OECD DEVELOPMENT CENTRE

Working Paper No. 343

Using Google data to understand governments' approval in Latin America

Nathalia Montoya, Sebastián Nieto-Parra, René Orozco and Juan Vázquez Zamora
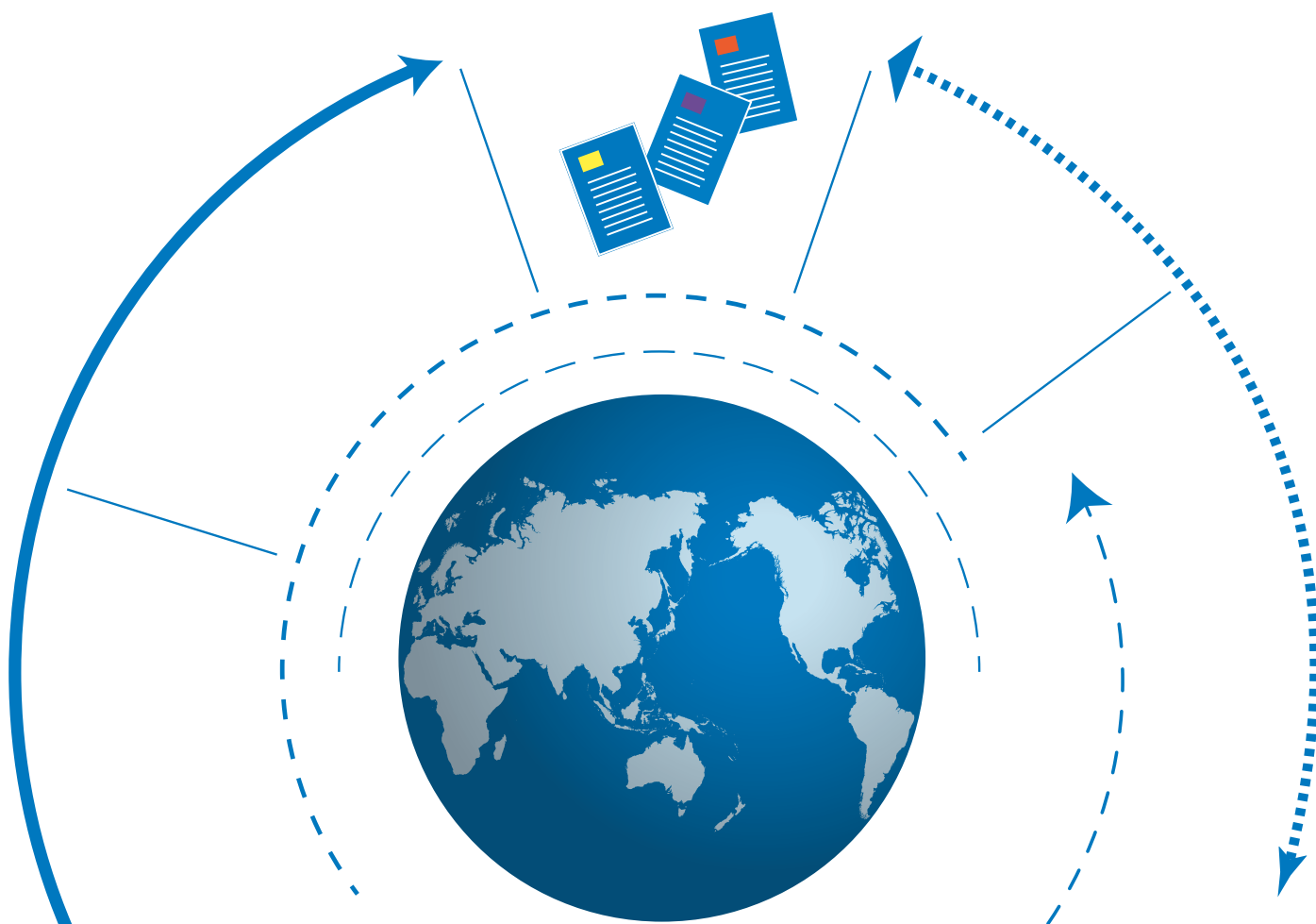
DEVELOPMENT CENTRE
WORKING PAPERS

This series of working papers is intended to disseminate the Development Centre's research findings rapidly among specialists in the field concerned. These papers are generally available in the original English or French, with a summary in the other language.

OECD Working Papers should not be reported as representing the official views of the OECD or of its member countries. The opinions expressed and arguments employed are those of the authors.

This document, as well as any data and map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Working Papers describe preliminary results or research in progress by the author(s) and are published to stimulate discussion on a broad range of issues on which the OECD works. Comments on Working Papers are welcomed, and may be sent to the OECD Development Centre, 2 rue André Pascal, 75775 PARIS CEDEX 16, France; or to dev.contact@oecd.org. Documents may be downloaded from: www.oecd.org/dev/wp.

CENTRE DE DÉVELOPPEMENT
DOCUMENTS DE TRAVAIL

Cette série de documents de travail a pour but de diffuser rapidement auprès des spécialistes dans les domaines concernés les résultats des travaux de recherche du Centre de développement. Ces documents ne sont disponibles que dans leur langue originale, anglais ou français ; un résumé du document est rédigé dans l'autre langue.

Les documents de travail de l'OCDE ne doivent pas être présentés comme exprimant les vues officielles de l'OCDE ou de ses pays membres. Les opinions exprimées et les arguments employés sont ceux des auteurs.

Ce document, ainsi que les données et cartes qu'il peut comprendre, sont sans préjudice du statut de tout territoire, de la souveraineté s'exerçant sur ce dernier, du tracé des frontières et limites internationales, et du nom de tout territoire, ville ou région.

Les documents de travail exposent des résultats préliminaires ou des travaux de recherche en cours menés par l'auteur/les auteurs et sont publiés pour stimuler le débat sur un large éventail de questions sur lesquelles l'OCDE travaille. Les commentaires sur les documents de travail sont bienvenus et peuvent être adressés au Centre de développement de l'OCDE, 2 rue André Pascal, 75775 PARIS CEDEX 16, France; ou à dev.contact@oecd.org. Les documents peuvent être téléchargés à partir de: www.oecd.org/dev/wp.

Table of contents

Tables

Figures

Acknowledgements

* Nathalia Montoya González (OECD Development Centre), Sebastián Nieto-Parra (OECD Development Centre), René Orozco (OECD Development Centre) and Juan Vázquez Zamora (OECD Development Centre).

Preface

Citizens' perceptions of economic, social and institutional outcomes are closely linked to government approval rates. Understanding citizens' perceptions is therefore essential for assessing policy impacts and the need for corrective action, as well as for defining policy priorities and how they are communicated. The urgency of engaging with citizens and fostering trust in government becomes particularly apparent when discontent with public services and institutions drives mass protests, as has been the case recently across Latin America. A more transparent and participatory approach to policy design and implementation is a key part of rebuilding a social pact in the region. Re-founding the social pact has become even more relevant in these challenging times. As the COVID-19 pandemic and global crisis erode development gains, the ability of governments to implement effective mitigation and recovery measures will critically require public trust and support.

An accurate understanding of the drivers of well-being requires moving beyond purely income indicators such as Gross Domestic Product (GDP) to exploit a wider range of metrics, including subjective measures of "how is life". As GDP per capita has increased over the last decades, non-income factors have become increasingly important to improve people's lives, as emphasised in the "development in transition" approach developed by the European Union, the OECD Development Centre and ECLAC. In fact, well-being outcomes gradually delink from income levels as countries become wealthier.

The digital transformation has made new information and analytical tools available, bringing unprecedented opportunities to analyse citizens' well-being, aspirations and perceptions. This paper uses new data sources to gauge citizens' perceptions of public policies. More specifically, it analyses the drivers of citizens' approval of governments in Latin American countries, by using Internet search query data to reveal citizens' main social concerns.

This paper contributes to the debate on trust in government and effective policy making in two ways. First, the empirical analysis shows that there is a negative and statistically significant correlation between the executive's approval rates and perceptions of corruption, of insecurity and complaints related to the quality of public services. Second, it puts forward the use of digital tools, such as Internet search query data, as a useful source to inform policy making, given its anonymity and the availability of high-frequency series in real-time.

This paper is a background document for the *Latin American Economic Outlook 2020*, the Development Centre's annual flagship report on the region. It also provides insights on citizens' perceptions in the context of the OECD's work at country level, such as its Multi-dimensional Country Reviews. Finally, it contributes to research on the use of new technologies to analyse subjective measures of well-being, including the project on "Policy Metrics for Well-being and Sustainable Development in Latin America and the Caribbean", undertaken as part of the EU Regional Facility for Development in Transition in Latin America and the Caribbean.

<div align="center">
Mario Pezzini
Director, OECD Development Centre
and Special Advisor to the OECD Secretary-General on Development
</div>

Abstract

This paper studies the potential drivers of governments' approval rates in 18 Latin American countries using Internet search query data from *Google Trends* and traditional data sources. It employs monthly panel data between January 2006 and December 2015. The analysis tests several specifications including traditional explanatory variables of governments' approval rates – i.e. inflation, unemployment rate, GDP growth, output gap – and subjective explanatory variables – e.g. perception of corruption and insecurity. For the latter, it uses Internet search query data to proxy citizens' main social concerns, which are expected to drive governments' approval rates. The results show that the perception of corruption and insecurity, and complaints about public services have a statistically significant association with governments' approval rates. This paper also discusses the potential of Internet search query data as a tool for policy makers to understand better citizens' perceptions, since it provides highly anonymous and high-frequency series in real-time.

JEL classification: D72; H11; O3.

Keywords: Social contract; governments' approval; citizens' perceptions; Latin America; big data

Résumé

Ce document analyse les déterminants potentiels des taux d'approbation des gouvernements dans 18 pays d'Amérique latine en utilisant les données de requête de recherche Internet de Google Trends et les sources de données traditionnelles. Il utilise des données de panel mensuelles entre janvier 2006 et décembre 2015. L'analyse teste plusieurs spécifications, y compris les variables explicatives traditionnelles d'approbation des gouvernements – i.e. inflation, taux de chômage, croissance du PIB, écart de production – et des variables explicatives subjectives (p.ex. perception de la corruption et de l'insécurité). Pour ce dernier, les données des requêtes de recherche sur Internet sont utilisées pour cerner les principales préoccupations sociales des citoyens, qui sont les déterminants attendus des niveaux d'approbation des gouvernements. Les résultats montrent que la perception de la corruption et de l'insécurité ainsi que les plaintes concernant les services publics ont en effet une association statistiquement significative avec les taux d'approbation des gouvernements. Sources anonymes, les données des requêtes de recherche sur Internet mettent à disposition des gouvernements des séries haute fréquence en temps réel, utiles pour suivre l'impact dans le temps de l'élaboration des politiques sur la perception des citoyens.

