Assessing Corruption with Big Data

March 2018

Assessing Corruption with Big Data

- ◆ **We build a Corruption Perception Index based on Google Trends Big Data** on searches about corruption. It covers more than 190 countries and, unlike traditional corruption indexes, it is available at real-time and with high-frequency since January of 2004.

- ◆ Data show that **the worldwide perception of corruption has been increasing since 2009-10**. There is a significant heterogeneity across countries, with a remarkable rise in such period **especially in regions such as Latin America.**

- ◆ We use our Corruption Perceptions Index to study **the case of Brazil**, where corruption scandals have been an important element of the political and economic environment in recent years.

- ◆ We show that **an increase in the perception of corruption has a significant effect on the government's approval rating in Brazil**. There is also **evidence that corruption perception impacts confidence indexes.**

Measuring corruption

BBVA | Research

# How do we build a Corruption Perception Index based on Google Trends Big Data?

◆ Searching for the topic "Corruption" at Google Trends (trends.google.com):
by searching for the topic rather than for the term "corruption" we make sure to take into account web searches including not only the exact term "corruption" but also the word "corruption" in other languages as well as misspellings and synonyms

◆ Defining the time range to extend from January 2004 until now

◆ Selecting the category "Law & Government" since our focus is on the misuse of public resources

◆ We first look at worldwide searches about the topic "corruption" and compare results for 191 countries. We then look at searches about "corruption" in Brazil, which will be our case study

◆ Google Trends provides relative rather than absolute data : "Numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means there was not enough data for this term"

BBVA | Research

Why do we build a Corruption Perception Index based on Google Trends Big Data?

◆ We build a Corruption Perception Index based on Google Trends to have a real-time, high frequency (monthly) indicator reflecting how people perceive corruption. Our indicator unveils some new features of corruption perception and allows for innovative analysis related to the issue.

◆ Most other corruption perception indexes are released with some delay, at annual frequency. On top of that, some of them do not exactly build on people's perception on corruption but rather on the opinion of experts or other, more structural, indicators

◆ In this sense, our index is more perceptional, more news-sensitive and potentially more volatile. We see it more suited to gauge effects of corruption on government approval ratings, confidence, electoral results, etc.

◆ Other corruption perception indexes, such as the most traditional one developed by Transparency International, are more structural, thus more suited to analysis of impact of corruption on inequality, development levels, etc.

BBVA | Research

The worldwide perception of corruption has been increasing since 2009-10

Corruption perception index, worldwide
(Index varying from 0 to 100)

