AGAINST ALL ODDS: FORECASTING BRAZILIAN PRESIDENTIAL ELECTIONS IN TIMES OF POLITICAL DISRUPTION

Contra todas as probabilidades: Previsão das eleições presidenciais brasileiras em tempos de ruptura política

Contra todo pronóstico: Pronosticando elecciones presidenciales brasileñas en tiempos de disrupción política

FREDERICO BERTHOLINI iD frederico.bertholini@unb.br [1]

LUCIO RENNÓ iD luciorenno@unb.br [1]

MATHIEU TURGEON iD mturgeo4@uwo.ca [2]

[1] Universidade de Brasília
[2] Western University

Abstract
When the number of observed elections is low, subnational data can be used to perform electoral forecasts. Turgeon and Rennó (2012) applied this solution and proposed three forecasting models to analyze Brazilian presidential elections (1994-2006). The models, adapted from forecasting models of American and French presidential elections, considers economic and political factors. We extend their analysis to the recent presidential elections in Brazil (2010, 2014 and 2018) and find that the addition of the three recent elections does not improve the accuracy of our forecast models although it strengthens the relationship between the explanatory variables and vote for the incumbent. We also find that models based on the popularity of the incumbent outperform those based on trial-heat polls and that electoral forecast models can survive earthquake elections like the 2018 election that led to the unexpected rise of "outsider" and extremist candidate Jair Bolsonaro.

Palavras-chave:
Previsão eleitoral; Erro de previsão; Eleições presidenciais; Metodologia; Brasil

Resumo
Quando o número de eleições observadas é baixo, pode-se usar dados subnacionais para realizar previsões eleitorais. Turgeon e Rennó (2012) aplicaram essa solução e propuseram três modelos de previsão para analisar eleições presidenciais brasileiras ocorridas entre 1994 e 2006. Os modelos, adaptados de modelos de previsão de eleições presidenciais americanas e francesas, consideram fatores econômicos e políticos. Estendemos esta análise para as recentes eleições presidenciais no Brasil (2010, 2014 e 2018) e demonstramos que a adição das três eleições mais recentes não melhora a precisão dos modelos preditivos, embora fortaleça a relação entre as variáveis explicativas e o voto no incumbente. Também concluímos que os modelos baseados na popularidade do incumbente superam aqueles baseados em pesquisas eleitorais e que os modelos de previsão eleitoral podem sobreviver a eleições com muito ruído, como a de 2018, que levou à ascensão inesperada de um candidato de extrema-direita, Jair Bolsonaro.

Palabras clave:
Pronóstico electoral; Error de pronóstico; Elecciones presidenciales; Metodología; Brasil

Resumen
Cuando el número de elecciones observadas es bajo, se puede usar datos subnacionales para hacer predicciones electorales. Turgeon y Rennó (2012) aplicaron esta solución y propusieron tres modelos de predicción para analizar las elecciones presidenciales brasileñas entre 1994 y 2006. Los modelos, adaptados de los modelos de predicción de elecciones presidenciales de Estados Unidos y Francia, consideran factores económicos y políticos. Extendemos este análisis a las recientes elecciones presidenciales en Brasil (2010, 2014 y 2018) y demostramos que la adicción de las tres elecciones más recientes no mejora la precisión de los modelos predictivos, aunque fortalece la relación entre las variables explicativas y el voto por el titular. También concluimos que los modelos basados en la popularidad del titular superan a los basados en encuestas electorales y que los modelos de predicción electoral pueden sobrevivir a elecciones ruidosas como la de 2018, que condujo al ascenso inesperado de un candidato de la extrema derecha, Jair Bolsonaro.

INTRODUCTION

Election forecasting in recently democratized countries is difficult —given the scarcity of elections and more unstable political environments— but not impossible (Bunker and Bauchowitz 2016; Cantu et al. 2016; Jastramskis 2012; Turgeon and Rennó, 2012; Toros 2012). To circumvent the low-N problem, Turgeon and Rennó (2012) moved to a lower level of analysis. By examining presidential Brazilian elections, the authors incorporated information from Brazil's 27 states. The authors relied on election forecast models that have been commonly used in settings of political stability because presidential elections in Brazil since 1994 have been dominated exclusively by two large parties—the *Partido dos Trabalhadores* (PT) and the Partido da Social Democracia Brasileira (PSDB).

How do models fare when more elections are analyzed, including elections followed by disruptive political events and deep institutional changes? This paper expands on Turgeon and Rennó's (2012) original study by including the Brazilian presidential elections of 2010, 2014 and 2018. Do prior results remain stable with more datapoints, showing that models designed for stable two-party systems hold elsewhere? Furthermore, the equilibrium that marked Brazilian elections between 1994 and 2014 ended abruptly in 2018 when an outsider from a marginal party won the presidency—after severe economic and political crises and deep institutional transformations. Do results that focus on stable party systems hold after an earthquake election, like the Brazilian presidential 2018?