Classification JEL : D72 ; H11 ; O3.

Mots clés : Contrat social ; approbation des gouvernements ; perceptions des citoyens ; Amérique latine ; big data

1 Introduction

Several countries of Latin America and the Caribbean (LAC) have been experiencing a growing sense of dissatisfaction of citizens with their governments. The share of the population with little or no trust in governments reached 74% in 2018, 20 percentage points higher than in 2010. Many citizens believe that governments lack integrity and that their actions are driven by motivations that deviate public resources from the public interest. In 2019, 53% of the population in LAC believed that corruption had increased and 65% thought that their government was run by and for a few private interests, with leaders acting in their own self-interest and at the expense of the citizens they serve (Transparency International, 2019[1]). Satisfaction with public services has also deteriorated (OECD, 2020[2]). Over 2006-18, the share of the population satisfied with the quality of healthcare services fell from 57% to 42%, well below levels in the OECD – stable at around 70%. The deterioration in citizens' perceptions of institutions and public services is a reflection of citizens' growing discontent, which translated into a wave of social protests across some Latin American countries in the second half of 2019 (OECD/CAF/UN ECLAC, 2018[3]; OECD et al., 2019[4]).

The causes behind these trends are difficult to disentangle. In a period in which overall standards of living have been rising across most of the LAC region, it seems paradoxical that trust in institutions has experienced erosion, or at least has not evolved accordingly. The expansion of the middle class appears as one of the core explanations (OECD/CAF/UN ECLAC, 2018[3]). Today, a third of the population in LAC belongs to this group. The middle class has higher and changing aspirations, and hence demands institutions that are ready to respond effectively to them.

Several megatrends are also shaping higher social demands. Digital technologies favour comparisons between citizens in areas with different levels of development, setting higher aspirations for better standards of living. Evidence has shown that the expansion of mobile Internet networks in the last decade has led to a significant reduction in government approval since it helps to expose the incidents of actual corruption to the public (Guriev, Melnikov and Zhuravskaya, 2019[5]). Moreover, the recent COVID-19 crisis is bringing additional challenges and most likely a rise in the demands of citizens for stronger public institutions and better quality public services.

The growing divide between citizens and institutions is one of the most pressing development challenges in the LAC region. It erodes the foundations of the social contract and fuels a vicious circle that operates as an "institutional trap": As citizens' trust and satisfaction declines, they gradually see less value in engaging in social duties such as paying taxes (i.e. what is commonly referred to as "tax morale"). The percentage of the population that found tax evasion justifiable in LAC increased from 46% in 2011 to 53.4% in 2016 (OECD/CAF/UN ECLAC, 2018[3]; OECD et al., 2019[4]). Tax revenues are thus negatively affected, limiting available resources for public institutions to provide better quality goods and services, and thus to respond to the rising aspirations of society. This vicious circle jeopardises the social contract in the region and generates a context of polarisation that can provide fertile ground to the emergence of populist political alternatives.

Indeed, LAC countries have embarked on substantial efforts to strengthen and modernise institutions to address this growing divide, though there are still significant weaknesses. In particular, integrity and anti-corruption have become a priority in the policy debate and during past elections (OECD, 2018[6]; OECD, 2019[7]). Digital transformation offers alternative solutions to address social challenges, and citizens demand that institutions adopt and adapt to these new possibilities.

The relevance of this topic from a public policy perspective is decisive. However, evidence about the drivers of social discontent is neither abundant nor conclusive, perhaps due to its complex, multifaceted and rapidly changing nature. This paper contributes to the literature analysing the drivers of citizen satisfaction with governments, with a focus on LAC, and to the literature exploring the potential of non-traditional data sources.

The paper is organised as follows: Section 2 briefly reviews the literature on the determinants of government and presidential approval rates over time, as well as the increasing literature on the use of Internet search query data and other Big Data sources to analyse social phenomena. Section 3 describes the data used and presents some stylised facts. Section 4 presents the empirical strategy of the analysis. Section 5 provides the results and robustness checks. Lastly, Section 6 concludes and outlines some policy implications.

2 Literature review

This paper contributes to two main strands of the literature. The first studies the potential drivers of presidential approval rates over time. So far, this strand has focused predominantly on trends within individual administrations and countries, stressing the effect of factors such as the honeymoon period or of economic performance on approval rates. The second is the growing literature that explores the use of new data sources, such as Internet search query data, to analyse socioeconomic phenomena.

Determinants of governments' approval and citizens' trust

On the first strand, several authors have investigated the determinants of governments' approval using different data sources and focusing on different countries and regions. The work of John Mueller has served as the baseline for subsequent empirical studies on governments' popularity. Mueller (1970[8]; 1973[9]) used the Gallup aggregate approval rating as the dependent variable for several regression analyses examining the determinants of presidential approval in the United States between 1945 and 1969. Mueller's findings suggest that presidential approval tends to decline systematically over the term of office and differs significantly between administrations. Furthermore, he found evidence of an asymmetric effect of the country's economic performance on the presidential approval rating: while a slow economy weakens presidential popularity, dynamic economic performance tends to leave popularity unaltered (Mueller, 1970[8]; Dickerson, 2016[10]).

The literature on this subject has advanced along with new academic streams, sources of data and empirical methods, by taking better account of political, social and psychological realities. Still, most of the literature has focused on the influence of economic variables on presidential popularity, primarily inflation and unemployment, due to the high frequency and quality of the data available for these two variables. Authors have often considered other crucial issues, including scandals (Smyth and Taylor, 2003[11]), crises (Newman and Forcehimes, 2010[12]), violence (Romero, Magaloni and Díaz-Cayeros, 2016[13]), wars (Geys, 2010[14]), populism (Algan et al., 2017[15]), political polarisation (Donovan et al., 2019[16]), and taxation (Geys and Vermeir, 2008[17]), among others. Often, they have also disentangled the "honeymoon" effect and the "rally around the flag" effect (Newman and Forcehimes, 2010[12]) on presidential approval.

Conversely, authors rarely include subjective factors in their analysis. Limitations in terms of data availability, frequency and comparability are possibly the main reason for the paucity of studies including subjective and perception explanatory variables. In the broader field of electoral studies, Ward (2015[18]) analysed how the popularity of governing parties is associated not only with the performance of the economy but also with the electorate's broader well-being. He found that a country's aggregate level of subjective well-being can account for more of the variance in government vote share than traditional macroeconomic variables.

After decades of empirical research on the drivers of presidential approval after Mueller's work, conclusions remain unclear. Berlemann and Enkelmann (2014[19]) reviewed and analysed 57 different empirical studies of the presidential popularity function in the US, published between 1970 and 2010. They found results are inconsistent and differ depending on the country or region analysed, as well as on the sample, period, data frequency and empirical techniques used for the analysis. In the case of the United States, for example, only about half of the studies found unemployment and inflation to have a significant effect on presidential

popularity. Berlemann and Enkelmann (2014[19]) also found that the length of the sample (in terms of time) is of crucial importance. While in the long run unemployment, inflation and budget deficit have a robust effect on presidential approval, this does not hold for shorter sub-periods.

On Latin America and the Caribbean (LAC), Carlin, Love and Martínez-Gallardo (2014[20]) analysed the relationship between presidential approval, economic performance and scandals. Faced with the obstacle of obtaining comparable cross-national approval survey data, they used the *Executive Approval Dataset* that includes measures of scandals, presidential approval and the economy for 84 presidential administrations in 18 LAC countries. They found that scandals only appear to damage presidential approval when inflation and unemployment are high. Using the same database, Carlin et al. (2018[21]) found evidence confirming the cyclicality over presidents' terms in the LAC region: high post-election approval rates (honeymoon effect) decrease over their mandate and experience an end-of-term boost when new elections approach. Murillo and Visconti (2017[22]) found evidence on the negative consequences of adverse economic conditions on the capacity of incumbents to sustain their electoral support in LAC. They tested this hypothesis at both the aggregate and individual level. Using electoral and survey data, they found evidence that adverse economic shocks erode support for incumbents at both levels of analysis.

Other authors have also conducted country-level analyses in the LAC region. Arce (2003[23]) analysed the impact of political violence on presidential approval in Peru, using monthly presidential approval data for the period 1985-97 for two presidencies. While controlling for variables frequently used in previous literature, his results suggest that higher levels of political violence hurt left-leaning governments, but not necessarily right-leaning governments. He concluded that it is likely that voters expect right-leaning governments to deal better with political violence in general and thus are supportive of their efforts. Romero, Magaloni and Díaz-Cayeros (2016[13]) analyse the impact that citizens' assessments of security issues have on presidential approval in Mexico. They found evidence suggesting that security issues matter more than partisanship or even the economy when determining citizens' presidential approval.

Although citizens' trust in government primarily reflects their approval of their country's leadership (OECD, 2015[24]), the concept of public trust is certainly much broader than presidential approval.[1] The OECD has proposed an analytical framework on trust, beyond presidential approval. This framework identifies two fundamental building blocks of citizens' trust in public institutions: competence and values (OECD, 2017[25]). Competence is the ability of governments to deliver to citizens the services they need, at the quality level they expect. It involves providing public services (responsiveness) and protecting citizens (reliability). Values refer to the principles that inform and guide government action. Values are reflected when governments use power and public resources ethically (integrity); inform, consult and listen to citizens (openness); and improve socio-economic conditions for all (fairness). Under this framework, therefore, citizens evaluate governments based on their experience of service delivery but also based on the efficacy and fairness of the policy-making process and its outcomes (OECD, 2017[25]; 2017[26]).

Although public trust is one of the foundations upon which political systems build legitimacy and sustainability, it is poorly understood and is not consistently measured in most countries (González and Smith, 2017[27]). Recently, the OECD Trustlab experiment (Murtin et al., 2018[28]) used data from six OECD countries, between November 2016 and November 2017, to analyse citizens' trust in governments. Their analysis supports the view that the most critical determinants of self-reported trust in government are the perception of high-level corruption, followed by the perception of government reliability, and government responsiveness. When it comes to public services, satisfaction with security and crime prevention services and with education and healthcare systems and provision of welfare benefits appear to be significant determinants of trust.

Although the OECD framework on trust contributes to critical elements in our analysis, it is essential to clarify that this paper focuses solely on presidential approval and not on the broader concepts of trust in governments and institutions.

