MASTER

APPLIED ECONOMETRICS AND FORECASTING

MASTER'S FINAL WORK

DISSERTATION

FORECASTING ELECTIONS IN A MULTIPARTY SYSTEM: THE CASE OF PORTUGAL AND BRAZIL

FERNANDO CARLOS ARAÚJO RAWICZ

SUPERVISION:

PROF. DR. RUI PAULO

JANUARY - 2022

Abstract

This work tries to forecast election results in Brazil and Portugal using two bayesian models and one frequentist in order to find out which one has better results. We will use older election‘s results and polls in order to check if there are sistematical biases towards certain parties. We also use macroeconomical data to check how influential this data is to forecast election. The analysis pointed out that there are no sistematical biases for any party in any polling company. We also found out that there is no significant relationship between macroeconomic data and the election results in these countries. Furthermore, the fact that both examples had few elections and have a lot of parties which are constantly being created and dismissed, there is not a "perfect" model, however, they all have very acceptable results.

Key Words: Elections, Forecasting, Bayesian, Portugal, Brazil.

Contents

ACKNOWLEDGEMENT

I want to first acknowledge Professor Doctor Rui Paulo for all the help and ideas througout this project. This work would not be possible without you.

I also want to thank my father, Flávio Rawicz, for all the support in my journey to Portugal. There is no way I could possible love you more. I would also like to thank Anna Zahri Paz for teaching me bayesian statistics and Tomas Neves for being a good friend in Portugal.

Finally, thank you for everything Vitão. We will always remember you.

"There’s no such thing as a vote
that doesn’t matter."

Barack Obama

1 INTRODUCTION

There is a large diversity of events which people try to forecast. These go from stock values and returns, football matches and championships winners, reality shows results and even extreme events like earthquakes. But there is no doubt that one of the most important and impactful events are the political elections. Their results affect the economy, the legal system, the diplomacy, and the results of the elections from one country might have impact all over the world. For this reason, companies, countries and individuals try to forecast electoral results. This is the subject of this thesis.

Usually, scientific articles about election forecast focus on the Presidential elections of the United States of America (US). This is understandable, since the US is arguably one of the most important democracies of the world. However, the american elections have some peculiarities and most of the models are tailored to to deal with these, which include bipartisanship and the high impact of results of the elections in each state. Additionally, there is a lot of data concerning these elections: the US is an old democracy and they always kept a good tracking of macroeconomic and other kinds of data. They are stable economically and politically, at least much more stable than other countries like the ones that are the subject of our analysis, namely, Portugal and Brazil.

Given that, we present the modeling of the Portuguese and Brazilian elections highlighting the difficulties encountered. For example, the fact that they are both fairly recent democracies, and as a result have held a relatively small number of elections. This makes it difficult to fit structural models, which will be discussed in Section 4.3, since the sample size is small. It is important to notice that the nomeclature "structural" is not related to the usual econometric use. It is used in the same was as in (Linzer, 2013), it will be better explained also in section 4.3. Another difficulty stems from the fact that in these countries there are more than two parties, in fact, there are several important parties. Additionally, along the decades, a number of parties has been created but also dismissed. Finally, we have economic instability, which affects every structural model, especially the Brazilian one, which had in one election a hyperinflation scenario with more than 100% and, in the next one, the inflation was at a much lower level.

Both in Portugal and in Brazil, the ability of the polling houses to produce independent and bias-free projections is often questioned. We constructed a statistical model to address these claims in the context of the presidential elections, and found no evidence of a structural bias in favor of a particular party by any of the polling houses.

In the end, we consider three differente models, one in the frequentist framework, which is based on ARIMA modeling, and two bayesian ones. The first one in the second framework in the Marginal Dynamic Linear Model, which considers only the individual time series of the vote share intention evolution from the polls for each candidate. The

final model in the Joint Dynamic Linear Model, in which we consifer all the candidate's vote share together in a multinomial based model. This model could incorporate the systematic polling company bias for each party and the prior information for the vote shares in the structural model. However, as we will see, the structural model is not statistically significant and there is also no significant bias.

2 HISTORY AND ELECTORAL SYSTEMS

As stated, most of the work on forecasting electoral results deals with the presidential elections in the United States. This is due to a variety of reasons, but most prominently because the US is one of the influential democracies on the planet, and as a consequence the whole world is watching what happens there. The outcome of the presidential election impacts governments, companies and people in general, so everyone needs to be prepared for what may occur as a consequence of said results.