Corruption perception index — Trend
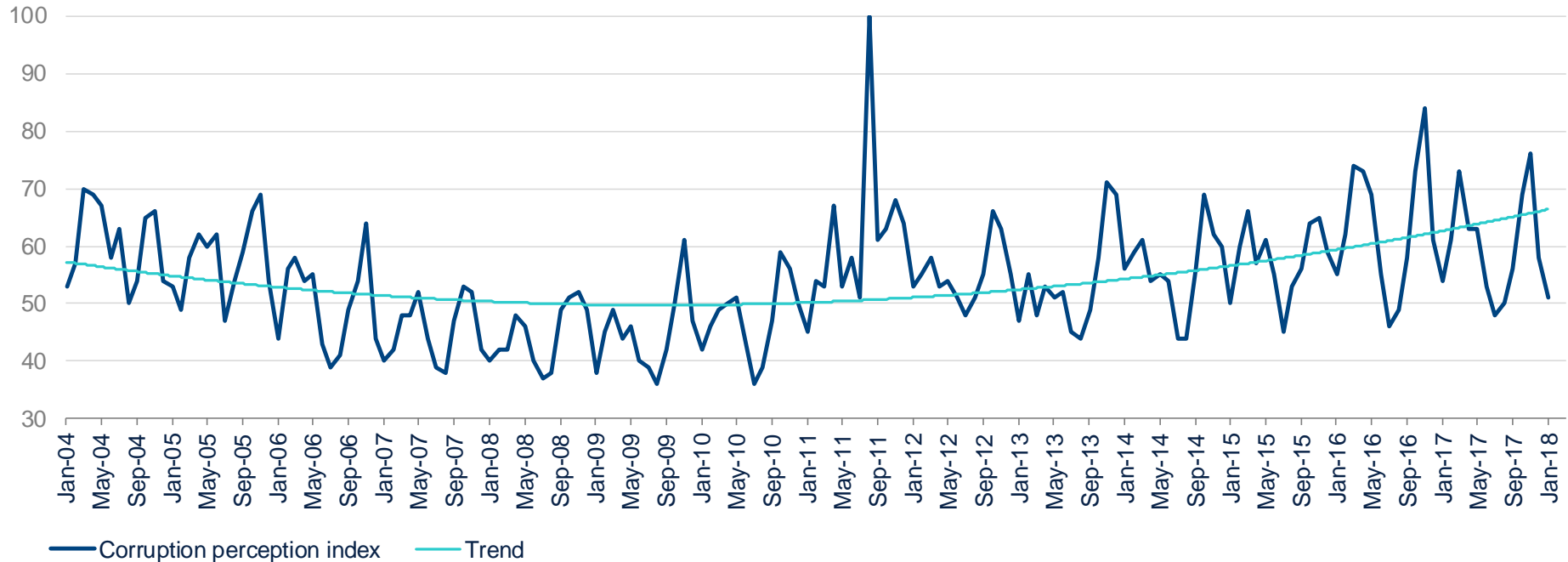
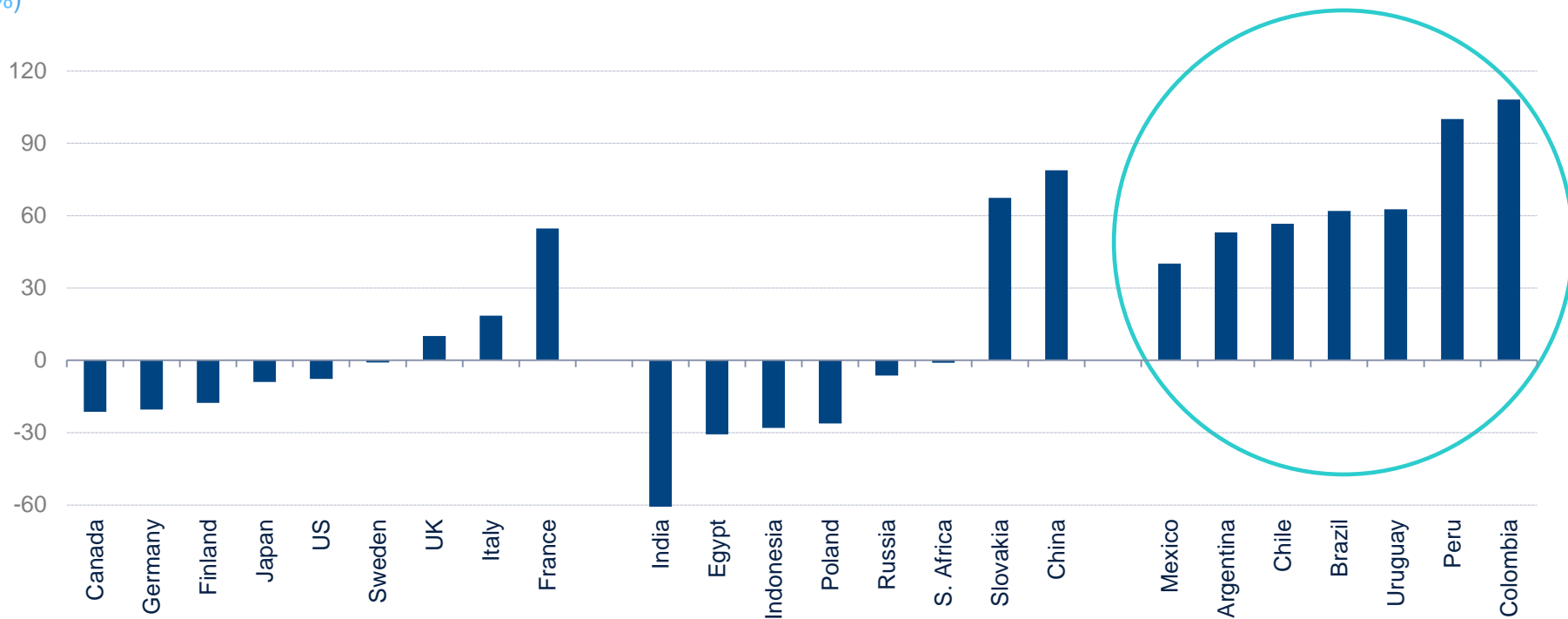Source: BBVA Research based on data from Google Trends

In this case, we look at worldwide searches including the topic "corruption"

Results show that the searches on "corruption" are becoming more and more common in comparison to other searches, suggesting an increasing concern worldwide about the issue

BBVA | Research

There exists an important heterogeneity across countries; the increase of the corruption perception in Latin America is particularly remarkable

Change in the corruption perception between 2017 and 2012, selected countries
(%)

Source: BBVA Research based on data from Google Trends

For a group of selected countries, we compare the frequency of searches on "corruption" in 2017 to the searches in 2012

Heterogeneity is significant. In regions such as Latin America there has been an important increase in the corruption perception, according to our index

7

Corruption perception is in general higher in less developed countries

Corruption perception index by country, 2017
(darker tones indicate higher perception of corruption)

Source: BBVA Research based on data from Google Trends.

When looking at worldwide searches on "corruption", Google Trends also provides data on the relative frequency of searches by country…

… which allows us to compare the perception of corruption for 191 countries. Results are unsurprising: in general, the perception of corruption is higher in less developed countries

Our corruption perception index compared to the one by Transparency International: positively correlated, but different by construction