Non-traditional data sources to analyse socio-economic phenomena

The second strand of literature studies the use of non-traditional data sources, such as Internet search query data (e.g. *Google Trends*) or social media data (e.g. Facebook, LinkedIn)*,* for academic research as well as to inform public policy making. Several authors have tested the accuracy of Google data to proxy social and economic variables, finding statistically significant associations between Google searches and other data (Ginsberg et al., 2009[29]; Seifter et al., 2010[30]; Choi and Varian, 2012[31]; Scheitle, 2011[32]). *Google Trends* variables can mirror traditional variables while providing a higher level of frequency. As shown in Figure 1, there are variables for which the actual trend and that of related Internet searches move very similarly. This is the case for the actual unemployment rate and searches for job-related issues in Colombia and Peru, for example. Regarding social media data, *LinkedIn* has become a powerful data source to analyse the pathways of graduates as they transition from education to employment and explore the relationships between their skills and qualifications and how they navigate the labour market (OECD, 2019[33]). Moreover, during the COVID-19 crisis, *Facebook* data has provided a unique window into the experiences of small and medium-sized businesses, since it allows to measure the pandemic's effects on business closures, revenue, employment and finances (Facebook, OECD and WB, 2020[34]).

Figure 1. Unemployment rate and job-wise Internet searches, 2006-15
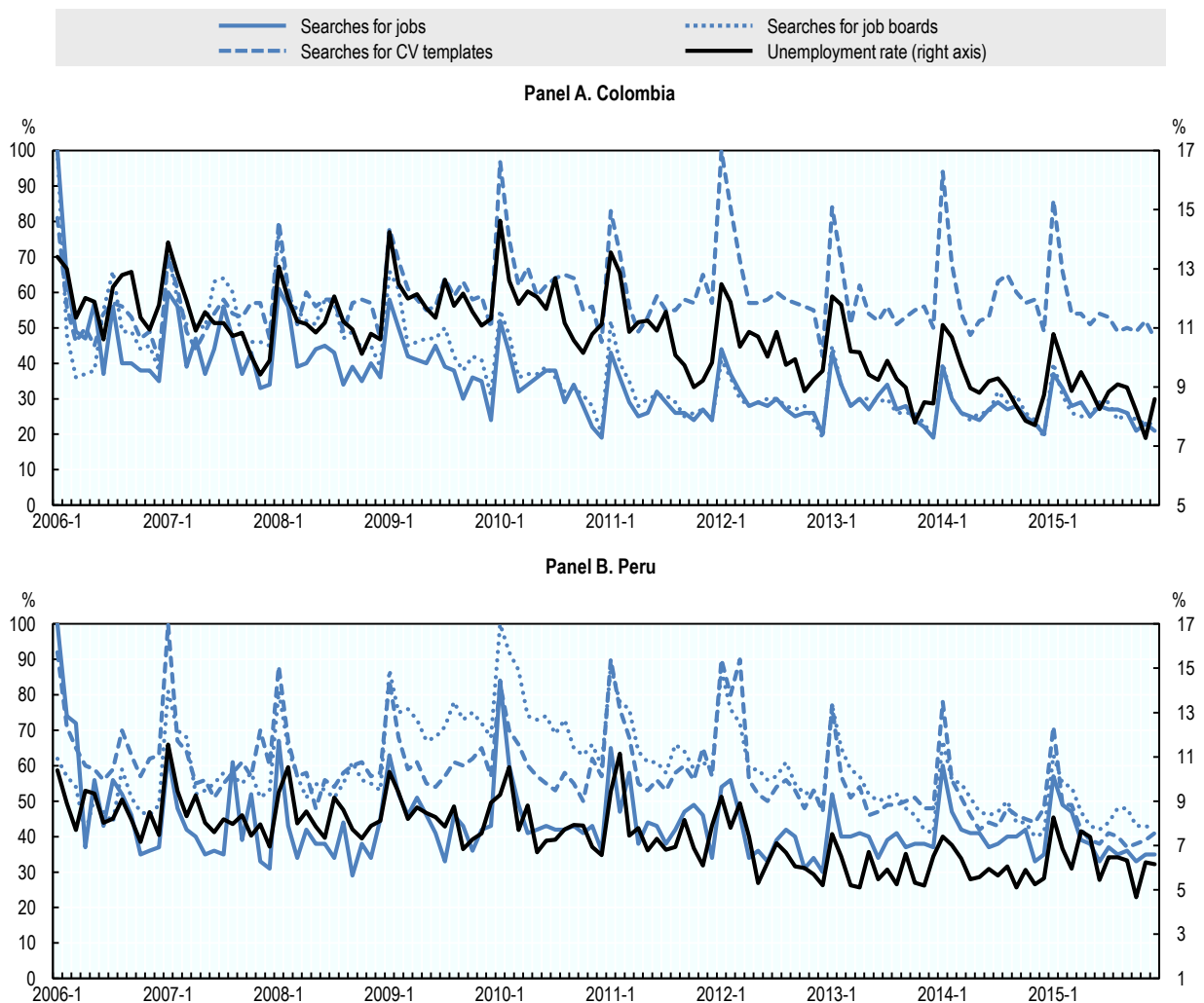
Source: Authors' calculations based on *Google Trends* and national data for the unemployment rate (DANE and INEI, National Statistics Offices for Colombia and Peru, respectively).

The use of *Google Trends* as a source of data has significantly evolved since its launch in 2006. Jun, Sun Yoo and Choi (2018[35]) analysed the evolution of research publications using *Google Trends* and found that *Google Trends* series are being used to analyse topics in diverse fields, including IT, communications, medicine, business and economics. The focus of research has shifted from merely describing and diagnosing trends to forecasting changes. Choi and Varian (2012[31]) showed that Internet searches had been used to forecast near-term values of economic indicators, such as automobile sales, unemployment claims, travel destination planning and consumer confidence. These searches are thus helpful for short-term economic prediction. The use of *Google Trends* has been gradually linked to other big data sources to overcome the limitations of using only research information, according to Jun, Sun Yoo and Choi (2018[35]). They also point to a critical discussion on the nature of big data as a public good.

Algan et al. (2019[36]) built an indicator of individual subjective well-being in the United States based on *Google Trends*. The indicator is a combination of keyword groups that are identified to fit with the subjective well-being measures disseminated by Gallup Analytics. They found that keywords associated with job search, financial security, family life and leisure are the strongest predictors of the variations in subjective well-being.

Several authors argue that Google searches can provide more reliable data on socially sensitive topics compared to traditional surveys. The absolute anonymity of these searches, and therefore the sincerity in their content, make the data generated more likely to be objective than traditional surveys (Stephens-Davidowitz, 2013[37]). These reasons explain why social scientists are increasingly using these data sources to know more about people's behaviours, and why they could potentially be used to inform policy makers and governments on various aspects of citizens' and public opinion (Reimsbach-Kounatze, 2015[38]). The use of big data for researching in different fields holds great potential, although interpretations may be cautious, considering the limitations and pitfalls of this data source.

Big data analytics is emerging with strength as a discipline and offers plenty of new possibilities to inform policy making and to transform the public sector. Similarly, Artificial Intelligence brings new unexplored possibilities to improve policy making (Pencheva, Esteve and Mikhaylov, 2018[39]). Höchtl, Parycek and Schöllhammmer (2016[40]) point to the fact that big data analytics can significantly transform policy making by replacing the traditional model of sequential execution of the individual stages of the policy cycle with a model of continuous evaluation. In the traditional policy cycle, the opportunity to adjust a policy only appears after the evaluation of results, while big data analytics can provide new, valuable information at every stage of the cycle, thereby shortening the decision-making process.

3 Data and stylised facts

In this analysis, citizens' satisfaction with governments is proxied by using presidential approval rates. On the left side of the equation, we use the *Executive Approval Index* (v1.0, Carlin et al. (2016[41])) as a proxy dependent variable. On the right-hand side, we include two types of independent explanatory variables. First, we include traditional socioeconomic explanatory variables, extracted from traditional data sources. This set of variables captures the influence of economic performance and other relevant phenomena on governments' approval. Second, we introduce subjective explanatory variables in the equation, built using Internet search query data, in particular with series from *Google Trends*.

The sample used covers ten years, from January 2006 to December 2015, with monthly frequency data of 18 LAC countries.[2] All countries included in the analysis have presidential systems. Therefore, in these countries, the president is both head of state and head of government (i.e. executive branch), and is elected and remains in office independently of the legislature. For that reason, the concepts of executive approval and presidential approval are used interchangeably.

Endogenous variable: The Executive Approval Index

The literature analysing the determinants of governments' approval is rather scarce in Latin America. The few published studies often focus on the analysis of a single country, perhaps owing to the scarcity of comparable cross-national and high-frequency data sources on executive approval. Indeed, each country has its survey sources and methodology, which do not always have the same frequency or the same specifications. Some databases provide cross-country data on this matter, but only with annual or quarterly frequency.

The *Executive Approval Database* (Carlin et al., 2016[41]) provides monthly executive approval series that combine the survey marginals of presidential popularity into a single smoothed series using Stimson's (2018[42]) dyads-ratio algorithm and WCalc6 software (Stimson, 2015[43]) for 18 Latin American countries (see Figure A A.1 in Annex A). This database combines survey data from major polling organisations (i.e. Gallup, LAPOP, *Latinobarómetro*, and IPSOS), as well as local polling firms from each country (e.g. *Borge y Asociados*, *Cifras y Conceptos*, *Ibarómetro*, *Parametría*, among others [see Table A A.1 in Annex A]).

Table 1 presents the main descriptive statistics at the country level. We use the approval rates from this database (Carlin et al., 2016[41]) as our dependent variable (i.e. the percentage of positive ratings for the current administration).

Table 1. Descriptive statistics of the *Executive Approval Index*, by country

Between January 2006 and December 2015

Country	Mean	St. Dev	Median
All	50.26	16.783	50.34
Argentina	49.10	10.109	49.52
Bolivia	54.88	8.753	56.07
Brazil	52.38	16.187	55.02
Chile	35.46	7.361	33.19
Colombia	51.38	11.651	50.26
Costa Rica	33.23	9.682	31.87
Ecuador	73.01	7.683	71.09
El Salvador	67.15	10.985	71.46
Guatemala	56.43	13.712	58.34
Honduras	37.95	12.795	35.45
Mexico	51.73	5.826	51.77
Nicaragua	63.25	6.586	64.02
Panama	50.38	10.969	52.10
Paraguay	36.37	14.071	39.33
Peru	31.48	13.190	26.63
Dominican Republic	68.35	19.787	79.51
Uruguay	55.12	6.991	54.68
Venezuela	36.18	7.079	37.97

Note: Satisfaction ranges from 0 (the worse score) to 100 (the best score), which refers to complete satisfaction.
Source: Carlin et al. (2016[41]), Executive Approval Database v1.0.

Traditional explanatory variables

Traditionally, most of the literature investigating the determinants of presidential approval has included inflation and unemployment as the two main explanatory variables. Although the relationship with these two variables is rather intuitive, one of the main reasons for their inclusion is the availability of reliable and high-frequency historical data for most countries. Less frequently, authors have included other economic variables, such as economic growth, the performance of financial markets, the tax burden or the public fiscal deficit (Berlemann and Enkelmann, 2014[19]).