The 2018 elections also pose an additional challenge for election forecasting models because the sitting president —Dilma Rousseff— was impeached mid-mandate in 2016, producing an electoral contest with potentially more than one party claiming responsibility for policies while in government and without a clear incumbent, given the disruption of prior governing and opposition coalitions. There were no individual candidate competing for reelection, although many candidates from the previous governments (Dilma's Worker Party and Temer's MDB) were seeking the presidency. Moreover, Brazil, like other parts of the world, has seen the rise of a populist and polarizing figure, Jair Bolsonaro. This change in the political environment might affect the ability of forecasting models, depending on whether they place more emphasis on the fundamentals of the economy and satisfaction with the incumbent or on trial-heat polls, as we have seen in the 2016 American presidential election (Campbell, 2017).

The inclusion of more elections may improve the precision of earlier forecasts by increasing sample size. The electoral earthquake of 2018, however, may hinder predictability, posing the question about how forecasting models fare in less stable environments. We can verify if forecasting models based on the fundamentals of the economy and popularity of the incumbent performed better than those reliant on trial-heat polls. Our findings show that the addition of recent elections does not contribute to improve our models' forecasts, confirming the strength of prior analysis based on fewer cases and showing that the subnational strategy of increasing sample size is a practical and valid solution for forecasting elections in young democracies. What it does, however, is to strengthen the theoretical arguments around our explanatory variables. Finally, our findings also show that the 2018 earthquake election is surprisingly better predicted by a model based on the fundamentals of the economy and popularity of the incumbent than those based on trial-heat polls.

FORECASTING ELECTIONS IN YOUNG DEMOCRACIES

As the literature on electoral forecasting increased exponentially in the established democracies (e.g., US and Great Britain), it has also gradually expanded to Second and Third Wave democracies like France and Germany in Europe and more recently Brazil, Turkey, and Lithuania[1]. The expansion of electoral forecasting models to more recently democratized countries poses new challenges. Two factors deserve particular attention. The new, younger democracies raise a methodological issue because they have held few elections to estimate electoral forecasting models. The question is one of degrees of freedom, reducing the precision of analysis and restricting model specification. A second significant problem concerns adapting forecast models that were initially designed for stable, two-party systems, to unstable, multiparty systems, especially those from the Third Wave democracies. In addition, electoral environments in younger democracies are usually more convoluted (more parties, more candidates, more volatility), rendering election results potentially less predictable. The transposition of electoral forecasting models to new democracies, therefore, provides for a rigorous test of their generalizability.

These challenges have been faced by the extant literature. For instance, a solution for the small-N problem is dealt with by using subnational measures of election outcomes and predictors. The study of the French case, for instance, has relied on local-level data to increase sample size (Jérôme and Jerome-Speziari, 2004; Auberger 2010; Foucault and Nadeau 2012). An identical approach has been adopted for Brazil (Turgeon and Rennó, 2012). These strategies model election outcomes for national level elections using information from subnational units. In a way, this strategy is like the one adopted in the United States to forecast Electoral College results, by modeling voting results at the state-level to anticipate the outcome of the Electoral College (Campbell, 1992; Berry and Bickers 2012; Jérôme and Jerome-Speziari, 2012, 2016; Jérôme et al., 2021). Other studies have attempted to deal with the small-N problem in a similar fashion but by examining other election outcomes. For instance, Jastramskis (2012) forecasts vote for a major national party in Lithuania, instead of focusing on the incumbent. Toros (2012), on the other hand, proposes to forecast mayoral elections in Turkey.

The second challenge concerns the expansion of the electoral forecasting models initially developed to understand older democracies to more recent democracies where elections are few, far in-between and frequently characterized by more complex institutional arrangements. Multiple parties, ballotage systems, party system instability all create specific problems that may limit the applicability

---

1. See the special issue of the *International Journal of Forecasting* 28:12, Election Forecasting in Neglected Democracies.

of established electoral forecasting models whose main premise is that of reward and punishment for good or bad times. Specifically, the more complex and less consolidated institutional traits of younger democracies may render accountability opaquer (Powell, 2004). Furthermore, young democracies are frequently characterized by weak parties with shallow social roots, increasing electoral volatility and thus rendering elections more unpredictable (Baker et al., 2020).

Finally, specific electoral episodes may be harder to predict given the unfolding of campaigns and the emergence of unexpected, outsider candidates. In this sense, even in strongly consolidated democracies, elections can be hard to forecast (e.g., the 2016 and 2020 US presidential elections)[2]. The literature on electoral forecasting, however, has generally found strong support for the main tenets of retrospective voting that underly most forecasting models. From Brazil to Norway, passing by France, Germany, Spain, and Turkey, forecasting models developed for the US and Great Britain tend to perform well in other contexts. Hence, even earthquake elections, in complex institutional environments, can be explained and modeled with the theoretical assumptions that the state of the economy and government evaluation are central to predicting vote for the incumbent party. The case of Brazil that we discuss hereafter exemplifies the virtues of forecasting models.

FORECASTING ELECTIONS IN TIMES OF POLITICAL DISRUPTION

Brazil has held eight presidential contests (1989, 1994, 1998, 2002, 2006, 2010, 2014 and 2018) since (re)democratization. Between 1994 and 2014 the *Partido da Social Democracia Brasileira* (PSDB) or the *Partido dos Trabalhadores* (PT) won the presidency: each party always reaching the second round (when such round was necessary). This pattern of polarization between the PT and the PSDB ended abruptly in 2018, but signs of political instability were first noticed in July 2013 with massive street protests ahead of the 2014 World Cup. There was a sense among Brazilians that the government was spending more on stadiums than in health services, education or urban transportation and that corruption was widespread. Meanwhile, the economy started to falter with increases in inflation, unemployment, and public debt. Then president Dilma Rousseff saw her popularity plumet and what was to be an easy reelection pledge, became a very close race.

