with HF proxies

## MEXICO: BBVA BIG DATA MACHINERY/CONSTRUCTION INVESTMENT INDEX AND PROXIES

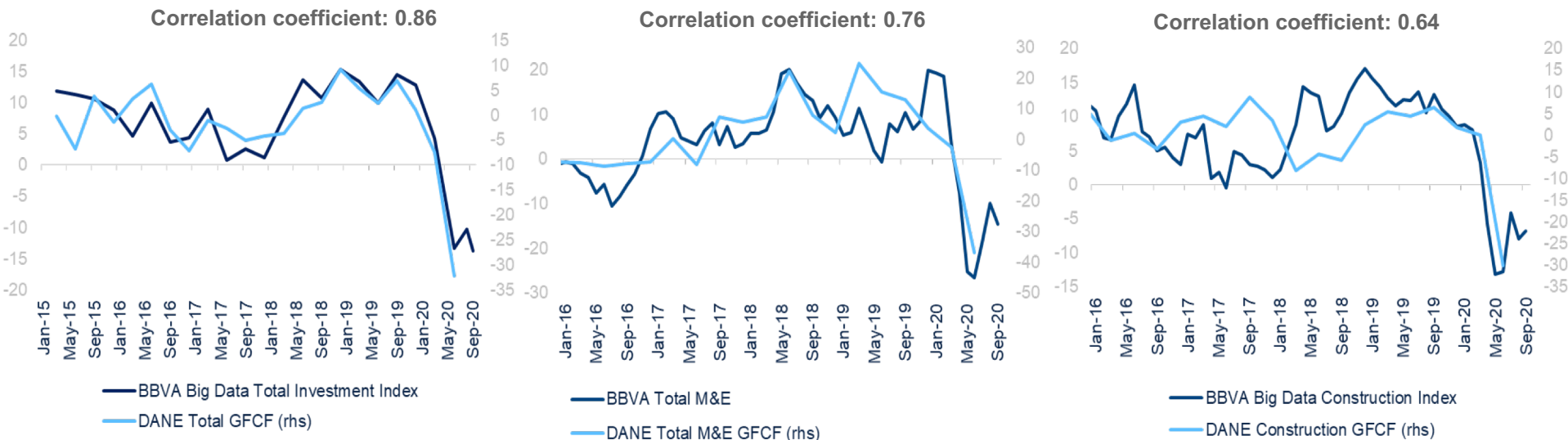
(YoY Real deflated by Producer Prices, 3 months moving average)



# Validation for Colombia: Positive results in Correlation & Comovement

## COLOMBIA: BBVA BIG DATA INVESTMENT INDEX AND GROSS FIXED CAPITAL INVESTMENT COMPONENTS

(YoY Real deflated by Producer Prices, 3 months moving average)



Conclusions  
Further Research  
References

## Conclusions and Further Research

- Big Data will become an increasing tool for economic analysis as they provide "real time" and "high definition" advantages
- We provide a novel and validated "real time" and "high definition" assessment of the Investment cycle for Turkey. Preliminary validation results show that this analysis can be extended to other countries
- The Big Data information can enhance the standard Nowcasting Models providing analyst and policymaker with an effective tool to react faster to shocks and design targeted policies
- Given the characteristic of Big Data information (HF, Granular but Short history) we need to explore the integration of the Big Data information in new models. Particularly Regularization Techniques to exploit high and rich dimension of the information (Large P) but in shorter samples (T) could enhance further the potential of Big Data for economic analysis

# References

- Akcigit,U & Yusuf Emre Akgunduz & Seyit Mumin Cilasun & Elif Ozcan Tok & Fatih Yilmaz, 2019. "[Facts on Business Dynamism in Turkey](#)," [Working Papers](#) 1930, Research and Monetary Policy Department, Central Bank of the Republic of Turkey.
- Andersen, A., Hansen, E. T., Johannesen, N., & Sheridan, A. (2020a). Consumer Responses to the COVID-19 Crisis: Evidence from Bank Account Transaction Data. CEPR Discussion Paper 14809.
- Andersen, A. L., Hansen, E. T., Johannesen, N., & Sheridan, A. (2020b). Pandemic, Shutdown and Consumer Spending: Lessons from Scandinavian Policy Responses to COVID-19. arXiv:2005.04630 [econ, q-fin].
- Baker, S. R., Farrokhnia, R., Meyer, S., Pagel, M., & Yannelis, C. (2020a). How Does Household Spending Respond to an Epidemic? Consumption During the 2020 COVID-19 Pandemic. Working Paper 26949, National Bureau of Economic Research.
- Baker, S. R., Farrokhnia, R. A., Meyer, S., Pagel, M., & Yannelis, C. (2020b). Income, Liquidity, and the Consumption Response to the 2020 Economic Stimulus Payments. Working Paper 27097, National Bureau of Economic Research.
- Banbura,M and Modugno (2014), [maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data](#), Journal of Applied Econometrics, 29, (1), 133-160
- Bodas, D., López, J. R. G., López, T. R., de Aguirre, P. R., Ulloa, C. A., Arias, J. M., de Dios Romero Palop, J., Lapaz, H. V., & Pacce, M. J. (2019). Measuring retail trade using card transactional data. Working Papers 1921, Banco de España;Working Papers Homepage.
- Bounie, D., Camara, Y., & Galbraith, J. W. (2020). Consumers' Mobility, Expenditure and Online-Offline Substitution Response to COVID-19: Evidence from French Transaction Data. SSRN Scholarly Paper ID 3588373, Social Science Research Network, Rochester, NY.
- Chetty, R., Friedman, J. N., Hendren, N., & 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.
- Chronopoulos, D. K., Lukas, M., & Wilson, J. O. S. (2020). Consumer Spending Responses to the COVID-19 Pandemic: An Assessment of Great Britain. Working Paper 20-012.
- Carvalho, V,Hansen,S,Ortiz,A, Garcia,JR, Rodrigo,T, Rodriguez Mora,S P,Ruiz (2020). Tracking the Covid-19 crisis with high resolution transaction data". CEPR DP No 14642
- Cox, N., Ganong, P., Noel, P., Vavra, J., Wong, A., Farrell, D., & Greig, F. (2020). Initial Impacts of the Pandemic on Consumer Behavior: Evidence from Linked Income, Spending, and Saving Data.
- Eurostat (2013) Handbook on quarterly national accounts
- Gezici F & Burçin Yazgı Walsh & Sinem Metin Kacar, 2017. "[Regional and structural analysis of the manufacturing industry in Turkey](#)," [The Annals of Regional Science](#), Springer;Western Regional Science Association, vol. 59(1), pages 209-230, July.
- Modugno, M., B. Soybilgen, and E. Yazgan (2016). Nowcasting Turkish GDP and News Decomposition. International Journal of Forecasting 32 (4), 1369–1384.
- Surico, P., Kanzig, D., & Hacıoglu, S. (2020). Consumption in the Time of COVID-19: Evidence from UK Transaction Data. CEPR Discussion Paper 14733.
- Watanabe, T., Omori, Y., et al. (2020). Online consumption during the Covid-19 crisis: Evidence from japan. Tech. rep., University of Tokyo, Graduate School of Economics.