To grasp the economic context in our analysis, we use a set of national and international economic variables. To capture the effects of price fluctuations in the economy, we included the monthly inflation rate, obtained from the national statistical offices (or the national institution in charge). We also included quarterly unemployment and real GDP growth rates, obtained from local sources and Datastream Infobase (Thomson Reuters, 2012[44]). To control for the economic cycle, we included the output gap, calculated as a deviation from GDP trend using the Hodrick-Prescott filter, with a lambda of 1600, on the constant quarterly GDP per country.

Another frequently used explanatory variable is the honeymoon effect. Evidence has shown that elected politicians have a period of high popularity in the early months of their mandate. Therefore, we included a set of eight quarterly dummy variables to capture the effects of the honeymoon period over the first two years after taking office. Including a set of eight dummy variables will allow capturing the potential decreasing nature of this effect over time.

Subjective explanatory variables: Internet search query data

While socioeconomic variables may contribute to determining executive approval, subjective and perception factors can also affect the popularity of governments. These factors have rarely been included in past literature because of the lack of comparable data on these topics.

In this analysis, the subjective determinants of governments' approval were chosen based on citizens' main concerns. In particular, we used *Latinobarómetro* opinion polls, which provide information on what citizens identify as the single, most significant problem of their country.[3] The polls suggest that the ten most prominent problems in Latin America, between 2006 and 2015, were insecurity, unemployment, economic environment, poverty, corruption, political environment, inflation, education, health and violence (see Figure 2).

Figure 2. Perception of the country's main problem in Latin American countries
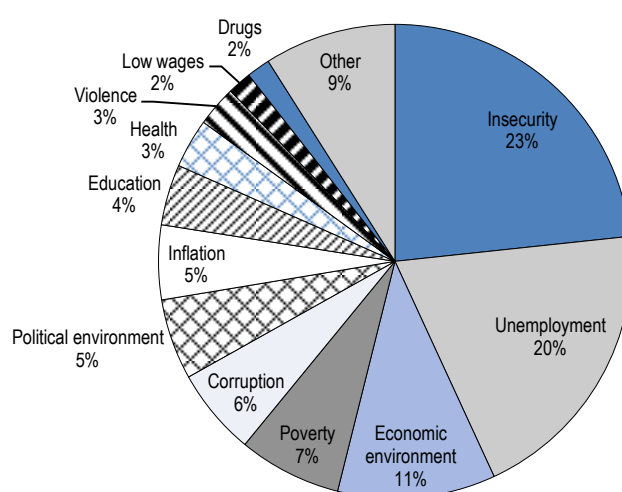
Source: Latinobarómetro database (2006, 2007, 2008, 2009, 2010, 2013, 2015).

In terms of data sources, this second set of explanatory variables was built using Internet search query data, in particular, *Google Trends* series. Although traditional survey measures exist for some of the chosen subjective determinants of presidential approval, they often do not comply with the cross-country comparability and monthly time-frequency required to support our empirical analysis. Instead, *Google Trends* data offers extensive coverage of countries across time, which allow us to capture the Latin Americans' perception of a wide array of issues. Thanks to its anonymity, this data is not subject to biases that traditional perception surveys present. In this paper, we used series for the searches for corruption, insecurity, jobs and complaints (see Annex B). These series were used as proxies for the perception of the main socioeconomic and institutional concerns in Latin America.

Google Trends series do not measure the number of searches for a given item, but rather its relative popularity in a given period within a specific area. Scores range between 0 and 100, where 100 is granted to the period of time with the highest number of relative searches for a specific word or topic (see Annex B). By plotting the chosen series, the data displays the expected negative relationship between governments' approval and searches for corruption, jobs, insecurity and complaints in the region. The higher the searches on these topics – which we assume reflect concern about them as they affect citizens' well-being – the lower the governments' approval (see Figure 3).

Figure 3. Governments' approval and relative popularity of searches for corruption, jobs, insecurity and complaints in LAC

2006-15, monthly variables
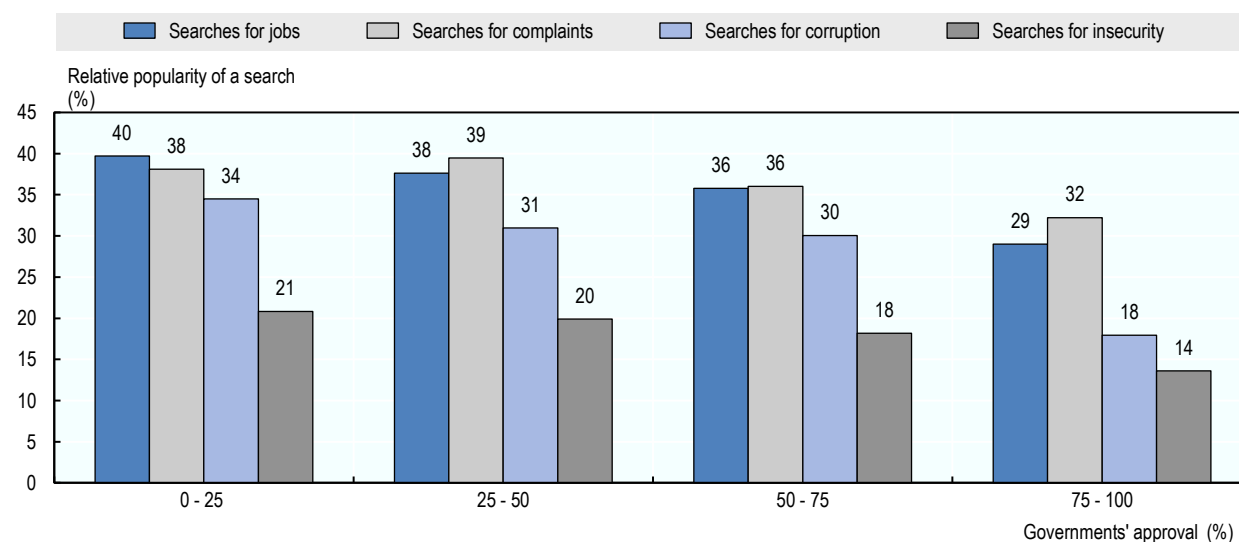
Note: In this graph, the total number of observations from our sample were divided in four groups using the quartiles of the dependent variable (executive approval). The bars represent the average value of each variable within each of these four groups.
Source: Authors' calculations based on *Executive Approval Database* (Carlin et al., 2016[41]) and *Google Trends* database.

Using the information obtained from Internet search engines presents some limitations that are worth signalling. First, the sample is unlikely to be representative of the population of a country, considering that Internet access is constrained for specific segments of the population. This is particularly applicable to the LAC region, given the low levels of Internet penetration (only 68% of the population in 2018 compared to 84% in the OECD economies) despite its progress over the last ten years. This means that middle- and high-income inhabitants are behind the sample of people performing Internet searches. However, these individuals are the ones that demand better governments and public services and are actively participating in the democracies of their countries (OECD et al., 2019[4]).

We also face the latent risk of capturing searches that are irrelevant to our analysis. Therefore, the *Google Trends* search series were refined as much as possible to reduce this risk, by excluding potentially misleading words to avoid biases generated by unrelated topics captured in the searches (see Annex B). Additionally, there are other characteristics of the *Google Trends* series, which were detailed by Algan et al. (2016[45]), such as spikes, cliffs, discontinuities, and time trends, that can affect the overall quality of the data negatively.

Table 2 presents the main descriptive statistics and the sources used for the period from January 2006 to December 2015. It presents simple averages for each of the variables included in the analysis, with searches for jobs and complaints being, on average, the most common in *Google Trends*. Searches for corruption or insecurity have a relatively lower average.

Table 2. Descriptive statistics of the dependent and independent variables

2006-15, monthly variables

Variable	Obs.	Mean	Std. Dev.	Min.	Max.	Source
Executive approval (%)	2 156	51.15	15.78	10.437	91.33	*Executive Approval Database*
Searches for corruption (0-100)	2 160	17.46	15.21	0	100	*Google Trends*
Searches for jobs (0-100)	2 160	36.32	21.31	0	100	*Google Trends*
Searches for insecurity (0-100)	2 160	18.74	16.35	0	100	*Google Trends*
Searches for complaints (0-100)	2 160	37.84	24.58	0	100	*Google Trends*
Annual Inflation rate (%)	2 125	8.36	11.87	-3.37	180.86	Nat. Stat. Offices and Datastream
Annual GDP growth (%)	2 115	4.55	3.91	-11.27	17.49	Nat. Stat. Offices and Datastream
Unemployment rate (%)	2 155	6.88	2.92	1.8	16.4	Nat. Stat. Offices and Datastream
Output gap	2 127	0.01	0.12	-0.82	0.66	Own calculations based on IMF World Economic Outlook Database

4 Hypotheses and empirical strategy

We used fixed-effects panel data analysis to address the potential endogeneity produced by time-constant omitted variables. However, panel data is likely to exhibit various issues with cross-sectional and temporal dependencies (Hoechle, 2007[46]). In fact, the serial (temporal) correlation has been a recurrent issue in the existing economic literature on presidential approval due to the usage of high-frequency time series in the analyses (Berlemann and Enkelmann, 2014[19]). Serially correlated error terms artificially inflate the *t-statistic*, showing statistical significance when there is none. As shown by Barlemamn and Enkelmann (2014[19]), many previous studies faced issues with temporal correlation given that most of the existing literature has been focused on a single country. Considering that we have panel data for 18 countries, our analysis faces the additional threat of cross-sectional (spatial) correlation. Moreover, heteroscedasticity stands out as another potential and recurrent misleading issue, given the characteristics of the data. Several tests[4] applied to the sample confirmed the presence of temporal and spatial correlation in the data, as well as of heteroscedasticity in the error distribution.

We solve these issues by using a fixed-effects regression with Driscoll and Kraay standard errors. We employ an extension of the conventional nonparametric covariance matrix estimation techniques, which yields standard error estimates that are robust to very general forms of spatial and temporal dependence, as the time dimension becomes large (Driscoll and Kraay, 1998[47]). Driscoll and Kraay's estimation is a simple adaptation of the standard heteroscedasticity and autocorrelation consistent covariance matrix estimation techniques, such as those in Newey and West (1987[48]) or Andrews (1991[49]), with the advantage of being consistent in the presence of spatial correlation.