Despite the adversity, Dilma Rousseff was reelected for a second term in 2014, with a very small margin of 3 percentage points against Aécio Neves, the PSDB candidate. The economy continued its downward trend in 2015 and 2016, making

---

2. See the PS October 2016 and PS January 2021 issues for the full set of forecasts of American Elections.

it harder for President Dilma to govern. Protesters hit the streets again and after the *Lava-Jato* Operation, a Brazilian *Mani Puliti* of sorts, had uncovered significant bribing schemes in the Brazilian oil giant, Petrobras, that benefitted the PT and its allies. The pressure became unbearable, leading to the impeachment of President Dilma Rousseff in August of 2016. Then Vice-President, Michel Temer of the PMDB (now MDB), took office, terminating a 22-year period of PSDB/PT reign.

The new President, Michel Temer, failed to restore economic growth and, after a year in office, his government was also deeply tarnished by corruption scandals that involved President Temer himself and one of the most important opposition leaders to Dilma, her runner-up contender in 2014, Aécio Neves. Temer faced two consecutive office-removal votes in Congress as a result of his involvement in corruption scandals but survived each time. Temer's ability to govern, however, was severely affected and, in his brief mandate, precious time was lost in defending himself instead of advancing legislation and policies to curb the economic crisis. Temer's popularity, which was never high to start out with, reached historical lows toward the end of his term.

The result of these concomitant and deep economic and political crises was that the two most powerful parties in Brazil (PT and PSDB) and to some extent the PMDB/MDB (Temer's party, a party that had been part of the governing coalition since 1995) were all affected by and blamed for the country's economic and political misfortunes. Dissatisfaction and frustration among the Brazilian public with its political elites was strong and created an ideal power vacuum for Jair Bolsonaro —an obscure and eccentric radical right-wing politician backed by an equally obscure political party (Partido Social Liberal, PSL)— to launch his presidential bid.

Immediately after Dilma's 2014 election, Jair Bolsonaro started a modest campaign for president that heavily relied on social media. He defended a right-wing, socially conservative agenda and made virulent attacks on the PT (Rennó, 2020). Bolsonaro became the spokesperson of *antipetismo* —the rejection of the Worker's Party (PT)— a movement headed up until now by the PSDB (Samuels and Zucco, 2018). Bolsonaro's radical positions against gender politics, homosexuals, and racial policies, coupled with a strong stance on fighting crime and corruption and free market policies resonated well with a dissatisfied and disillusioned population (Rennó, 2020). His incessant campaigning on social media and in person across the country made him a top contender for the 2018 election, to the surprise of most political pundits and elites.

The 2018 presidential election was also the first election after significant political reforms had been adopted (Rennó, 2020). Donations were now more restricted, prohibiting corporations from financing campaigns, spending limits were established by elected office, gender quotas on campaign finance where instituted, party performance clauses were put in place and the official campaign period was reduced in half, from 90 to 45 days. All these factors influenced the impact

of traditional electoral resources in Brazil, like TV and radio advertising and how campaign resources are allocated, more generally.

The 2018 elections were also marked by unprecedented and unexpected events. First, the PT nominated Lula da Silva as his presidential candidate. The problem was that Lula was incarcerated in Curitiba, Paraná, based on allegations of corruption. Surprisingly, Lula led the polls for most of the election, closely followed by Jair Bolsonaro. Lula's candidacy was eventually barred by the Supreme Court and Fernando Haddad became the official candidate for the PT only two weeks before the first round. Lula's popularity was enough to qualify Haddad for the second round.

Second, in the early days of the first-round election, Jair Bolsonaro was stabbed while campaigning. Bolsonaro was hospitalized for a long period of time but eventually survived the attack. During this time, Bolsonaro's popularity skyrocketed, as the press and the other presidential candidates softened their discourse about him (given his frail health situation) while his supporters took full advantage of the situation to portray him as a martyr and savior. Bolsonaro's attack also allowed him to escape from participating in the presidential debates, thus managing to keep a positive image of him among voters. In the end, Bolsonaro handily defeated Haddad in the second-round election, becoming Brazil's 38th President.

Admittedly, the 2018 Brazilian elections were marked by many factors that render it very hard to predict using conventional election forecasting models. First, it was an election with no clear incumbent since President Dilma was removed from office in mid-mandate and President Temer (her then Vice-President) did not seek the office. Both the MDB and the PT could claim credit for their policies while in government, but both were also blamed for the misfortunes of the preceding years. Temer decided not to run, but put his support behind Minister of Finance, Henrique Meirelles. Second, a complete outsider, with no partisan support and very few resources, rode to victory without much difficulty. Third, the campaigning and electoral rules were markedly different than in past elections. And fourth and finally, the attack on Bolsonaro changed dramatically the dynamic of the presidential election. Undoubtedly, the 2018 scenario contrasts greatly with the earlier period from 1994 to 2014 where Brazilian presidential elections have been systematically centered around two opposing forces: the PT and PSDB. Thus, we ask: how do electoral forecast models fare in this adverse scenario?