Rankings of corruption perception, 2016
(higher position in the rankings represent higher corruption perception)
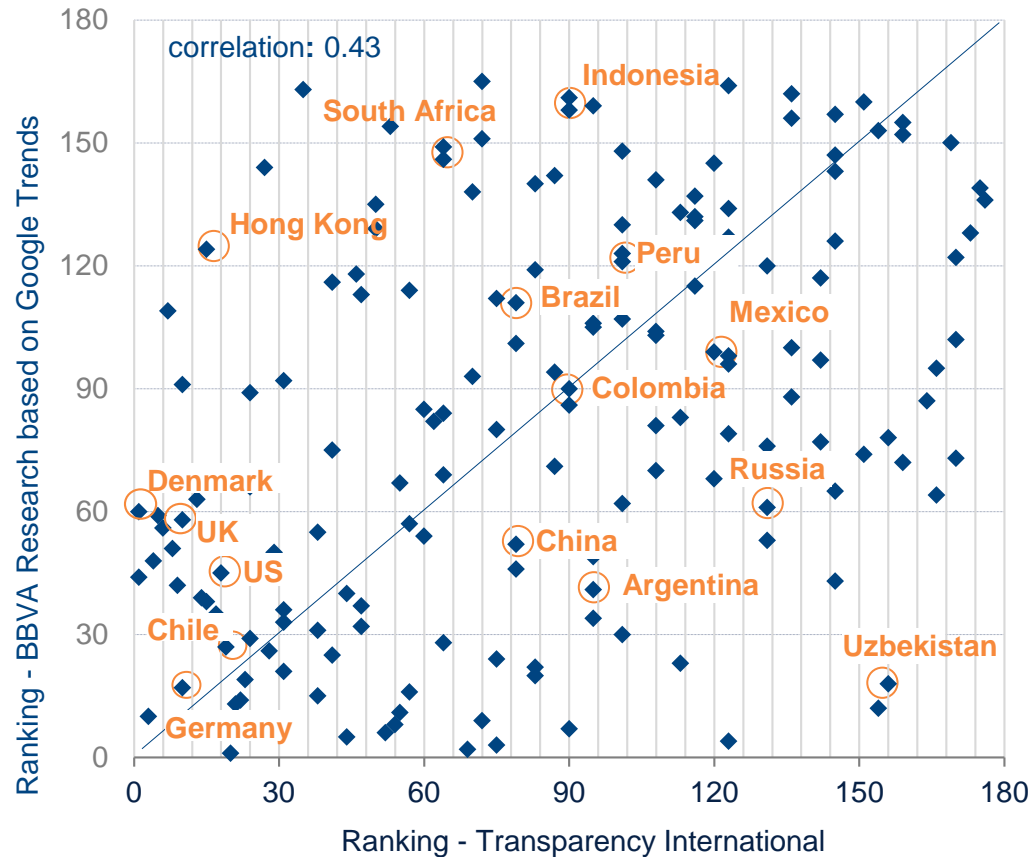
- We rank countries according to our index and then compare to the corruption perception ranking released by Transparency International

- Although different, the two rankings are positively correlated

- In some cases, such as in UK, US, Hong Kong, Indonesia, South Africa, etc. The perception of corruption based on web searches is higher than the Transparency International' index suggests. Taking comments on page 5 into account, in these places corruption is relatively more perceptional than structural

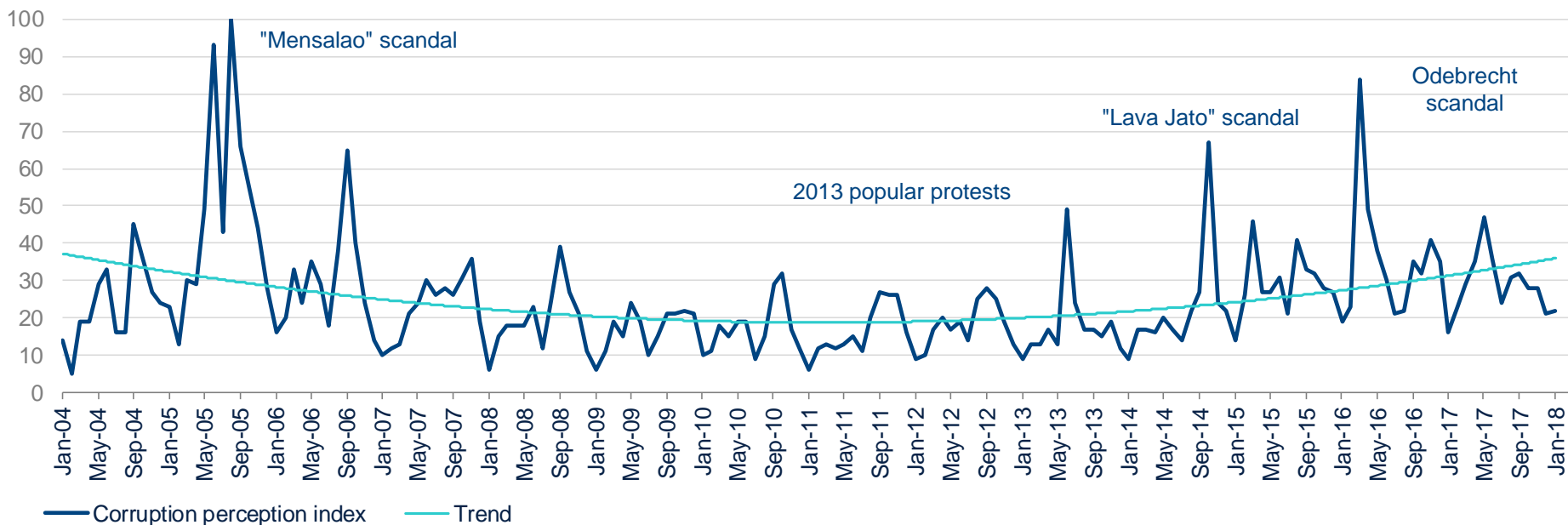- On the other hand, in China, Argentina, Uzbekistan, etc. corruption seems to be relatively more structural than perceptional