We use the executive approval index as our dependent variable, which contains comparable data on the monthly popularity rates of presidents in LAC. In the first specification, the equation includes only the traditional economic explanatory variables that are used in most of the literature about the determinants of presidential approval. Specifically, we included the inflation rate, economic growth, unemployment rate, and the output gap (economic cycle). We also included a set of eight quarterly dummy variables (i.e. 1-3, 4-6, 7-9, 10-12, 13-15, 16-18, 19-21, 22-24 months after taking office) to capture the so-called honeymoon effect, as follows:

$$Executive\ Approval_{i,t} = \alpha + \beta_1\ Inflation_{i,t} + \beta_2\ GDP\ growth_{i,t} + \beta_3\ Unemployment_{i,t} +$$
$$\beta_4\ Output\ gap_{i,t} + \sum_{j=9}^{16}\beta_j\ Honeymoon_{q,i} + \delta_t + \varphi_i + \varepsilon_{i,t}\quad(1)$$

In addition to the traditional variables included in the first specification, the second specification includes the set of subjective explanatory variables (Internet search query data). With this extension, we aim to capture the effects of the subjective determinants of presidential approval (in particular, perception of corruption, of unemployment, of insecurity and consumer unrest). The second specification is as follows:

$$Executive\ Approval_{i,t} = \alpha + \sum_{j=1}^{4}\beta_j X_{j,i,t} + \beta_5\ Searches\ for\ corruption_{i,t} +$$
$$\beta_6\ Searches\ for\ jobs_{i,t} + \beta_7\ Searches\ for\ insecurity_{i,t} + \beta_8\ Searches\ for\ complaints_{i,t} +$$
$$\sum_{j=9}^{16}\beta_j\ Honeymoon_{q,i} + \delta_t + \varphi_i + \varepsilon_{i,t}\quad(2)$$

Where $\sum \beta_j X_{j,i,t}$ represents the set of economic variables included in the first specification. In both specifications, δ_t and φ_i represent sets of dummy variables to capture the time and the country fixed effects respectively, and $\varepsilon_{i,t}$ stands for the error term.

The first hypothesis links the honeymoon effect with executive approval:

> **_Hypothesis 1:_** _Politicians have a period of high popularity in the early months of their mandate._

The first hypothesis is validated when $\sum_{j=9}^{16} \beta_j\, Honeymoon_{q,i}$ are positive and statistically significant in both specifications (equations 1 and 2). By introducing a set of eight quarterly dummy variables to capture the honeymoon effect, results enable analysing the decreasing nature of the honeymoon effect over the two years after taking office.

Furthermore, we build on the belief that the heterogeneity among the results and conclusions of previous economic research on presidential approval may be associated with the omission of relevant subjective explanatory variables. Therefore, with our second hypothesis, we expect that the introduction of subjective determinants to the empirical analysis should provide more consistent and accurate results:

> **_Hypothesis 2:_** _Citizens' perceptions regarding socioeconomic and institutional aspects of their country are strongly associated with executive approval after controlling for traditional economic explanatory variables._

The second hypothesis is validated when β_5, β_6, β_7 and β_8 (searches for corruption, jobs, insecurity and complaints respectively) are negative and statistically significant in the second specification.

5 Results and robustness checks

Table 3 reports the results of the fixed-effects regression with Driscoll and Kraay standard errors estimations.

Table 3. Main regressions

Executive approval	(1)	(2)
Constant	41.146 (4.293)***	45.908 (4.557)***
Searches for corruption		-0.099 (0.028)***
Searches for jobs		0.019 (-0.019)
Searches for insecurity		-0.045 (0.018)**
Searches for complaints		-0.088 (0.025)***
Inflation	-0.231 (0.109)**	-0.221 (0.105)**
GDP growth	0.688 (0.239)**	0.635 (0.232)**
Unemployment	0.259 (-0.357)	0.023 (-0.348)
Output gap (economic cycle)	-2.334 (-3.667)	-1.954 (-3.696)
Honeymoon effect — First trimester (Month 1-3)	8.343 (1.650)***	8.063 (1.601)***
Second trimester (Month 4-6)	7.65 (1.654)***	7.489 (1.574)***
Third trimester (Month 7-9)	7.223 (1.815)***	7.191 (1.754)***
Fourth Trimester (Month 10-12)	5.078 (1.639)***	5.415 (1.580)***
Fifth trimester (Month 13-15)	4.569 (1.764)**	4.722 (1.746)**
Sixth trimester (Month 16-18)	3.51 (-2.105)	3.785 (2.050)*
Seventh trimester (Month 19-21)	3.888 (1.993)*	3.902 (1.927)*
Eighth trimester (Month 22-24)	3.667 (1.906)*	3.662 (1.824)*
Within R2	0.1816	0.2041
N	2 059	2 059
Time fixed effects	x	x
Presidential dummies	-	-

Note: Fixed-effects regression with Driscoll and Kraay standard errors. Dependent variable: Executive approval. Standard errors in parentheses. *, **, *** indicate significance at the 90%, 95% and 99% level, respectively. The within R-squared of a panel regression shows how much of the variation in the dependent variable within each country is captured by the model.

Column 1 presents results from the first regression, including traditional economic and political explanatory variables. Results suggest that selected economic variables and the political-time effect have an impact on presidential approval. In particular, the inflation rate, economic growth, and the so-called honeymoon effect have a statistically significant association with presidential approval rates in Latin America. The coefficients of the dummy variables confirm the decreasing nature of the honeymoon effect after taking office. In contrast, the unemployment rate and the economic cycle variable (output gap) do not exhibit a statistically significant correlation with the dependent variable.

Column 2 adds the set of subjective explanatory variables to the previous regression. Under this new specification, the positive and decreasing relationship between honeymoon effect and executive approval remains statistically significant at the 1% level from the first to the fourth trimester, significant at the 5% level in the fifth trimester, and weakly significant afterwards. Therefore, hypothesis 1 is validated. Besides, results show that Google searches for corruption, insecurity, and consumer unrest have a negative and statistically significant relationship with executive approval. Therefore, the higher the number of searches for corruption, insecurity and complaints, the lower the level of presidential approval. Hypothesis 2 holds for these three variables. Conversely, searches for jobs do not have a statistically significant correlation with the executive approval rates of the region. A possible explanation for this could be the effect of large labour informality in the region. This could imply that most job searches among most vulnerable groups (i.e. those whose jobs are informal and are at more risk when there is economic instability and political uncertainty) do not take place through Internet searches, but rather via informal channels such as friends, family or past jobs.

Variance decomposition

For investigating more in-depth into the results of our analysis, we included a Shorrocks-Shapely decomposition of the within R-squared obtained from a simplified version[5] of the second specification. This type of analysis provides an additive decomposition of the statistic, allowing us to see the relative contribution of each regressor. According to the R-squared decomposition, traditional explanatory variables are the main drivers of variation, followed by the political-time effect, and the subjective explanatory variables. Table 4 reports the detailed results of this analysis.

Table 4. Decomposition of the within R-squared

	Shapley value (estimate)	Per cent (estimate)
Subjective explanatory variables	**0.02802**	**28.42%**
Searches for corruption	0.00881	8.94%
Searches for jobs	0.00185	1.87%
Searches for insecurity	0.00221	2.24%
Searches for complaints	0.01515	15.37%
Traditional explanatory variables	**0.04006**	**40.65%**
Inflation	0.02839	28.81%
GDP growth	0.00683	6.93%
Unemployment	0.00403	4.09%
Output gap (economic cycle)	0.00081	0.82%
Honeymoon effect	**0.03045**	**30.89%**
Within R-squared (Baseline model without time fixed effects)	**0.09856**	**100%**

Note: The Shorrocks-Shapely decomposition of the within R-squared has a maximum of 20 variables to be included in the analysis; therefore, this decomposition analysis does not include the time fixed effects in the analysis, which results in a lower within R-squared. Also, we only included one honeymoon dummy variable for the first six months after taking office.

Robustness checks

Table 5 presents some robustness checks made to the second specification to test both hypotheses. The estimations of the set of honeymoon dummy variables maintain their significance and the coefficients their decreasing nature over time in all the estimations presented below. All models include country and time fixed effects as included in the two main regressions.

Column 1 of Table 5 presents an alternative specification, where the set of subjective explanatory variables was replaced with a set of index variables, constructed based on a more extensive selection of search terms. We constructed an index for corruption-related searches, which includes searches for corruption and bribes. Similarly, for the case of the unemployment-related search index, we have included searches for CV templates, job boards and jobs. For the insecurity-related search index, we included the searches for insecurity, delinquency, thefts, kidnapping, murdering and drug trafficking. The results of the regression analysis show that, when these indexes are introduced to replace the single search variables, results presented in Column 2 of Table 3 remain consistent. In addition, in the case of the insecurity variable, the series constructed from searches for several terms (index) seems to offer more robust data in this analysis than the simple variable of searches for insecurity.

Column 2 of Table 5 presents the results of a specification that includes a measure of the quality of government (Teorell et al., 2019[50]), in order to capture the structure and behaviour of public administrations. This variable, scaled from 0 to 1, has an annual frequency and condenses the information on corruption, law and order, and quality of bureaucracy. Results show that the quality of government is not significant, which can be partly explained by the annual periodicity of the variable. However, there is no monthly or quarterly data available regarding experts' opinions on the quality of governments. Other results remain consistent with the two main regressions.

In order to capture the effects of the inherent characteristics of each governor on the approval rates, several authors have included a set of "presidential dummies" to their analysis on presidential approval (Berlemann and Enkelmann, 2014[19]). We have tested the robustness of our results by including a set of presidential dummy variables in the regression (Column 3 of Table 5). These dummy variables are intended to capture characteristics inherent to each president that can impact their approval among citizens, such as their charisma and assertiveness, among others. Results remain consistent with the baseline model exhibited in Column 2 of Table 3. This model also yields a considerably higher within R-squared, which suggests the personality and other characteristics inherent to each president can explain a significant share of the variation of governments' approval rates.

In order to tackle a potential multicollinearity issue in the analysis, we checked the robustness of the results after dropping either unemployment or perception of unemployment (searches for jobs). Results are shown in Columns 4 and 5 of Table 5, respectively. For both cases, results are consistent with the primary regression analysis shown in Column 2 of Table 3.

We have also considered an alternative analysis in order to test the consistency of the results when observations are more recent. To do so, we have dropped the oldest half of our sample (i.e. before January 2011). One of the benefits of using only data that is more recent is having a more representative sample since the penetration of the Internet has been rising over the last decades in LAC. Therefore, searches on the Internet may be reflecting the concerns of an increasingly wider share of the population. Results are shown in Column 6 of Table 5. They remain consistent for the baseline model presented in Column 2 of Table 3, except for GDP growth, which losses significance.