DATA AND MODELS

Our interest lies in using data from the 1994 to the 2018 Brazilian presidential elections to evaluate three election forecasting models tested in Turgeon and Rennó (2012) over a shorter period. The data come from various sources including

the *Tribunal Superior Eleitoral*[3] for election data, the *Instituto Brasileiro de Geografia e Estatística* (IBGE)[4] and the *Instituto de Pesquisa Econômica Aplicada*[5] (IPEA) for economic data and Fernando Rodrigues's *Poder360* website for polling data.[6] It is worth noting that we are limited in the type of data that we can rely on (and Turgeon and Rennó (2012) did) given the necessity for the data to be available at the subnational state level.

The three models used in Turgeon and Rennó (2012) are based on existing models to predict U.S. and French presidential elections, with slight modifications. The models are parsimonious and have been around for some time. But, more importantly, they make use of data that are available in the Brazilian context (both at the national and subnational levels). Table 1 presents the details about the models, including information about the authors of said models, predictors used in the original and proposed adapted models, and the level of analysis and election years for which the models are estimated.

Table 1. Election Forecasting Models

Models and Authors	Original model	Adapted model (level of availability)	Years
Model 1: Abramowitz (2008)	1. Popularity of the incumbent 2. Second quarter GDP 3. Third term dummy	1. Popularity of the incumbent (national) 2. Real annual GDP growth (state) 3. Third term dummy	1994-2018
Model 2: Campbell (2008)	1. Trial-heat poll 2. Second quarter GDP	1. Trial-heat poll (state) 2. Real annual GDP growth (state)	2006-2018
Model 3: Lewis-Beck et al. (2008)	1. Popularity of the incumbent 2. Unemployment	1. Popularity of the incumbent (national) 2. Real annual GDP growth (state)	1994-2018

Source: Authors.

All three models use the same dependent variable: the first-round vote for the incumbent candidate in the presidential elections, as measured as the percentage of all votes received by the incumbent in each of the 27 states. In 1994, 1998 and

---

3. The TSE website is http://www.tse.gov.br/internet/index.html.
4. The IBGE website is www.ibge.gov.br/english/
5. The IPEA datasets are available at http://www.ipeadata.gov.br/
6. https://www.poder360.com.br/pesquisas-de-opiniao/

2002, the incumbent party is the PSDB (with Fernando Henrique Cardoso and Jose Serra as candidates); and, in 2006, 2010 and 2014 the incumbent party is the PT (with Lula da Silva and Dilma Rousseff as candidates). As discussed extensively above, 2018 is peculiar in the sense that there is no clear incumbent running for office. One possibility is to consider the PT as the incumbent party because Dilma Rousseff was elected in 2014 although removed from office mid-mandate. Another possibility is to consider the party of Michel Temer —PMDB/MDB— as the incumbent party since it was the party in the presidency at the time of the election. Both the PT, with Fernando Haddad, and the PMDB/MDB, with Henrique Meirelles, had candidates running in the 2018 election. Thus, in theory, we could consider two potential incumbent parties running in 2018, adding another layer of complexity to our forecasting models. We have, however, a theoretical preference for the model with Fernando Haddad (PT) as the incumbent for two reasons. First, the *Partido dos Trabalhadores* (PT) is the party that won the 2014 election and, if Dilma Rousseff had not been removed from office, the PT would have been the official incumbent. In sum, the PT was the legitimate incumbent in 2018, especially given the dubious accusations President Rousseff faced during her impeachment trial. Second, and maybe more importantly, Brazilian politics since 1994 has been structured around the PT and some other competing party (the PSDB from 1994 to 2014 and the PSL with Bolsonaro in 2018 (Duque and Smith, 2019). Thus, the PT has been the central figure of national politics and the party to "beat." The same cannot be said of the MDB (former PMDB), Temer/Meirelles' party. Although the MDB played a supporting role in each and every governing coalition since 1995 —including the one formed by Dilma Rousseff during her second term, before being impeached— the party has never been a serious contender for the presidency (Pereira and Bertholini, 2018). For these reasons, we have a theoretical preference for the model with Fernando Haddad (PT) as the incumbent candidate. In the Appendix (Figure A1), we also present the results using Henrique Meirelles (PMDB/MDB) instead as the incumbent candidate.

The first model, inspired from Abramowitz (2008) accounts for economic activity, popularity of the incumbent and a dummy for an incumbent party seeking a third term (2002 and 2010). Economic activity is the real annual growth of the GDP from the year preceding the election to the election year, measured at the state level[7]. Unfortunately, the popularity of the incumbent is not available at the state level. Instead, we use a measure at the national level. The polls are from Datafolha and were conducted in August, about two months prior to the election[8]. The model

---

7. It generally takes a little over two years for state-level GDP to be released by the IBGE. Thus, for 2018, we calculated state GDP by multiplying the state average share of the national GDP (calculated for 2016 and 2017) by the 2018 national GDP.
8. We took the average of the August polls when there is more than one.

is estimated for all 7 elections (1994, 1998, 2002, 2006, 2010, 2014 and 2018), generating 189 observations (27 states over seven years). Given the panel structure of the data (repeated observations, election years, on the same cross section, states), we use random effects GLS with standard errors adjusted for clustering at the state-level[9].