Assessing the impact of corruption in Brazil

BBVA | Research

Assessing the impact of corruption perception: the case of Brazil

Corruption perception index, Brazil
(Index varying from 0 to 100)

Source: BBVA Research based on data from Google Trends

As an example of the application of our Corruption Perception Index, we focus in the case of Brazil and estimate the impact it had on some political and economic variables

When we look at web searches about the topic "corruption" only in Brazil, we can see a clear upward trend in the last few years, suggesting Brazilians are increasingly concerned about the issue

11

The Corruption Perception Index for Brazil is negatively correlated with government approval ratings in the country

Corruption perception index and government approval ratings in Brazil
(Indexes varying from 0 to 100)

correlation: -0.40

Legend: Government approval rating — Corruption perception index
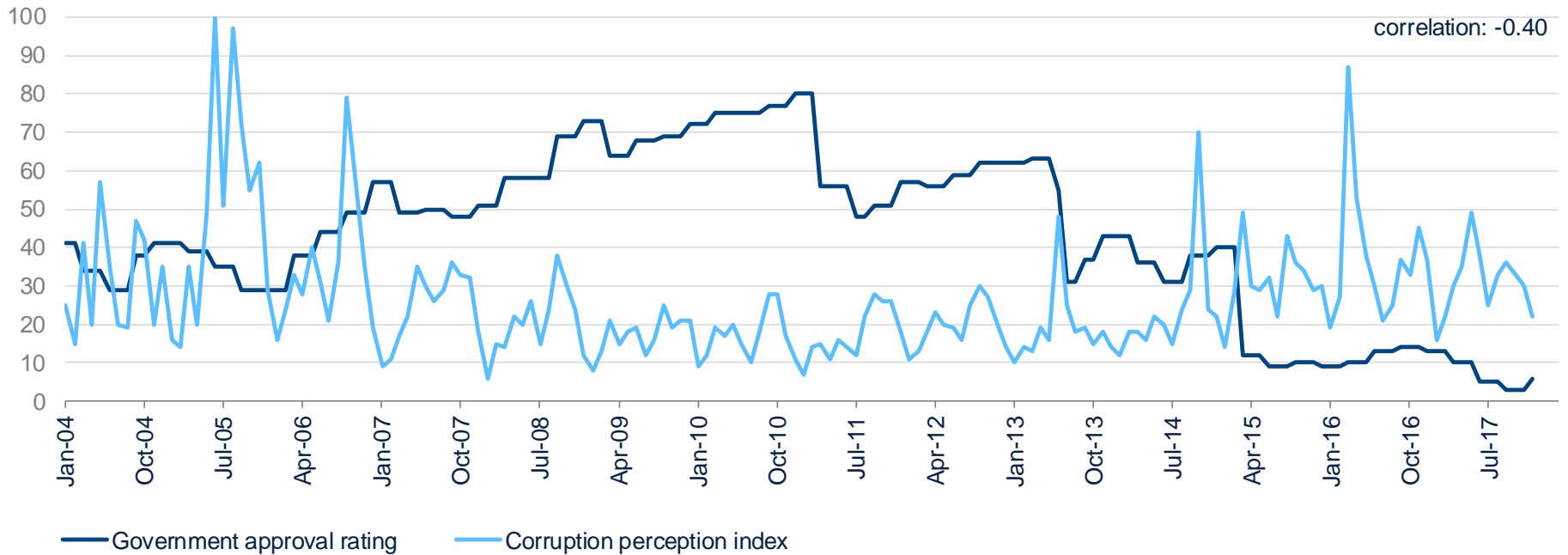
Source: BBVA Research based on data from Google Trends; CNI

A higher perception of corruption could be causing a drop in government approval ratings

The correlation between the two variables (-0.40) reinforce this claim

We perform some econometric exercises to formally test whether corruption perception affects government approval ratings in Brazil

◈ Following the literature on the issue and taking into account the availability of data, we propose an econometric model in which approval ratings are determined by i) our index of corruption perception, ii) the unemployment rate, iii) annual inflation, iv) Brazil's terms of trade, v) a dummy variable indicating the period in which Dilma Rousseff was the president, vi) other similar dummy variable for Michel Temer, vii) a dummy for the six first months of each government (to check for a possible "honeymoon effect"), and viii) a dummy variable for the Lehman Brothers crisis.