We have tested two additional specifications by introducing lags of one ($t-1$, Column 7) and two months ($t-2$, Column 8) to all explanatory variables (see Table 5). The level of significance of perception of corruption decreases in both cases compared to the original models. In contrast, the significance of GDP growth seems to be more robust in these two alternative specifications. These results may suggest that macroeconomic variables are likely to have a lagged impact on citizens' sentiments since they measure

actual economic performance and not citizens' perception of the economy in t. Moreover, to further test the counter-intuitive results about unemployment and perception of unemployment, we have gradually introduced up to ten lags to both variables. Results remain consistent compared to the hypotheses tested in Column 2 of Table 3.

Table 5. Robustness checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7) $(t-1)$	(8) $(t-2)$
Constant	45.323 (5.066)***	44.368 (8.638)***	32.892 (4.018)***	45.839 (3.044)***	46.772 (4.305)***	42.499 (6.808)***	47.379 (3.055)***	47.042 (2.885)***
Searches for corruption		-0.099 (0.028)***	-0.077 (0.023)***	-0.096 (0.027)***	-0.099 (0.028)***	-0.246 (0.052)***	-0.079 (0.028)**	-0.049 (0.025)*
Searches for jobs		0.019 (-0.019)	0.029 -0.02	0.018 -0.019		0.038 (-0.043)	0.02 (-0.02)	0.004 (-0.019)
Searches for insecurity		-0.045 (0.018)**	-0.037 (0.017)**	-0.045 (0.018)**	-0.046 (0.018)**	-0.073 (0.038)*	-0.055 (0.021)**	-0.058 (0.020)**
Searches for complaints	-0.079 (0.026)***	-0.088 (0.025)***	-0.032 -0.027	-0.087 (0.025)***	-0.088 (0.025)***	-0.158 (0.041)***	-0.095 (0.025)***	-0.099 (0.024)***
Searches for corruption (index)	-0.14 (0.048)***							
Searches for jobs (index)	0.085 (-0.055)							
Searches for insecurity (index)	-0.145 (0.038)***							
Inflation	-0.208 (0.099)*	-0.22 (0.104)**	-0.177 (0.097)*	-0.22 (0.100)**	-0.224 (0.105)**	-0.188 (0.093)*	-0.201 (0.102)*	-0.179 (0.097)*
GDP growth	0.651 (0.231)**	0.634 (0.232)**	0.748 (0.164)***	0.665 (0.232)**	0.625 (0.230)**	0.315 -0.394	0.716 (0.217)***	0.778 (0.206)***
Unemployment	-0.102 (-0.368)	0.027 (-0.346)	0.627 (-0.373)		0.075 (-0.344)	1.916 (0.774)**	-0.1 (-0.337)	-0.088 (-0.33)
Output gap (economic cycle)	-1.922 (-3.656)	-1.956 (-3.701)	-2.637 (-2.995)	-2.117 (-3.702)	-1.957 (-3.676)	-3.271 (-7.127)	-3.418 (-3.919)	-5.071 (-3.808)
Quality of government		3.522 (-17.156)						
Within R2	0.2647	0.2882	0.5002	0.2534	0.2716	0.2724	0.2086	0.2115
N	2 059	2 059	2 059	2 064	2 059	1 050	2059	2059
Time fixed effects	x	x	x	x	x	x	x	x
Set of honeymoon dummies	x	x	x	x	x	x	x	x
Set of presidential dummies			x					

Note: Fixed-effects regressions with Driscoll and Kraay standard errors. Dependent variable: Executive approval. Standard errors are in parentheses. *, **, *** indicate significance at 90%, 95%, and 99% level, respectively. The within R-squared of a panel regression shows how much of the variation in the dependent variable within each country is captured by the model.

6 Conclusions and areas for further research

This paper exploits the potential of Internet search query data to understand better the social unrest and dissatisfaction with governments in Latin America and the Caribbean, particularly when it is combined with traditional explanatory variables. The paper draws three main conclusions:

First, the use of data generated by Internet searches has remarkable potential for research purposes and can inform policy making with readily available, anonymous and high-frequency data.

Second, the perception of corruption, insecurity and complaints about public services have a negative and statistically significant association with governments' approval. Indeed, these issues frequently appear in traditional opinion surveys among the primary sources of citizens' concerns. The fact that Internet searches on these topics appear as relevant drivers of citizens' approval of governments confirms that using this type of indicators is a promising path.

Third, the analysis proves that politicians have a period of high popularity in the early months of their mandate after controlling for traditional economic explanatory variables and perceptions. As shown, there is a positive and statistically significant relationship between the honeymoon period and the governments' approval. Besides, results proved the decreasing nature of the honeymoon effect after taking office, especially after the third quarter. These results reinforce the importance of using the political capital held by governments at the initial stages to move forward reform agendas.

However, strengthening all the stages of the policy-making cycle to preserve trust throughout is essential (OECD, 2017[25]). The governments of Latin America and the Caribbean need to strengthen trust among their citizens in order to advance more forcefully towards their development goals, especially given the weakened social contract resulting from the current context. For that purpose, governments are called upon to improve their openness to inform, consult and listen to citizens; their integrity to use power and public resources ethically; their responsiveness to provide public services; their reliability to protect citizens; and their fairness to improve socio-economic conditions for all (OECD, 2017[25]; 2017[26]).

As the use of the Internet and the availability of new data sources increase, the importance and potential benefits of using these resources in the policy making process grows. This paper has shed light on some of those potential uses, as other authors have highlighted the utilisation of these new data sources to measure people's subjective well-being (Algan et al., 2019[36]). The present work can be taken as the first step in a growing research agenda in this field. Future research could include the measurement of a "contagion effect". In other words, the reaction of citizens (e.g. navigating the Internet) to a political crisis in a neighbouring country. Similarly, these type of phenomena could be analysed from a different perspective, by looking at the type of issues that people search for in times of political or social unrest.

Annex A. Executive approval series

Figure A A.1. Executive approval rates in Latin America

Monthly data, between January 2006 and December 2015
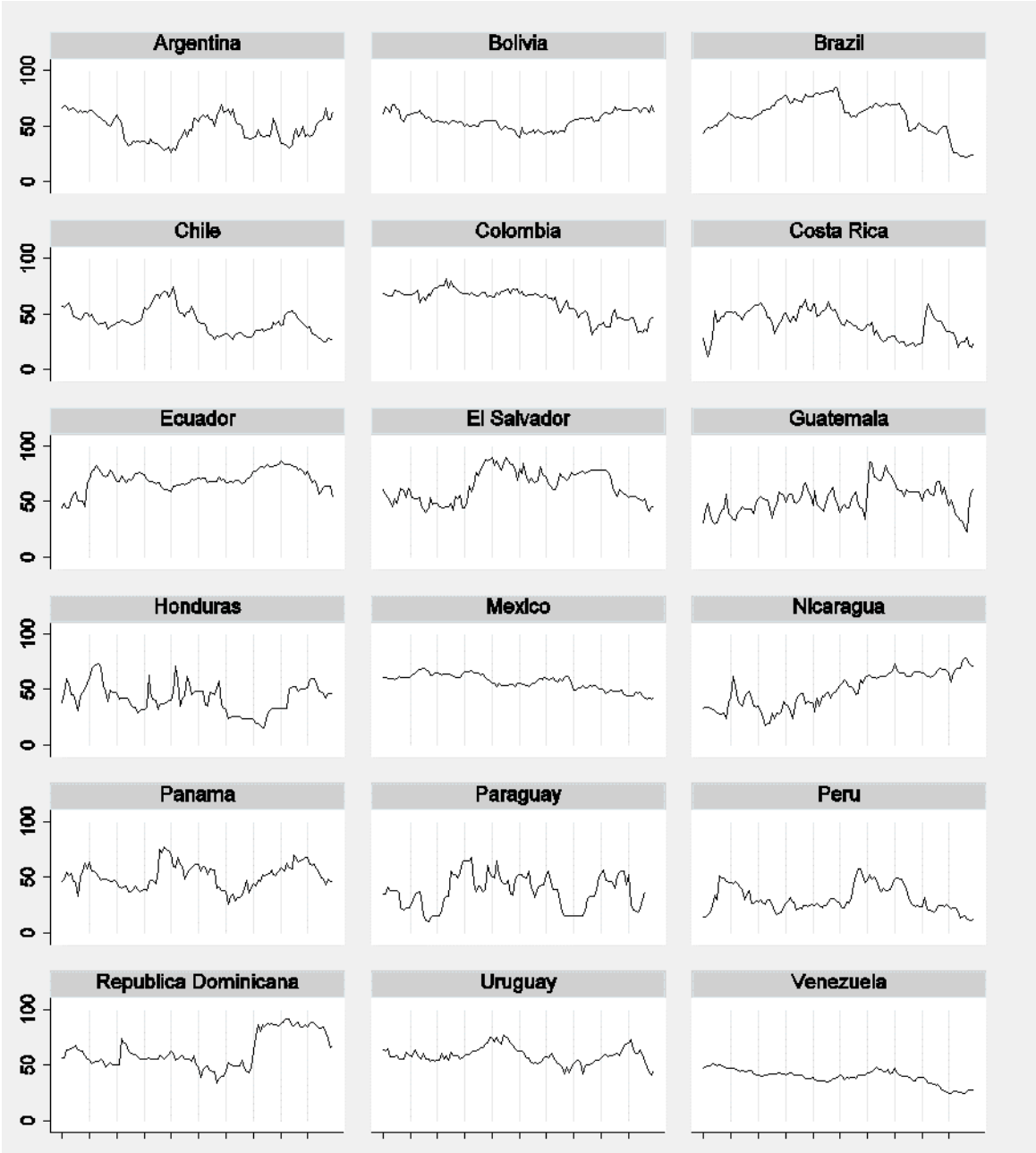
Note: Gridlines separate years from 2006 to 2015.
Source: Carlin et al. (2016[41]).