There are positive aspects to forecasting American elections. The United States held its first presidential election in 1788-89 and the quadrennial elections were held ever since. This means that we have a lot of data to inform statistical modes. Another positive aspect is bipartisanship, which has been present for almost all recent elections and makes it easier to model the results.

Although it is important to forecast the north american elections, this is not the goal for this master's thesis. Instead, we focus on Portugal and Brazil. For this reason, we will start with a brief explanation of how these electoral systems work and some background history.

2.1 Portugal

In this subsection, we will only make a quick summary of the recent democracy in portuguese history. All the historical information and terms were used based on Mata and Valério (2011).

Portugal got its freedom from the fascist regime on April 25, 1974. This was the beginning of the Third Republic [1] , which is the current regime. The Constitution was promulgated on April 2nd, 1976. And it was in this same year, on June 27, that the country held its first presidential election.

In the Portuguese Political System, we have the President of the Republic as chief of state and the Prime Minister as chief of government. According to the Portuguese Constitution, the Prime Minister is nominated by the President, while taking into account

[1]The oficial name of the curent regime in Portugal is just Portuguese Republic, but some authors, like Mata and Valério (2011) use Third Republic.

the electoral results.

In this work, we will only look at the Presidential election. For this reason, it is important to understand the presidential constitutional duties. According to the Portuguese constitution, in a free translation, "The President of the Republic represents the Portuguese Republic, guarantees its national independence, the unity of the State and the regular functioning of the democratic institution and is, inherently, the Supreme Commander of the Armed Forces."

The election is based on a universal suffrage for Portuguese people aged 18 and above. The candidate is declared President if he or she has more than half of the valid votes (50% plus 1 vote). If none of the candidates has this amount of votes, there is a second suffrage ("segunda volta" in European Portuguese) where only the two most voted candidates will be on the run.

The President can only run for one reelection, that is, two consecutive mandates of five years each. When 15 years have passed since when the term is over, he or she may run again for a third term. This has never happened.

Since 1976 there were 10 presidential elections in Portugal. Only one had a second suffrage, in 1986, as we are going to see.

This regime in Portugal has a difficulty in terms of forecasting in comparisson of regimes where the president is the chief of government, since the president may not be seen as the person responsible for the country's economic or social results, be they positive or negative.

Portugal has a peculiarity, which makes the analysis of the data much more difficult: independent candidates are allowed. Sometimes they can be supported by a party or by a group of parties. The first election was an example of this; three of the four candidates were independent. The two most voted were António Eanes - supported mainly by the Socialist Party (PS) and the Democratic Social Party (PPD now PSD), the main players in the Portuguese political scenario, and Otelo Saraiva de Carvalho, who was supported by several left and far left-parties. Eanes won with 61,59%.

The 1980 election was at the end of the year in which Portugal started looking more closely into entering the European Economic Community. Eanes ran with the support of some parties, including a good portion of the socialists - except for Mário Soares, who at the time was secretary-general of the Socialist party. Eanes won again in the first round, this time with 56.44%.

The first half of the 1980s was marked by the success of the stabilization policies, the effect of the dollar devaluation in the short term, falling interest rates, and increasing integration in the European Community. Even with a good economic situation, the elections held in January 1986 were not easy and we had the only second suffrage after the

re-democratization. In the first round, Diogo Freitas do Amaral, a candidate supported by center-right and right parties had 46.3% while the left party votes were divided between Mário Soares (25.4%), Francisco Salgado Zenha (20,9%) and Maria de Loures Pintasilgo (7.4%). All the values are rounded. But in the second round, with only the first two candidates, almost all the votes went straight to Mário Soares, who won with 51,18%.

The economic prosperity continued in the second half of the 1980s. The problems in foreign payments were over and there were a lot of public investments like the basic highway network. These were all aimed to make Portugal join the European Monetary System, which finally happened in 1992. But one year before, in January 1991, Mário Soares was again candidate and had even the support of the former center-right Prime Minister Cavaco Silva. He won with 70,35% of the votes easily.