◈ We use monthly data ranging from January 2004 to December 2017

◈ More formally, this is our proposed econometric model:

$$Approval\ Rating = \alpha_0 + \boldsymbol{\alpha_1(Corruption\ Perception)} + \alpha_2(Unemployment\ Rate)$$

$$+\alpha_3(Inflation) + \alpha_4(Terms\ of\ Trade) + \alpha_5\ (dummy\ for\ D.Roussef) + \alpha_6\ (dummy\ for\ M.Temer)$$

$$+\alpha_7(dummy\ for\ honeymoon\ effect) + \alpha_8\ (dummy\ for\ Lehman\ Brothers\ crisis) + \mu$$

BBVA | Research

The results support the claim that corruption perception negatively affects approval ratings in Brazil

◆ The coefficient associated to the corruption perception index is negative and statistically significant, supporting the claim that corruption perception negatively affects approval ratings

◆ The coefficients of other variables are also significant and in line with expectations:

- Higher unemployment and higher inflation both drive approval ratings down;

- Higher terms of trade (which reflect a better external environment) drive approval ratings up;

- There is a negative effect related to the governments of Rousseff and mainly of Temer (in comparison to the government of Lula);

- There exists a honeymoon effect: approval ratings are higher during the first six months of each government;

- The Lehman Brothers crisis had a positive effect on approval ratings, in line with findings for other countries showing that approval ratings increase during adverse periods (war, external crisis, etc.)

◆ More details, including additional estimations, in the Annex

| OLS estimation results: "approval ratings" model	
(independent variables)	(associated coefficients)
Corruption Perception	-0.13 ***
Unemployment	-1.56 ***
Inflation	-1.20 **
Terms of Trade	1.18 ***
Rousseff	-31.2 ***
Temer	-52.1 ***
Honeymoon effect	6.31 ***
LB crisis	18.5 ***

*** Significant at 1%. ** Significant at 5%. * Significant at 10%. Sample size: 168 months (Jan 2004 to Dec 2017). R2=0.86. Source: BBVA Research.

BBVA | Research

On top of political consequences, corruption perception could also have an economic effect; we test its impact on confidence indicators
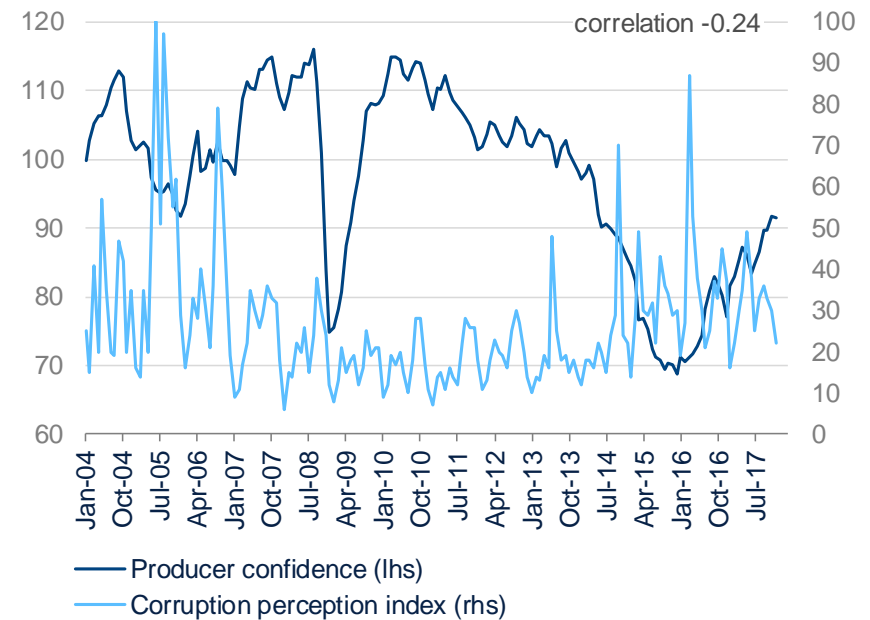
Corruption perception index and consumer confidence
(CPI ranges from 0 to 100; confidence ranges from 0 to 200)

correlation -0.47

— Consumer confidence (lhs)
— Corruption perception index (rhs)
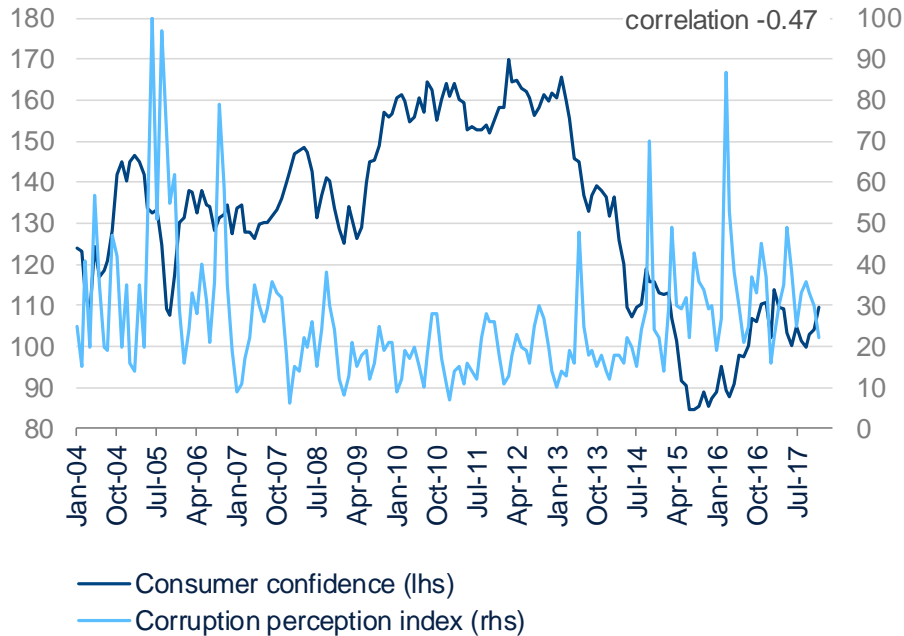
Corruption perception index and producer confidence
(CPI ranges from 0 to 100; confidence ranges from 0 to 200)

correlation -0.24

— Producer confidence (lhs)
— Corruption perception index (rhs)