The second model, inspired this time from Campbell (2008), also includes the real annual growth of the GDP from the year preceding the election to the election year (at the state level) as the economic activity measure but adds state-level trial-heat polls. The trial-heat polls are from IBOPE and were fielded in August[10]. Because state-level poll data are only available from 2006, the second model is only estimated for 2006, 2010, 2014 and 2018 election, for a total of 108 observations. Just like in the first model, coefficient estimates are obtained by random effects GLS with standard errors adjusted for clustering at the state-level.

The third and last model, for its part, is inspired from Lewis-Beck, Bélanger and Fauvelle-Aymar (2008). It deviates more from the original model than the first two in that it uses real annual growth of the GDP instead of the unemployment rate as the economic activity measure (for data availability at the state-level). In addition, the model includes the same popularity measure used in the first model. This last model is estimated for all 7 elections (1994, 1998, 2002, 2006, 2010, 2014 and 2018), producing 189 observations, and using, again, random effects GLS with standard errors adjusted for clustering at the state-level[11]. Descriptive statistics about our predictor variables are presented in the Appendix (Table A1).

RESULTS

Predicted vote shares for the incumbent candidates in the first round are presented in Figure 1. Figure 1 also indicates (with a star) the actual vote share received by the incumbent candidate, allowing for a visual assessment of each of our three model specifications. Across all three models, the average within-sample forecasting error is 5.55. Details about the accuracy of the forecast models are presented in Table 2 by model and election. First, it is important to note that *all* forecasts are within-sample forecasts. Just like in Turgeon and Rennó (2012), we find that Model 1 is the most accurate model. Model 1 predicts five of the seven

---

9.   Hausman tests indicate support for the random effects model over fixed effects.
10.   We used poll results from July or September, in that order, when no August poll are available. When there are more than one poll for August, we considered the average value.
11.   All models were estimated by weighting the data by state population. The model generates a predicted value by year for the incumbent for each of the 27 states, and these are averaged, in turn, to compute the final prediction, accounting again for state population.

elections within three percentage points, including two within one percentage point (1994 and 2010). This is quite impressive. Model 1 only performs badly for the 2014 when President Dilma Rousseff was seeking reelection with a forecast error of 6.81 percentage points. Interestingly, it performed quite well in 2018 with a forecast error of 2.87 percentage points even though this election represents a break in the PT-PSDB equilibrium that characterized Brazilian presidential elections for over 20 years. Overall, the average forecast error for Model 1 is 2.79 percentage points. Model 2, for its part, comes as the next most accurate model. It predicts nearly perfectly the 2010 election but fails badly at predicting the 2014 and 2018 election. Its average forecast error is one percentage point higher than that for Model 1, at 3.82. Model 2 has an additional downside, it cannot be used to forecast the 1994, 1998 and 2002 elections because trial-heat polls at the state level are not available for these elections. Finally, Model 3 performs poorly with average forecast errors of 7.18 percentage points. Finally, it is worth noting that the addition of the last three elections did not improve the perform of these models, as can be seen by comparing the average forecast error obtained from the 1994-2010 elections. To the contrary, the performance of all three models got substantially worst, especially for models 1 and 3.

Finally, Table 3 presents the coefficient estimates for the three models. For Model 1, we find, as expected, a strong effect for the Third Term Dummy, meaning that candidates/parties seeking a third term lose votes, on average. Economic growth is also statistically significant in Model 1 and has the expected sign: economic growth leads to a larger vote share for incumbent candidates. Economic growth is equally statistically significant in Model 2 and in expected ways but fails to reach conventional levels of statistical significance in Model 3. The Trial-heat variable in Model 2 is statistically significant and, not surprisingly, shows a very strong effect on the incumbent's vote share. Lastly, the popularity of the incumbent in Models 1 and 3 is statistically significant and exert, as expected, a positive effect on the incumbent's vote share. It is worth noting that the addition of the last three elections has proven beneficial. In Turgeon and Rennó (2012), economic growth was not statistically in all model specifications (it is now in two of the three) and the popularity of the incumbent had failed to reach statistical significance in Model 1. In sum, the addition of the last three elections has demonstrated the theoretical value of the three models although it did not improve their forecasting capabilities. It is worth noting that the regression estimates from the 2012 article and the ones presented in Tables 3 are not readily comparable because some of the data from the 2012 have been updated and a few minor errors were found in the original dataset.

In sum, we find that Model 1 outperforms the other models. This finding is important because it suggest that models that rely on the popularity of the incumbent as superior to those that rely on trial-heat polls once the fundamentals of the economy

are also taken into account. As for the uncharacteristically volatile 2018 election, our findings suggest that electoral forecasting models can survive such earthquake elections. The forecasts from both Models 1 and 3 are within 3 percentage points. This is quite surprising given how the 2018 election changed nearly entirely the equilibrium between the PT and PSDB that lasted for over two decades.