Table A A.1. Polling organisations

List of data sources used to build the Executive Approval database (Carlin et al., 2016[41]), by country

	Polling companies		
Argentina	ANALOGIASGES	FRAGA	ISONOMIAIMAG	OPSMADMIN
	ANALOGIASIMAG	GALLUPAPP	LAPOP	POLIARQUIA
	ARAGON	GALLUPDESEMP	LB	ROMER
	ARESCO	GIACOBBE	MGTFITAPP	ROMER3GBA
	CARLOSFARA	HUGOHAIME	MGTFITIMAG	RROUVIER
	CATTERBERG	HUGOHAIME_AP	MORAYARAUJOECNAC	RROUVIERIMAG
	CEOP	HUGOHAIME_DIR	MORIGES	SOCMERCECONPLAN
	CEOPIMAG	IBAROMETRO_GOB	MORIIMAG	SOCMERCGOVAPP
	DATAMATICA	IBAROMETRO_GOB2	OPSM	SOCMERCPRESIMAG
	EQUIS	IBAROMETRO_IMG	OPSM3	
Bolivia	CAPTURA	DATASIETE	LB	RADIOFIDES
	CIMAAPOYOAPP	EQUIPOS_MORI	MERCADOSMUESTRAS	TALCUAL
	CIMAAPOYODESMP	LAPOP	MORAYARAUJO	UNIVERSIDADES
Brazil	CIMAAPBR	DATAFOLHA	INSTSENSUS	VOXPOPULI
	CIMADESPEM	GALLUP	INSTSENSUS3	
	CNT-MDA	IBOPE	LAPOP	
	CNT-VOXPOPULI	IBOPEAPP	LB	
Chile	ADIMARK	FUNDFUTURO	IPSOSPARTOPP	LATERCERA
	CADEM	GIROPAIS	ISUCIMA	LBMORI
	CEP	ICCOM	ISUCMONTH	UDD
	CERC	IMAGINACCION	ISUCPANEL	UDP
	CIMAAPP	IPSOSMORICON	ISUCPANELFAV	
	CIMADESEMP	IPSOSMORIDES	ISUCSPEC	
	ELMERCURIOOPINA	IPSOSMORIMAN	LAPOP	
Colombia	CIFRASYCONCEPTOS_IMG	CNCIMAG	GALLUPFAV	NAPOLEON_FRANCO_TRAB
	CIMACNCAPP	DATEXCOIMG	LAPOP	YANHAAS
	CIMACNCDESEMP	DATEXCOMAN	LB	YANHAASAP
	CNCGES	GALLUPDES	NAPOLEON_FRANCO_IMG	
Costa Rica	BORGEYASOC	CIDGALLUPFAV	LAPOP	UNIMER
	BORGEYASOCAGR	CIMAAPP	LB	
	CIDGALLIPAPP	CIMADESEMP	UCR	
	CIDGALLUP	DEMOSCOPIA	UNACIONAL	
Dominican Republic	CIBAO	CIDGALLUPFAV	CIMASIGMA2DR	LAPOP
	CIDGALLUP	CIMAS2DESM	KELLER	LB
Ecuador	CEDATOS	CMS	INFORME_CONF3	MARKETAPP
	CEDATOS_FORMA	GALLUP	INFORME_CONFNAC	MARKETGES
	CIMACEDATOSAPP	INFCONF_GUAYAQ	LAPOP	PDO
	CIMACEDATOSDESEMP	INFCONF_QUITO	LB	SANTIAGOPEREZ
El Salvador	CIDGALLIPAPP	CIMABORGEDESEM	IUDOPGOB	LB
	CIDGALLUP	IUDOPACT	JABESPERF	LPGDATOS
	CIDGALLUPF	IUDOPDES	JABESSAT	MITOFSKY
	CIMABORGEAPP	IUDOPGES	LAPOP	
Guatemala	BORGEYASOCAGR	CIMAAPP	LAPOPFAV	VOXLATINAAPP
	BORGEYASOCGES	CIMADESEMP	LB	
	CIDGALLUP	GALLUPFAV	PRODATOS	
	CIDGALLUPAPP	LAPOP	PRODATOSPRESINDEX	
Honduras	CIDAGALLUPAPP	CIDGALLUP	CIMADESEMP	LB
	CIDAGALLUPFAV	CIMAAPP	LAPOP	
Mexico	ARCOP2	GEAISA_IMP	OPOPRM3VIVACDGOB2	OPOPRMNVIVAPPDES2

	ARCOP3	GEAISA3	OPOPRM4VIVACDGOB2	OPOPRMNVIVAPPGOB2
	BERUMEN	INDEMERC5	OPOPRM6VIVACDGOB2	OPOPRMNVIVAPPTRAB3
	BGCAPPDESAP	IPSOSBIMSA	OPOPRMNTELACDGOB2	PARAMETRIA
	BGCCALIF	LAPOP	OPOPRMNTELAPPDES2	PARAMETRIAIMAG
	BUENDIA	LB	OPOPRMNTELAPPDES3	REFORMA
	DEMOTECNIA	MITOFSKY	OPOPRMNTELAPPTRAB3	UNIVERSAL
	GEAISA_APP	MXPS	OPOPRMNVIVACDGOB2	VARELAYASOC
Nicaragua	BORGEYASOC	CIDGALLUPFAV	CIMAGESPRES	M&RAPPGES
	BORGEYASOCFAV	CIMABORGEAPP	LAPOP	M&RP3
	CIDGALLUP	CIMABORGEDESEMP	LB	M&RSUPP
Panama	CIMAPSMS2DESM	DICHTERNEIRAGES	LAPOP	UNIMER
	CIMAPSMSIGMA2	DICHTERNEIRALAB	LB	
	DATEXCOIMG	IPSOS	PSMSIGMA2	
Paraguay	ATISNEAD	CIMADESEMP	COIN	LAPOP
	ATISNEADAPP	CIRDCALIF	FIRST	LB
	CIES	CIRDPRCONT	GEO	MORI
	CIMAAPP	CIRDPRTRAB	ICA	
Peru	APOYOLC	DATUMINTL	IMAMARKAPP	ULIMA3LC
	APOYONAC	DATUMLC	IMASEN	ULIMALC
	CIMAAPYAPP	GFKAPR	LAPOP	ULIMANAC
	CIMAAPYDES	GFKEVAL	LB	VOXPOPULI
	CPILIMA	IDICE	PUCPLIMA	
	CPINAC	IMAMARK3	PUCPNAC	
Uruguay	CIFRA	CIMADESEMP	INTERCONSULT	MORI
	CIFRASIMP	FACTUM	LAPOP	MORIMONTEVIDEO
	CIMAAPP	GRUPORADAR	LB	OPCION
Venezuela	C21	DATOS	HERCON	KELLER
	CIMAAPP	DOXA	HINTERLACES	LAPOP
	CIMAGES	EVANSMCD	HINTERLACESAGRADO	LB
	DATANALISIS	GISXXIDES	ICS	POLINOMICS
	DATANALISISCARACAS	GISXXIGES	IVAD	VARIANZASGESTION
	DATIN	GQRR	IVADGES	VARIANZASPERFILIMAG

Source: Carlin et al. (2016[41]).

Annex B. Google Trends series

A reference variable to enable comparability

Google Trends series do not measure the number of searches for a given item, but rather its relative popularity in a given period within a specific area. Scores range between 0 and 100, where 100 is granted to the period of time with the highest number of relative searches. For example, if we download the series of searches for corruption in Mexico between 2006 and 2015, the value(s) 100 in the series will correspond to the popularity peak of searches for corruption in Mexico in that period. In this way, if there are higher peaks before or after the covered period, the maximum (100) will continue to be determined by the popularity peak between 2006 and 2015. *Google Trends* interface allows downloading a maximum of five series simultaneously. In cases where more than one series is downloaded, the 100 (maximum) will be determined by the highest popularity peak of all series, which will affect the scale of the other series downloaded.

One of the limitations when analysing *Google Trends* series is the non-comparability on the scale of two series that were downloaded separately. While the 100 in a popular series may represent millions of searches, the 100 in an unpopular series may have just a few hundred behind.

To minimize the negative effects of this characteristic in our analysis, we have downloaded each variable simultaneously with a chosen series in common. To choose it, we consider a list of series that had a stable popularity throughout the year (non-seasonal) and the period analysed (stationary). We chose the term *migraña* (headache), which has a relatively stable fluctuation during the whole period.

Internet search query variables

For the corruption series, two searches were used simultaneously for each country: "corrupcion" and "migraña". In the case of Brazil, we used the searches "corrupção + corrupto + corrupssão + corupção + corupsão + corrupição + corrupçao + corrupcao" and "enxaqueca".

For the insecurity series, we used the searches "inseguridad" and "migraña". In the case of Brazil, we used the search "insegurança + inseguransa" and "enxaqueca".

For jobs series, we used the search "busco trabajo + busco empleo". In the case of Honduras, this search does not yield results, so the series corresponds to "busco trabajo" only. In the case of Brazil, we used the search "procuro emprego + procuro trabalho + busco emprego + busco trabalho" and "enxaqueca".

For the complaints series, we used the search "reclamos + quejas + reclamo + queja". This search did not work for four countries. For the series of Nicaragua, Honduras and El Salvador, only the word "reclamo" was used (for two reasons: (1) it was the search with a higher average search between "queja" and "reclamo" and (2) it had more observations, that is, fewer zeros kept in the series). In the case of Ecuador, the search was "reclamo + queja + reclamos" because the word "quejas" in Ecuador is popular to refer to a well-known poem in that country. For the case of Brazil, we used the search "reclamação + reclamamassão + reclamassao".

Notes

[1] Trust relationships in society (public trust) include citizens towards government, citizens towards other citizens, citizens toward political leaders, businesses towards government, government towards citizens, within government, and among government leaders. These are frequently clustered into two broader categories: (i) interpersonal trust, in the realm of human and social interactions, and (ii) systemic or institutional trust, in the realm of public and political institutions (OECD, 2017[25]).

[2] Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Dominican Republic, Uruguay and Venezuela.

[3] In particular, we have used the open question "What do you think is the most important problem in your country?" (Latinobarómetro, 2018[51]).

[4] To test the presence of heteroscedasticity in the residuals, we applied the Modified Wald test for group-wise heteroscedasticity in a fixed effect regression model. To test the presence of temporal correlation in the time series, we applied the Wooldridge test for serial (temporal) correlation in panel-data models. To test the presence of spatial correlation in the cross-sections, we applied the Pesaran's test of cross-sectional dependence in panel data models (Pesaran, 2004[52]).

[5] Considering that we could only include 20 variables in this analysis, we could not include the set of time dummies. Also, we only included one honeymoon dummy variable for the first six months after taking office.