Source: BBVA Research based on data from Google Trends; Producer confidence index: FGV; Consumer confidence index: FECOMERCIO

In theory, higher corruption perception could have a negative effect on both producer and consumer confidence indexes

We test that using two models, one with consumer confidence and other with producer confidence as dependent variable (we keep the same independent variables used in the model for approval ratings)

BBVA | Research

The results show that corruption perception negatively affects consumer confidence

♦ The coefficient associated to the corruption perception index is negative and statistically significant, supporting the claim that corruption perception negatively affects consumer confidence.

♦ Regarding the coefficients of other variables:

- Unemployment, honeymoon effect and the Lehman Brothers crisis do not significantly drive consumer confidence;

- Higher inflation drives consumer confidence down;

- Higher terms of trade drive consumer confidence up;

- There is a negative effect related to the governments of Rousseff and mainly of Temer (in comparison to the government of Lula)

♦ More details, including additional estimations, in the Annex

OLS estimation results: "consumer confidence" model	
(independent variables)	(associated coefficients)
Corruption Perception	-0.15 **
Unemployment	1.10
Inflation	-1.42 **
Terms of Trade	1.74 ***
Rousseff	-22.2 ***
Temer	-46.8 ***
Honeymoon effect	-0.92
LB crisis	-3.80

*** Significant at 1%. ** Significant at 5%. * Significant at 10%. Sample size: 168 months (Jan 2004 to Dec 2017). R^2=0.73
Source: BBVA Research

BBVA | Research

There exists also evidence on a negative effect of corruption perception on producer confidence

◆ The coefficient associated to the corruption perception index is negative and statistically significant, supporting the claim that corruption perception negatively affects producer confidence.

◆ Regarding the coefficients of other variables:

- higher unemployment drives producer confidence up (maybe not surprisingly given that producers can benefit from less tight labor markets);

- higher inflation drives consumer confidence down;

- higher terms of trade drive producer confidence up;

- there is a negative effect related to the governments of Rousseff and mainly of Temer (in comparison to the government of Lula);

- honeymoon effect is not significant;

- the Lehman Brothers crisis had a negative effect on producer confidence

◆ More details, including additional estimations, in the Annex

OLS estimation results: "producer confidence" model	
(independent variables)	(associated coefficients)
Corruption Perception	-0.07 **
Unemployment	1.82 ***
Inflation	-1.60 ***
Terms of Trade	0.88 ***
Rousseff	-12.7 ***
Temer	-29.2 ***
Honeymoon effect	1.16
LB crisis	-22.3 ***

*** Significant at 1%. ** Significant at 5%. * Significant at 10%. Sample size: 168 months (Jan 2004 to Dec 2017). R2=0.81. Source: BBVA Research .

Final comments on the impact of corruption perception in Brazil

◆ Although we are focusing in only one of the many channels through which corruption can affect some variables, our results show that corruption does have an immediate and negative political and economic impact in Brazil

◆ The coefficients associated with corruption perception in approval ratings, consumer confidence and producer confidence models are consistent with (average) elasticities of -0.21, -0.04 and -0.02 respectively (*)

◆ At the beginning of 2016, for example, the perception of corruption increased around 340%. According to our results, that reduced approval ratings by 50%, consumer confidence by 14% and producer confidence by 7%

Impact of a 10% increase in corruption perception on approval ratings, consumer and producer confidence (*)
(%)
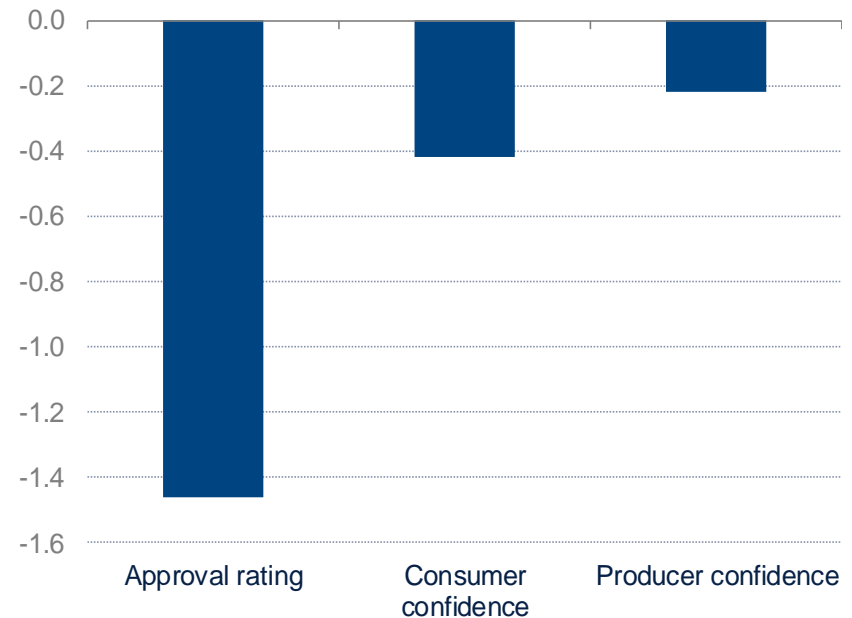
(*) Based on the coefficients of OLS regressions as well as on the coefficients of IV-GMM regressions displayed in the Annex
Source: BBVA Research.