In 1996, the economical and political scenarios were less favorable and happy than before. According to Mata and Valério (2011), the early 1990s were marked by an economic downturn. This downturn was felt in the whole world but it was worse for Europe, which also had to deal with the short-term negative effects of monetary union and the consequences of the German unification, which made the monetary union much more difficult for some time. The GDP growth was upper-bounded and the Portuguese currency at the time, the Escudo, was devaluated in the conversion to prevent speculative attacks. The elections were held in the middle of all this. With Mário Soares unable to run again, the PS candidate, Jorge Sampaio, ran against the former Prime Minister Cavaco Silva. They were the only candidates and the final results were much closer than the 1991's. PS won again with 53,91% of the votes.

After the downturn in the early 1990s, the second half of the decade was better for the economy. In order to enter the European Monetary Union, Portugal had to reduce its public debt and continue the privatization process — remembering that this was with a socialist government. This all had a good impact on the economy which was also glooming worldwide. In this scenario, the 2001 elections were held. Following the usual Portuguese tradition, Jorge Sampaio was easily reelected with 55,76% of the votes; Joaquim Ferreira do Amaral, the second most voted, had only 34.54%.

The beginning of the 2000s was marked by the completion of the monetary union, and in 2002 we had the euro in the European Union. But it was not an easy period neither economically nor politically. The situation had the extreme in 2004 when the President, Jorge Sampaio, dissolved the Parliament. It had a great impact on the PS public opinion. When the elections were held, in January 2006, the left was fragmented, and for this reason, Aníbal Cavaco Silva - a former Prime Minister - found a clear path to end the decades of PS Presidents in the chair and won by 50.54% of the votes.

The economic situation was not much better after the new President got the chair.

The crisis of 2008 affected harshly the country to the level of being unable to refinance its debt. Two major Portuguese banks, the Banco Português de Negócios (BPN) and Banco Privado Português (BPP) went bankrupt and the government had to bail them out. Portugal had to take up a lot of austerity measures in 2010 with tax hikes and salary cuts for public servants. Even in this situation, Cavaco Silva was reelected with 52.95% of the votes.

The 2010s saw a little recovery in the economy. In 2012 Portugal was able to regain access to financial markets and the general situation was a little better. In this situation, it was easy for the center-right to elect a new president, Marcelo Rebelo de Sousa, with 52% of the votes in January 2016.

The most recent election, in January 2021, will be used as the Portuguese example for forecasting in the thesis. The election was held in the middle of the Covid-19 crisis. With a good public evaluation in general of how he handled the pandemic, Marcelo Rebelo de Sousa followed the tradition and was reelected with 60,66% of the votes.

In Figure 1, it is possible to see that although the very first poll on December 15 had some discrepant results. it was, however, more than one month before the election. After the first poll, the were no large movements in polled public opinion. That is very different from the brazilian case, as we will see next.