Figure 1. Forecasts for Brazilian Presidential Elections, by forecast models: 1994-2018 (with Fernando Haddad (PT) as incumbent in 2018)
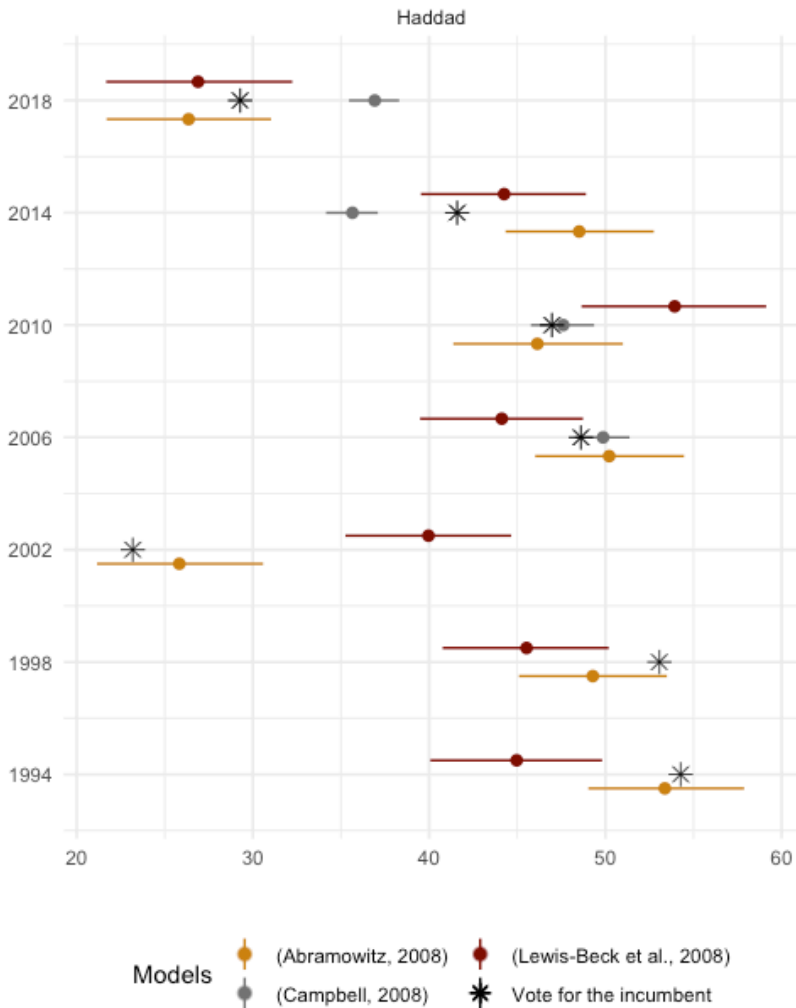
Source: Authors.

Table 2. Within-Sample Forecasts, by Model and Election

Election year	Vote for the Incumbent (%)	Forecast, in % (Forecast error, in %)		
		Model 1 (Abramowitz, 2008)	Model 2 (Campbell, 2008)	Model 3 (Lewis-Beck et al. 2008)
1994	54.27	53.31 (-0.96)		44.99 (-9.28)
1998	53.06	49.10 (-3.96)		45.52 (-7.54)
2002	23.19	25.69 (+2.5)		40.04 (+16.85)
2006	48.61	50.11 (+1.5)	49.82 (+1.21)	44.18 (-4.43)
2010	46.97	46.07 (-0.9)	47.31 (+0.34)	54.08 (+7.11)
2014	41.59	48.40 (+6.81)	35.66 (-5.93)	44.25 (+2.66)
2018	29.28	26.41 (-2.87)	37.06 (+7.78)	26.88 (-2.4)
Average (1994-2018)		2.79	3.82	7.18
(SD)		3.65	5.61	9.21
Average (1994-2010)*		1.21	0.13	1.27
(SD)		0.56	-	0.9

*Recalculated results for Turgeon and Rennó 2012 with new data. Values shown here differ from those in Turgeon and Rennó (2012) for two reasons. First, the analyses were conducted using updated data from the IBGE. Second, the authors found a few errors in the original dataset for some state-level values.

Source: Authors.

Table 3. Forecast Models for Brazilian Presidential Elections: 1994-2018

	Model 1 (Abramowitz 2008)	Model 2 (Campbell 2008)	Model 3 (Lewis-Beck et al. 2008)
Real annual GDP growth	0.456[*]	0.388[*]	-0.157
	(0.212)	(0.184)	(0.249)
Third term dummy	-19.565[**]		
	(2.027)		
Popularity of the incumbent	0.238[**]		0.183[**]
	(0.021)		(0.025)
Trial-heat poll		0.968[**]	
		(0.054)	
Constant	46.022[**]	0.321	43.261[**]
	(1.788)	(2.312)	(1.941)
R2 (overall)	0.7412	0.9896	0.5785
R2 (within)	0.9721	0.9896	0.9293
N	189	108	189
Log Likelihood	-785.221	-402.229	-824.305
Akaike Inf. Crit.	1,582.442	814.457	1,658.609
Bayesian Inf. Crit.	1,601.892	827.868	1,674.818

Entries are random effects GLS with standard errors adjusted for clustering at the state-level.

*p<0.05 **p<0.01 (two-tailed).

Source: Authors.