Assessing Corruption with Big Data

- **We build a Corruption Perception Index based on Google Trends Big Data** on searches about corruption. It covers more than 190 countries and, unlike traditional corruption indexes, it is available at real-time and with high-frequency since January of 2004.

- Data show that **the worldwide perception of corruption has been increasing since 2009-10**. There is a significant heterogeneity across countries, with a remarkable rise in such period **especially in regions such as Latin America.**

- We use our Corruption Perceptions Index to study **the case of Brazil**, where corruption scandals have been an important element of the political and economic environment in recent years.

- We show that **an increase in the perception of corruption has a significant effect on the government's approval rating in Brazil**. There is also **evidence that corruption perception impacts confidence indexes.**
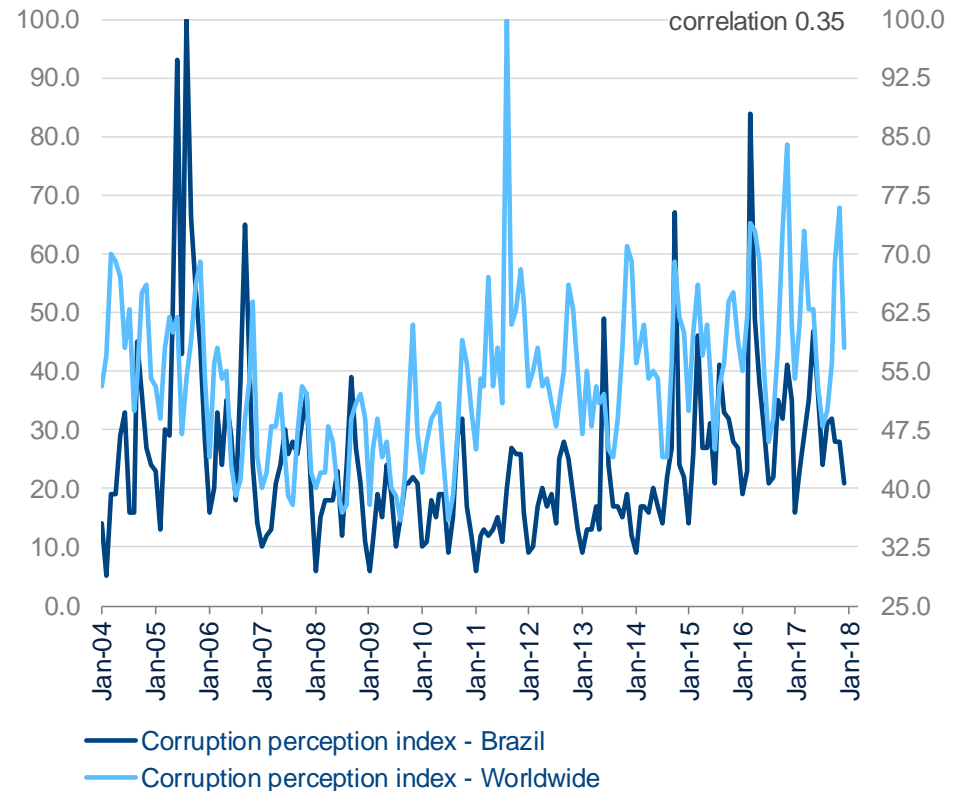
Annex

BBVA | Research

In our model, the corruption perception variable could be endogenous, so we use worldwide corruption perception as instrumental variable

◆ Worldwide corruption perception impacts the corruption perception in Brazil (formal econometric tests support this claim) and is not impacted by government approval ratings and confidence indicators in Brazil. Thus, it should be a valid instrumental variable (IV)

◆ We also use the lags of corruption perception in Brazil as IVs

◆ We reestimate previously proposed models using IV-GMM

◆ In fact, to address potential problems due to residuals' heteroscedasticity or serial correlation, we use the Newey-West' heteroscedasticity and autocorrelation consistent (HAC) estimator

Corruption perception index, Brazil and worldwide
(Indexes varying from 0 to 100)

correlation 0.35

— Corruption perception index - Brazil
— Corruption perception index - Worldwide

Source: BBVA Research based on data from Google Trends

BBVA | Research

IV-GMM estimations reinforce previous results, in particular it supports that corruption perception negatively affects approval ratings