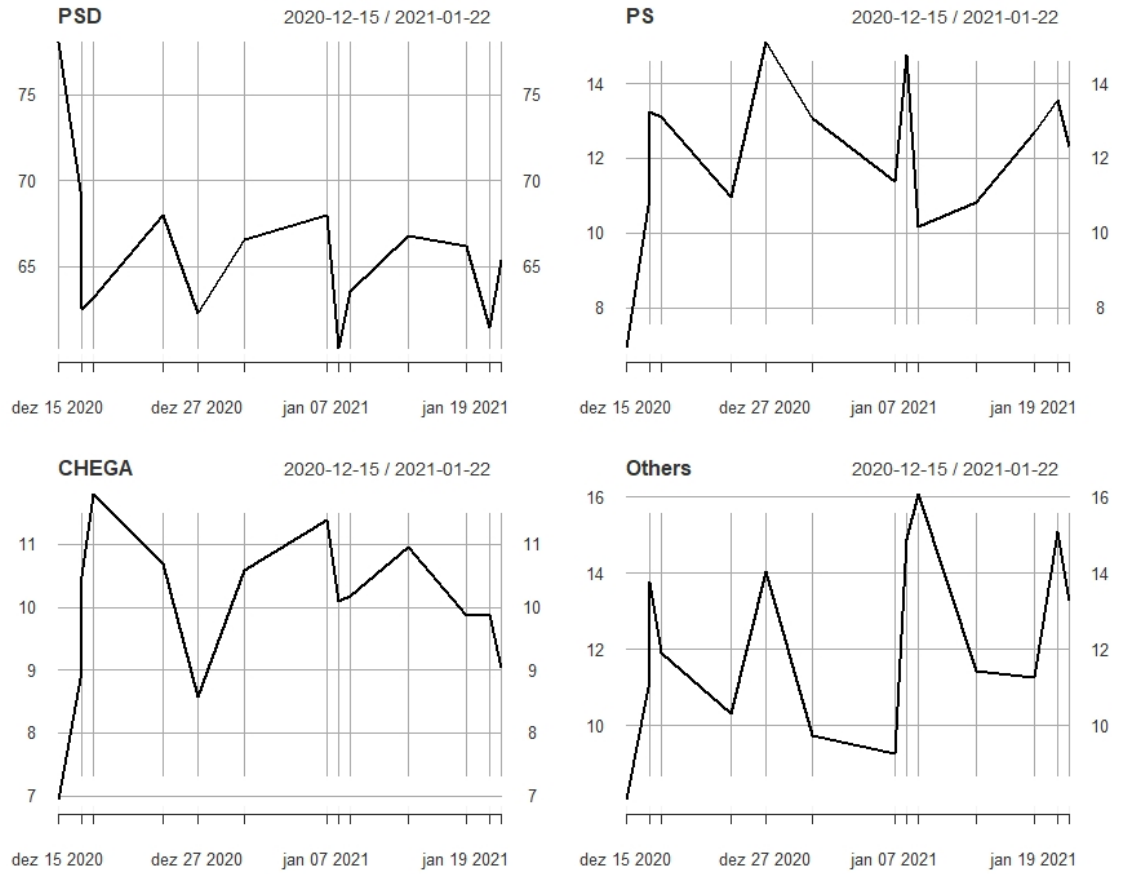

FIGURE 1: Evolution of the polls for the main candidates in Portuguese 2021 election in % (valid votes)

2.2 Brazil

Just like Portugal, Brazil is a recent democracy. The country got its freedom from a Military Regime after a "slow and steady" political openness in which the citizens were slowly getting their rights back.

One of the first acts towards democracy was to allow the return of political exiles. After that, local governors were able to be elected directly and legal multi-party system was once again legal. Finally, in 1988, the new constitution was promulgated by the parliament, and in 1989 Brazil had its first direct presidential election since the military coup in 1961.

Unlike Portugal, there is no prime minister, the president is both chief of executive power and chief of state. The brazilian president has a lot of attributions which are all listed in the article 84 of the brazilian constitution.

The electoral run is very similar to the portuguese one. The candidate with more than half of the valid votes becomes the president. If none of the candidates have this amount, there is a second suffrage ("segundo turno" in brazilian portuguese) with the two most

voted candidates. A elected president can only run for reelection, that means only two consecutive mandates are allowed.

Since 1989, there were 8 presidential elections in Brazil. Only two of them didn't have a second suffrage and two impeachments were performed.

The first presidential election was the first vote for many Brazilians and the first try of many candidates. More than 20 candidates tried to reach the presidential chair. But the main four were Fernando Collor, from Partido da Reconstrução Nacional (PRN), a young politician who tried to be anti-politics, Luiz Inácio Lula da Silva, from Partido dos Trabalhadores (PT), a famous leftist and unionist, Leonel Brizola, from Partido Democrático Trabalhista (PDT), leftist and famous in some Brazilian States, namely Rio Grande do Sul and Rio de Janeiro, states where he was governor, and Mário Covas, from Partido da Social Democracia Brasileira (PSDB), a more intellectual centrist candidate, whose ideas were considered center-right by some and center-left by others. The scenario was not good, there was hyperinflation in place for some years and a lot of failed plans tried to get it back to a normal index.

Collor and Lula went to the second round and, after a hard and controversial campaign, Collor was the winner. But only two years later, he resigned from the role after some scandals and right after that he was impeached and lost his political rights for eight years. His Vice-President, Itamar Franco, took the role and put in place an ambitious plan to stop the hyperinflation, the "Plano Real", which greatly succeeded. His Minister of Economy, Fernando Henrique Cardoso, also called FHC, was one of the main faces of this plan.

In 1994, mainly FHC and Lula ran for the presidency. But with the popularity gained from the stabilization of the inflation, FHC won with 54.24%. In 1998, even after some scandals, like accusations from trying to buy the congress to allow reelection — that's why it was 4 years after, not 5 like between 1989 and 1994 — FHC won again easily — with 53.06% in the first round.

The 2002 election was a time for a change of scenarios. Lula ran again with a more centrist and pro-market mentality and campaign. After a series of crises in the early 2000s, the PSDB lost some popularity and Lula easily won. In 2006, Lula was easily reelected with high popularity even after the corruption scandal called "Mensalão".