CONCLUSION

In this paper, we extended Turgeon and Rennó's (2012) earlier work on forecasting Brazilian presidential elections by incorporating data from the three most recent elections (2010, 2014 and 2018). Our findings show that the addition of recent elections contributes to improve the explanatory power of the electoral forecasting models, that is, their theoretical value, despite the inclusion of an election with a high degree of uncertainty (2018). Interestingly, the addition of the last three elections did not improve, however, their forecasting capabilities. This is true for models that are more heavily based on polls, but also for those with measures of the state of the economy and popularity of the incumbent. The recent 2018 presidential election in Brazil restructured the political landscape. The Worker's Party (PT) remains a central actor, but the PSDB has lost its prominent status as the anti-PT force. The new forces are now centered around Jair Bolsonaro since he defeated Fernando Haddad (PT) in the 2018 presidential election. This earthquake election was surprisingly well predicted by two of our three forecast models.

There remain important problems in using sub-national data to predict national results. First, economic data measured at the state level, as noted earlier, become available only two years *after* the election. To make election forecasts, one must estimate first a state-level measure of economic growth. Second, public opinion data is not systematically available at the state level, shortening the period that such models can use. The tendency is for this second problem to lessen with the passage of time. Also, as the number of elections increase, we will be able to properly incorporate time into our estimations, using not only state-effect, but also time-effect[12]. Also, it would be interesting to compare the forecasts from our models with those from vote intention polls conducted about three months prior to the election. We leave this task for a future update of this paper.

Finally, we are optimistic about our ability to forecast the upcoming 2022 election. First, it is very likely that the 2022 will have President Bolsonaro as the incumbent candidate. Second, the PT remains a very strong political force in Brazil and will certainly present to voters a strong candidate, most likely former President Lula. What might distinguish 2022 from prior elections, however, is the possibility of other credible anti-PT forces to compete for votes (at the expense of Bolsonaro). The center-right PSDB is expected to orchestrate a strong comeback, possibly led by São Paulo Governor João Dória Jr. The latter has gained some popularity since the beginning of the COVID-19 pandemic and has portrayed himself as the loudest critic of the Bolsonaro government. Therefore, there is a strong possibility that the

---

12.   We tested specifications using both state and time-effects, although we do not present the findings here because our data afford only 7 elections, fewer than the minimum suggested by Beck (2001) for panel/time-series-cross-sectional (TSCS) analysis.

2022 will have two strong candidates from the right —Bolsonaro and Doria— competing for votes. Will it be enough for the PT to make a comeback? Only history (and possibly our forecast models) will tell.

REFERENCES

Abramowitz, A. I. (2008). Forecasting the 2008 Presidential Election with the Time-for-Change Model. PS: *Political Science and Politics*, *41*, 691-695.

Auberger, A. (2010). Popularity and vote: forecasting the 2007 French presidential election. *Canadian Journal of Political Science/Revue canadienne de science politique*, *43*(1), 123-136.

Baker, A., Ames, B., & Rennó, L. (2020). Persuasive Peers. Princeton University Press.

Beck, N., (2001). Time-series–cross-section data: What have we learned in the past few years? *Annual Review of Political Science*, *4*(1), 271-293.

Berry, M. J., & Bickers, K. N. (2012). Forecasting the 2012 presidential election with state-level economic indicators. *PS: Political Science & Politics*, *45*(4), 669-674.

Bunker, K., & Bauchowitz, S. (2016). Pronósticos Electorales y Seguimiento de Opinión Publica en Latino América: Una aplicación a Chile. *Política. Revista de Ciencia Política*, *54*(2), 207-233.

Campbell, J. E. (1992). Forecasting the presidential vote in the states. *American Journal of Political Science*, *1*, 386-407.

Campbell, J. E. (2008). The Trial-Heat Forecast of the 2008 Presidential Vote: Performance and Value Considerations in an Open-Seat Election. *PS: Political Science and Politics*, *41*, 697-701.

Campbell, J. E. (2017). A recap of the 2016 election forecasts. *PS*, *Political Science & Politics*, *50*, 331-338.

Cantu, F.; Hoyo, V. and Morales, M. A. (2016). The Utility of Unpacking Survey Bias in Multiparty Elections: Mexican Polling Firms in the 2006 and 2012 Presidential Elections. *International Journal of Public Opinion Research*, *28*, 96-116.

Dubois, E. and Fauvelle-Aymar, C. (2004). Vote Functions in France and the 2002 Election Forecast. In M. Lewis-Beck (ed.), *The French Voter: Before and After the 2002 Elections* (pp. 205-230). Palgrave-MacMillan.

Duque, D, and Smith, A. (2019). The Establishment Upside Down: A Year of Change in Brazil. *Revista de Ciencia Política*, *39*(2), 165-189.

Foucault, M., & Nadeau, R. (2012). Forecasting the 2012 French presidential election. *PS: Political Science & Politics*, *45*(2), 218-222.

Jastramskis, M. (2012). Election forecasting in Lithuania: The case of municipal elections. *International Journal of Forecasting*, *28*(4), 822-829.

Jérôme, B., & Jérôme-Speziari, V. (2004). Forecasting the 2002 elections: Lessons from a political economy model. In M. Lewis-Beck (ed.), *The French Voter: Before and After the 2002 Elections* (pp. 178-204). Palgrave Macmillan.

Jérôme, B., & Jerôme-Speziari, V. (2012). Forecasting the 2012 French presidential election: Lessons from a region-by-region political economy model. *French Politics*, *10*(4), 373-377.