IV-GMM estimation results: "approval ratings" model			
(independent variables)	IV: worldwide corruption perception (associated coefficients)	IV: three first lags of corruption perception in Brazil (associated coefficients)	IV: worldwide corruption perception AND three first lags of corruption perception in Brazil (associated coefficients)
Corruption Perception	-0.30 **	-0.36 ***	-0.35 ***
Unemployment	-1.34 **	-1.40 **	-1.38 **
Inflation	-1.05 *	-1.01	-1.01
Terms of Trade	1.10 ***	1.04 ***	1.05 ***
Rousseff	-30.36 ***	-30.85 ***	-30.58 ***
Temer	-50.76 ***	-50.97 ***	-51.05 ***
Honeymoon effect	5.16 *	4.94	5.01 *
LB crisis	16.1 ***	14.59 ***	14.95 ***

In all cases, the Newey-West HAC estimator with three lags is used. J-specifications tests support the validity of the used instruments.
*** Significant at 1%. ** Significant at 5%. * Significant at 10%. Sample size: 168 months (Jan 2004 to Dec 2017).
Source: BBVA Research

BBVA | Research

IV-GMM estimations in general support the claim that corruption perception negatively impacts consumer confidence

	IV-GMM estimation results: "consumer confidence" model		
(independent variables)	IV: worldwide corruption perception (associated coefficients)	IV: three first lags of corruption perception in Brazil (associated coefficients)	IV: worldwide corruption perception AND three first lags of corruption perception in Brazil (associated coefficients)
Corruption Perception	-0.07	-0.39 **	-0.32 **
Unemployment	1.00	1.69	1.21
Inflation	-1.48	-1.29	-1.59
Terms of Trade	1.77 ***	1.65 ***	1.66 ***
Rousseff	-22.63 ***	-20.04 ***	-18.44 ***
Temer	-47.33 ***	-45.49 ***	-44.18 ***
Honeymoon effect	-0.45	-2.97	-3.12
LB crisis	-2.83	-6.78 *	-4.57

In all cases, the Newey-West HAC estimator with three lags is used. J-specification tests support the validity of the used instruments.
*** Significant at 1%. ** Significant at 5%. * Significant at 10%. Sample size: 168 months (Jan 2004 to Dec 2017).
Source: BBVA Research

BBVA | Research

IV-GMM estimations in general also reinforce previous results regarding the negative effect of corruption perception on producer confidence

	IV-GMM estimation results: "producer confidence" model		
(independent variables)	IV: worldwide corruption perception (associated coefficients)	IV: three first lags of corruption perception in Brazil (associated coefficients)	IV: worldwide corruption perception AND three first lags of corruption perception in Brazil (associated coefficients)
Corruption Perception	-0.02	-0.13 **	-0.11 *
Unemployment	1.77 ***	2.04 ***	2.03 ***
Inflation	-1.64 ***	-1.78 ***	-1.90 ***
Terms of Trade	0.90 ***	0.85 ***	1.05 ***
Rousseff	-12.97 ***	-11.43 ***	-11.00 ***
Temer	-29.51 ***	-29.03 ***	-29.15 ***
Honeymoon effect	1.43	1.41	1.44
LB crisis	-21.79 ***	-22.89 ***	-22.35 ***

In all cases, the Newey-West HAC estimator with three lags is used. J-specification tests support the validity of the used instruments.
*** Significant at 1%. ** Significant at 5%. * Significant at 10%. Sample size: 168 months (Jan 2004 to Dec 2017).
Source: BBVA Research

This report has been produced by the South America Unit

Enestor Dos Santos
enestor.dossantos@bbva.com

BBVA-Research
Jorge Sicilia Serrano

Macroeconomic analysis
Rafael Doménech
r.domenech@bbva.com

Global Economic Situations
Miguel Jiménez
mjimenezg@bbva.com

Global Financial Markets
Sonsoles Castillo
s.castillo@bbva.com

Long-Term Global Modelling and Analysis
Julián Cubero
juan.cubero@bbva.com

Innovation and Processes
Oscar de las Peñas
oscar.delaspenas@bbva.com

Financial Systems and Regulation
Santiago Fernández de Lis
sfernandezdelis@bbva.com

Digital Regulation and Trends
Álvaro Martín
alvaro.martin@bbva.com

Regulation
Ana Rubio
arubiog@bbva.com

Financial Systems
Olga Cerqueira
Olga.gouveia@bbva.com

Spain and Portugal
Miguel Cardoso
miguel.cardoso@bbva.com

United States
Nathaniel Karp
Nathaniel.karp@bbva.com

Mexico
Carlos Serrano
carlos.serranoh@bbva.com

Turkey, China and Big Data
Álvaro Ortiz
alvaro.ortiz@bbva.com

Turkey
Álvaro Ortiz
alvaro.ortiz@bbva.com

Asia
Le Xia
Le.xia@bbva.com

South America
Juan Manuel Ruiz
juan.ruiz@bbva.com

Argentina
Gloria Sorensen
gsorensen@bbva.com

Chile
Jorge Selaive
jselaive@bbva.com

Colombia
Juana Téllez
juana.tellez@bbva.com

Peru
Hugo Perea
hperea@bbva.com

Venezuela
Julio Pineda
juliocesar.pineda@bbva.com