In 2010, Lula's popularity was one of the highest ever seen. After some social programs and even being called by Obama "The most popular politician in the world", he could easily choose a successor and the choice was Dilma Rousseff. With a not that difficult campaign, she won against José Serra, the same candidate who lost against Lula in 2002.

Dilma ran again in 2014. With a bad situation economically and lower popularity, the whole run was very close against Aécio Neves, a center-right candidate and former

governor and senator for Minas Gerais. Rousseff's image was dilapidated by a series of protests which happened mainly in 2013. In the end, the result was very close, close enough to raise some suspicions and create a even higher and more violent polarization process in Brazil. This polarization became even greater when Rousseff was impeached in 2015 with an amazingly low popularity. Her Vice, Michel Temer, took the chair.

The 2018 election was the high point of this polarization. With Lula being arrested some months before the election, the candidate for a replacement chosen by PT was Fernando Haddad, a former mayor of the city of São Paulo who made his entire campaign based on the assumption "voting for me is the same as voting for Lula". Meanwhile, PSDB and everyone from the center-right lost space to Jair Bolsonaro, a candidate who started growing in popularity a few years before with some extremist and populist ideas. But these ideas found a place in a country plagued by a bad economic situation, devaluation of currency, and high "antipetismo", the name given to the PT aversion. Jair Bolsonaro was the first candidate to be in the top two who was not from neither PT nor PSDB since Collor in 1989.

One of the interesting things to keep in mind for this election is the graph for the time series for the main candidates, as you can see in Figure 2

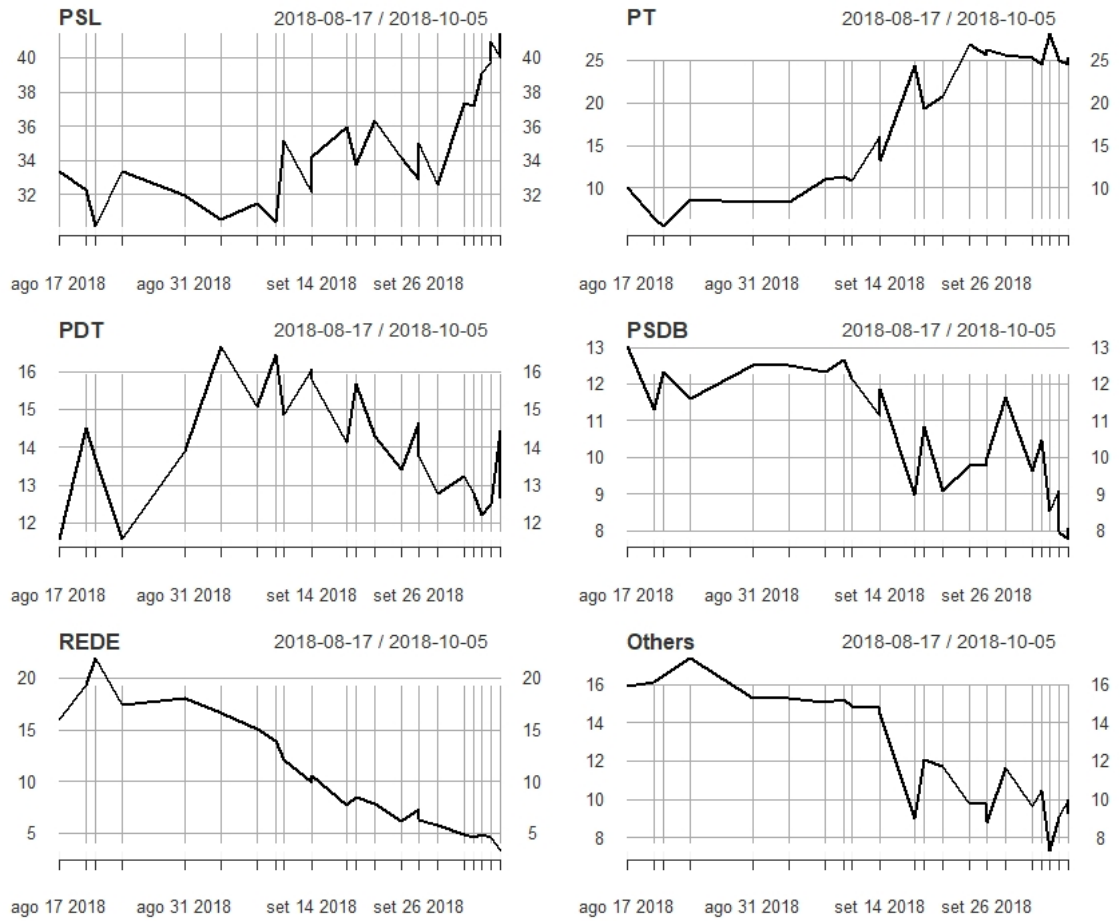

FIGURE 2: Evolution of the polls for the main candidates in Brazilian 2018 election in % (valid votes)