Jérôme, B., & Jerôme-Speziari, V. (2016). State-Level Forecasts for the 2016 US Presidential Elections: Political Economy Model Predicts Hillary Clinton Victory. *PS: Political Science & Politics*, *49*(4), 680-686.

Jérôme, B., Jérôme, V., Mongrain, P., & Nadeau, R. (2021). State-level forecasts for the 2020 us presidential election: Tough victory ahead for biden. *PS: Political Science & Politics*, *54*(1), 77-80.

Lewis-Beck, M. S., Bélanger, É. and Fauvelle-Aymar, C. (2008). Forecasting the 2007 French Presidential Election: Ségolène Royel and the Iowa Model. *French Politics*, *6*, 106-115.

Pereira, C. and Bertholini, F. (2018). Coalition Management in Multiparty Presidential Regimes. In B. Ames (ed.) *Routledge Handbook of Brazilian Politics* (pp. 313-330). Routledge.

Powell Jr, G. B. (2004). The quality of democracy: The chain of responsiveness. *Journal of Democracy*, *15*(4), 91-105.

Rennó, L. R. (2020). The Bolsonaro voter: Issue positions and vote choice in the 2018 Brazilian presidential elections. *Latin American Politics and Society*, *62*, 1-23.

Samuels, D. and C. Zucco. (2018). *Partisans, antipartisans, and nonpartisans: voting behavior in Brazil*. Cambridge University Press.

Toros, E. (2012). Forecasting Turkish Local elections. *International Journal of Forecasting*, *28*, 813-821.

Turgeon, M. and Rennó, L. (2012). Forecasting Brazilian presidential elections: Solving the N problem. *International Journal of Forecasting*, *28*(4), 804-812.

APPENDIX

Figure A1. Forecast Models for Brazilian Presidential Elections: 1994-2018. Scenario: Meirelles (PMDB) as incumbent in 2018
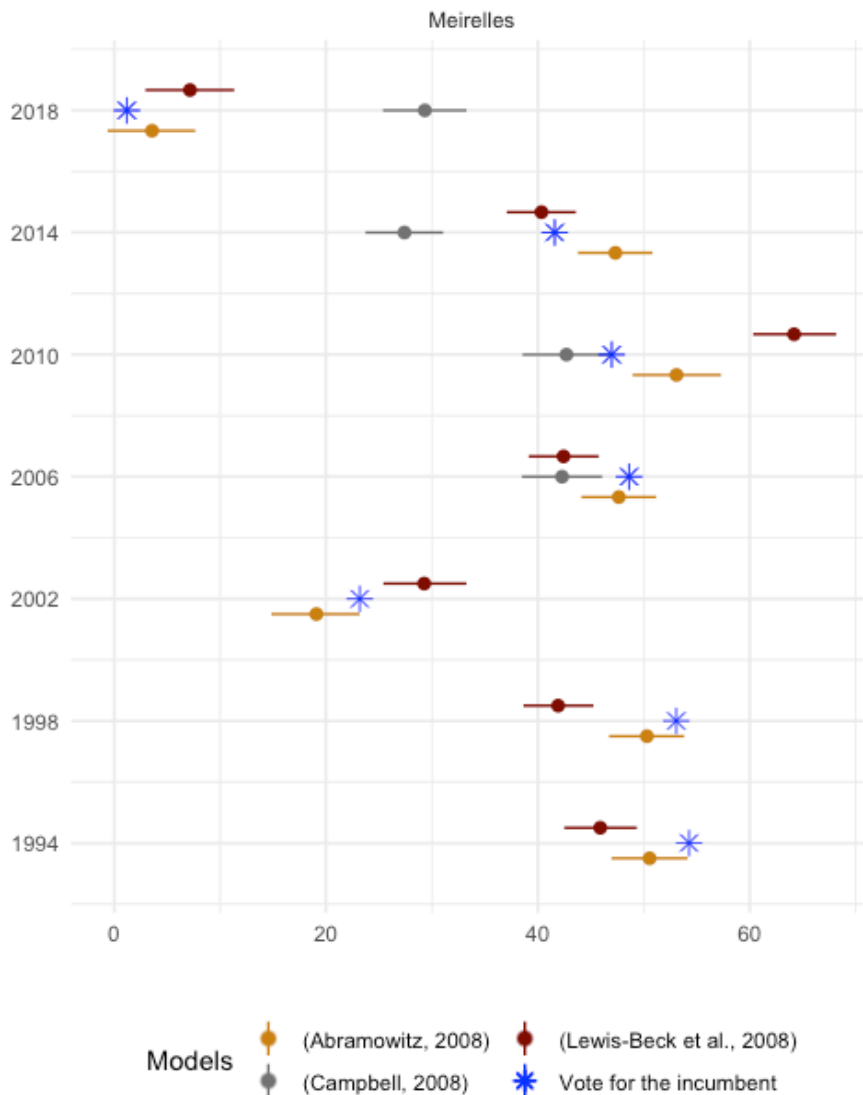
Source: Authors.

Table A1. Descriptive statistics of the predictor variables

	Mean (SD)	N
Real annual GDP growth	2.73 (5.16)	189
Popularity of the incumbent	9.57 (42.6)	189
Incumbent vote	44.4 (17.2)	189
Trial-heat poll	45.0 (14.4)	108

Source: Authors.