References

Algan, Y. et al. (2016), *Big data measures of well-being: Evidence from a Google well-being index in the United States*, OECD Statistics Working Papers 2016/03, Paris.   [45]

Algan, Y. et al. (2019), "Well-being through the lens of the internet", Vol. 14/1, https://doi.org/10.1371/journal.pone.0211586.   [36]

Algan, Y. et al. (2017), "The European Trust Crisis and the Rise of Populism", *Brookings Papers on Economic Activity*, Vol. 2, pp. 309-400.   [15]

Andrews, D. (1991), "Heteroskedasticity and autocorrelation consistent covariant matrix estimation", *Econometrica*, Vol. 59/3, pp. 817-858.   [49]

Arce, M. (2003), "Political violence and presidential approval in Peru", *The Journal of Politics*, Vol. 65/2, pp. 572-583.   [23]

Berlemann, M. and S. Enkelmann (2014), "The economic determinants of US presidential approval: A survey", *European Journal of Political Economy*, Vol. 36, pp. 41-54.   [19]

Carlin, R. et al. (2018), "Public support for Latin American presidents: The cyclical model in comparative perspective", *Research & Politics*, Vol. 5/3, http://dx.doi.org/2053168018787690.   [21]

Carlin, R. et al. (2016), *Executive Approval Database 1.0*, Available for download at www.executiveapproval.org.   [41]

Carlin, R., G. Love and C. Martínez-Gallardo (2014), "Cushioning the fall: Scandals, economic conditions, and executive approval", *Political Behavior,*, Vol. 37/1, pp. 109-130.   [20]

Choi, H. and H. Varian (2012), "Predicting the present with Google Trends. Economic Record", Vol. 88/s1, pp. 2-9.   [31]

Dickerson, B. (2016), "Economic perceptions, presidential approval, and causality: The moderating role of the economic context", *American Politics Research*, Vol. 44/6, pp. 1037-1065.   [10]

Donovan, K. et al. (2019), *Motivated Reasoning, Public Opinion, and Presidential Approval*, Political Behavior, https://doi.org/10.1007/s11109-019-09539-8.   [16]

Driscoll, J. and A. Kraay (1998), "Consistent covariance matrix estimation with spatially dependent panel data", *Review of economics and statistics*, Vol. 80/4, pp. 549-560.   [47]

Facebook, OECD and WB (2020), *Global State of Small Business Report*, Facebook, the Organisation for Economic Co-operation and Development (OECD), and the World Bank (WB), https://dataforgood.fb.com/global-state-of-smb/. [34]

Geys, B. (2010), "Wars, presidents, and popularity: The political cost (s) of war re-examined", *Public Opinion Quarterly*, Vol. 74/2, pp. 357-374. [14]

Geys, B. and J. Vermeir (2008), "Taxation and presidential approval: separate effects from tax burden and tax structure turbulence?."", *Public Choice*, Vol. 135/3-4, pp. 301-317. [17]

Ginsberg, J. et al. (2009), "Detecting influenza epidemics using search engine query data", *Nature*, Vol. 457/7232, pp. 1012-1014. [29]

González, S. and C. Smith (2017), *The accuracy of measures of institutional trust in household surveys: evidence from the oecd trust database*, OECD Statistics Working Papers, No. 2017/11, OECD Publishing, Paris, https://doi.org/10.1787/d839bd50-en. [27]

Guriev, S., N. Melnikov and E. Zhuravskaya (2019), *3G Internet and Confidence in Government (June 30, 2019)*, Available at SSRN: https://ssrn.com/abstract=3456747 or http://dx.doi.org/10.2139/ssrn.345674. [5]

Höchtl, J., P. Parycek and R. Schöllhammmer (2016), "Big data in the policy cycle: Policy decision making in the digital era", *Journal of Organizational Computing and Electronic Commerce*, Vol. 26/1-2, pp. 147-169. [40]

Hoechle, D. (2007), "Robust standard errors for panel regressions with cross-sectional dependence", *Stata Journal*, Vol. 7/3, p. 281. [46]

Jun, S., H. Sun Yoo and S. Choi (2018), "Ten years of research change using Google Trends: From the perspective of big data utilizations and applications", *Technological Forecasting and Social Change*, Vol. 130, pp. 69-87. [35]

Latinobarómetro (2018), *Latinobarómetro 2018 (database)*, Corporación Latinobarómetro, Santiago de Chile. [51]

Mueller, J. (1973), "War, presidents, and public opinion", *John Wiley & Sons*. [9]

Mueller, J. (1970), "Presidential Popularity from Truman to Johnson", *American Political Science Review,*, Vol. 64/1, pp. 18-34. [8]

Murillo, M. and G. Visconti (2017), "Economic performance and incumbents' support in Latin America", *Electoral Studies*, Vol. 45, pp. 180-190. [22]

Murtin, F. et al. (2018), *Trust and its determinants: Evidence from the Trustlab experiment*, OECD Statistics Working Papers 2018/02, Paris, https://doi.org/10.1787/869ef2ec-en. [28]

Newey, W. and K. West (1987), "Hypothesis testing with efficient method of moments estimation", *International Economic Review*, pp. 777-787. [48]

Newman, B. and A. Forcehimes (2010), ""Rally round the flag" events for presidential approval research", *Electoral Studies*, Vol. 29/1, pp. 144-154. [12]

OECD (2020), *Government at a Glance: Latin America and the Caribbean 2020*, OECD Publishing, Paris. [2]

OECD (2019), *Benchmarking Higher Education System Performance, Higher Education*, OECD Publishing, Paris, https://doi.org/10.1787/be5514d7-en. [33]

OECD (2019), *La Integridad Pública en América Latina y el Caribe 2018-2019: De Gobiernos reactivos a Estados proactivos*, OECD Publishing, Paris. [7]

OECD (2018), *Integrity for Good Governance in Latin America and the Caribbean: From Commitments to Action*, OECD Publishing, Paris, https://dx.doi.org/10.1787/9789264201866-en. [6]

OECD (2017), *OECD Guidelines on Measuring Trust*, OECD Publishing, Paris, https://doi.org/10.1787/9789264278219-en. [26]

OECD (2017), *Trust and Public Policy: How Better Governance Can Help Rebuild Public Trust*, OECD Public Governance Reviews, OECD Publishing, Paris, https://doi.org/10.1787/9789264268920-en. [25]

OECD (2015), *Government at a Glance 2015*, OECD Publishing, Paris, https://doi.org/10.1787/gov_glance-2015-en. [24]

OECD/CAF/UN ECLAC (2018), *Latin American Economic Outlook 2018: Rethinking Institutions for Development*, OECD Publishing, Paris, https://dx.doi.org/10.1787/leo-2018-en. [3]

OECD et al. (2019), *Latin American Economic Outlook 2019: Development in Transition*, OECD Publishing, Paris, https://dx.doi.org/10.1787/g2g9ff18-en. [4]

Pencheva, I., M. Esteve and S. Mikhaylov (2018), "Big Data and AI–A transformational shift for government: So, what next for research?", *Public Policy and Administration*, http://dx.doi.org/0952076718780537. [39]

Pesaran, M. (2004), *General diagnostic tests for cross section dependence in panels*, Institute for the Study of Labor (IZA), Discussion Paper No. 1240, Bonn, http://ftp.iza.org/dp1240.pdf. [52]

Reimsbach-Kounatze, C. (2015), "The Proliferation of "Big Data" and Implications for Official Statistics and Statistical Agencies: A Preliminary Analysis", *OECD Digital Economy Papers*, No. 245, OECD Publishing, Paris, https://dx.doi.org/10.1787/5js7t9wqzvg8-en. [38]

Romero, V., B. Magaloni and A. Díaz-Cayeros (2016), "Presidential Approval and Public Security in Mexico's War on Crime." Latin American Politics and Society 58.2 (2016): 100-123.", *Latin American Politics and Society*, Vol. 58/2, pp. 100-123. [13]

Scheitle, C. (2011), "Google's Insights for Search: A Note Evaluating the Use of Search Engine Data in Social Research", *Social Science Quarterly,*, Vol. 92/1, pp. 285-295. [32]

Seifter, A. et al. (2010), "The utility of "Google Trends" for epidemiological research: Lyme disease as an example", *Geospatial health,*, Vol. 4/2, pp. 135-137. [30]

Smyth, D. and S. Taylor (2003), "Presidential popularity: what matters most, macroeconomics or scandals?", *Applied Economics Letters*, Vol. 10/9, pp. 585-588. [11]

Stephens-Davidowitz, S. (2013), *Essays Using Google Data (Doctoral dissertation)*, Department of Economics, Harvard University, Cambridge. [37]

Stimson, J. (2015), *Wcalc6.1: The Dyad Ratios Algorithm for creating time series from scattered survey marginals*, http://stimson.web.unc.edu/software/. [43]

Stimson, J. (2018), "The Dyad ratios algorithm for estimating latent public opinion: estimation, testing, and comparison to other approaches", *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, Vol. 137-138/1, pp. 201-218. [42]

Teorell, J. et al. (2019), *The Quality of Government Standard Dataset, version Jan19*, University of Gothenburg: The Quality of Government Institute, http://www.qog.pol.gu. [50]

Thomson Reuters (2012), *Thomson Reuters' Datastream Infobase [Online]*, Available at: Subscription Service. [44]

Transparency International (2019), *Corruption Perceptions Index (database)*, Transparency International, Berlin, https://www.transparency.org/cpi2018. [1]

Ward, G. (2015), "Is happiness a predictor of election results?", *CEP Discussion Paper No 1343. Centre for Economic Performance. London School of Economics and Political Science*, https://cep.lse.ac.uk/pubs/download/dp1343.pdf. [18]

Other titles in the series

The former series known as "Technical Papers" and "Webdocs" merged in November 2003 into "Development Centre Working Papers". In the new series, former Webdocs 1-17 follow former Technical Papers 1-212 as Working Papers 213-229.

All these documents may be downloaded from:

https://www.oecd-ilibrary.org/development/oecd-development-centre-working-papers_18151949

or obtained via e-mail (dev.contact@oecd.org).

Recent working papers:

Working Paper No. 342, "The Social Institutions and Gender Index (SIGI) 2019: A revised framework for better advocacy, by Gaëlle Ferrant, Léa Fuiret and Eduardo Zambrano, September 2020.

Working Paper No. 341, "Economic globalisation, inequality and the role of social protection", by Andreas Bergh, Alexandre Kolev and Caroline Tassot, November 2017.

Working Paper No. 340, "No sympathy for the devil! Policy priorities to overcome the middle-income trap in Latin America", by Angel Melguizo, Sebastián Nieto-Parra, José Ramón Perea and Jaime Ariel Perez, September 2017.

Working Paper No. 339, "The grant element method of measuring the concessionality of loans and debt relief", by Simon Scott, May 2017.

Working Paper No. 338, "Revisiting personal income tax in Latin America: Evolution and impact", by Alberto Barreix, Juan Carlos Benítez and Miguel Pecho, March 2017.

Working Paper No. 337, "The pursuit of happiness: Does gender parity in social institutions matter?", by Gaelle Ferrant, Alexandre Kolev and Caroline Tassot, March 2017.

Working Paper No. 336, "Fiscal policy and the cycle in Latin America: The role of financing conditions and fiscal rules", by Enrique Alberola, Iván Kataryniuk, Ángel Melguizo and René Orozco, February 2017.

Working Paper No. 335, "The economic effects of labour immigration in developing countries: A literature review", by Marcus H. Böhme and Sarah Kups, January 2017.

Working Paper No. 334, "Harnessing the digital economy for developing countries", by Carl Dahlman, Sam Mealy and Martin Wermelinger, December 2016.

Working Paper No. 333, "The cost of air pollution in Africa", by Rana Roy, September 2016.