That is the story of how polarizing the election was. Some traditional candidates and parties, usually considered more "center" in the political spectrum, like PSDB, PDT and Marina Silva, a traditional runner who this time ran for a new party, REDE, lost their votes to the right wing or far right wing Bolsonaro (PSL) or the leftist Workers Party (PT). They went to the second round and, in the end, Bolsonaro won.

2.3 United States

A short description of the United States electoral system is necessary. Not because it will appear in our model, but because most of the statistical models deal with forecasting this specific presidential election. This is the case for the famous paper Abramowitz (2008) and also for the recent model used for example for "The Economist", Linzer (2013).

All the information here will be a summary of the information available at the White

House website [2] . In the american democracy, the executive branch of the federal government is vested in the president of the United States. Similar to the brazilian system, the President is also the Head of state and commander in chief of the armed forces.

The presidential election is held every four years in November. But there is a big difference: every xtate and local governments are allowed to conduct the specificities of their elections. This happens because there is an indirect system to elect the president called "electoral college".

In this system, every state of the USA and the District of Columbia are allowed to have some of the 538 electors and at least 270 are necessary to elect the president. The number of electors for each State is distributed according to the number of voters, with three being the lowest amount possible. In almost all states, the winner of the state takes all the electors.

Another difference is that, despite the fact many parties are allowed, there have been two which have been responsible for most of the votes at least in the last one hundred years, the Republican party and the Democratic party. This is well explained in Mellow and Trubowitz (2005). Most of the literature about forecasting american elections ignores any other party and treats everything as if the voting was in one of the two parties.

2.4 Summary of the electoral systems

In the end we have three distinct scenarios and it is important to make this short summary about the them.

United States:

Bipartisanship with every victory in each state being more valuable than the overall number of votes.

Portugal:

Multipartisanship with many independent candidates. Despite this, almost all elections were decided in a single sufragy. Overall number of votes decides the president.

Brazil:

Multipartisanship with almost all of the elections being decided in the second sufragy. Overall number of votes decides the president.

3 STATISTICAL ANALYSIS OF ELECTORAL DATA

In this chapter, we are going to describe what kinds of data are available for Brazilian and Portuguese elections, and will also present a literature review on the material available for statistical analysis of electoral data or for forecasting electoral results.

[2]The White House website is available at (https://www.whitehouse.gov/, n.d.)

This chapter is also important because there are many models being fit and they have their own specificities. This chapter is useful to clarify some of the difficulties.

3.1 Literature review

The first attempts at using statistical or econometric models to forecast electoral results date back to the late 1980s; see, for example Abramowitz (1988), an author who continued to be important in the field. The literature on the subject had a fair growth in the 1990s with some important articles, like Abramowitz (1996) and Lewis-Beck and Rice (1993). These models were heavily based on economic and microeconomic theory and tried to find relations between economic and social variables with the outcome of the US elections. Abramovitz continued doing this exercise and one of his works – Abramowitz (2008) – is still one of the most cited articles in the area.

It is only in 1996 that we find the first article which incorporated an autoregressive term but for the results of the two previous elections, alongside with some economics and social variables — see Norpoth (1996).

In the 1990s we had also some instances of incorporating poll data in the forecast; one of the most important in the genre is Holbrook and DeSart (1999), which combined poll data, and the results of the previous election, in their model.

So far, we described models which used time series techniques, economic and social theory, and some which used poll data. Only one piece of our puzzle in this thesis is missing: the bayesian framework. This last piece is found in Linzer (2013), who combines the evolution of the polling in a Dynamic Linear Model with a "structural" model based on, again, economic and social variables for the mean. This specific model and also Abramowitz (2008) are going to be better explained in Section 4. Linzer's model is widely used and is even the basis for the model used in the famous magazine "The Economist".

All the literature described so far describes tools to forecasting election results in a bipartisanship scenario. Most tools are really only applicable in the context of a stable and old democracy, for which there is a lot of data. This is far from the reality for most the democracies. The article Stoetzer, Neunhoeffer, Gschwend, Munzert, and Sternberg (2019) tries to use a multinomial-based model, also based on Linzer's, to try to overcome the bipartisanship restriction. In Section 4, we will explain this model. Our own model, described in Section 4.5 is inspired by this model; we will be additionally incorporating polling houses' bias.

3.2 Types of data

For modeling both Portuguese and Braziilian elections, we have mainly four kinds of data. The first one is the economic and social data, with historical data of GDP, inflation

rate, and the unemployment rate for both countries. This will be used in the "fundamentals based model" or "structural model", which will be discussed in subsection 4.3, where we also discuss the nomenclature for the model.

The second type of data we are going to use is the results of the elections from past years. Not only the presidential elections but also regional elections and legislative elections for both parties. This will be important for both the structural model in subsection 4.3 and the polling bias model in subsection 4.5.

The third type of data is the evolution of the polls for both countries in the elections of 2018 in Brazil and 2021 for Portugal. We are always going to use the forecast of the valid votes. This means that we are proportionally dividing the undecided voters and voters who are not going to vote for any candidate. This is because the final result of the elections counts only the valid votes in order to decide the winner. This data is the most important one in this work and it will be the basis for the Frequentist Model in section 4.1, the Marginal Dynamic Linear in section 4.7 and the Joint Dynamic Linear Model in section 4.9.

The last data type is the last poll before the elections for the elections on the past years, the ones from the second type. This is of extreme importance for the model in subsection 4.5, In which we will use the comparison of the election results with the last poll to model the bias for each party and each polling company.

The data from the polls were gathered in each polling company site for the brazilian case and in the ERC [3] (Entidade Reguladora de comunicação) website for the portuguese case.

3.3 Coding and Analysis

All the models were developed using R programming ((R. C. Team, 2017)) language in RStudio (R. Team (2020)). The main packages used to fit bayesian models were Rjags (Plummer (2019)) and RStan (Stan Development Team (2020)). All the codes and data used are available at the GitHub page Rawicz (2021) [4] .

4 STATISTICAL MODELS AND RESULTS

In this section we are going to present all the models and their individual results. The choice to present the result this way is because there are a lot of models and this way the rationality behind each one and the choices regarding them are more clearly presented. In the sections 4.1 and 4.2 we are going to present the frequentist model and show its results.

[3] www.erc.com

[4] https://github.com/frawicz/Elections

It is the only main model which uses the frequentists framework.

After that, we are going to present the two models that are not going to be used to forecast the elections, but instead they would be used as inputs for the Joint Dynamic model. These are the Structural Model - section 4.3 - and the Polling Company bias model - section 4.5. But as we are going to see, the results of the Structural model are not valid statistically, therefore its results are going to be ignored. For the Polling Company bias, the results are good, but they fortunately show that the bias for every party and every company amongst the ones analyzed are also not significant, therefore we will keep the bias as a zero mean process prior, as we will explain better in the section 4.9.

After we are going to present our both bayesian models, the Marginal Dynamic in section 4.7 and the Joint Dynamic in section 4.9. The first one considers each series individually for each candidate as a common DLM and the second one considers all the series of every candidate as one process coming from a multinomial DLM. The joint model was the one which could use the data from both Structural and Polling company bias model.

4.1 Frequentist Model

The first model we are going to fit is in the frequentist framework, the ARIMA modeling. Just like we will see in the subsection 4.7, this approach takes each evolution of each pary vote share intention individually. The approach we are going to use is the "Box-Jenkins method". It was defined in Box, Jenkins, and Reinsel (1994). We will follow three steps in order to identify which ARIMA model is more appropriate for each of the series of the candidates:

1. Model Identification
    - Identifying stationarity
    - Identifying possible AR and MA orders
2. Model Estimation
    - Estimate through Maximum Likelihood
3. Model Validation
    - Test Residuals
    - Compare probable models

But before doing this we have to face before another problem: the missing data. While in the bayesian paradigm this is not a big issue when dealing with dynamic linear models,

since we can can input the missing data as part of the fitting process, "borrowing strength" from our priors. In the frequestist approach, however, we will have to deal with this, since our model is based on daily data, but a lot of days in the campaign do not have data.

On the section 4.2 we can clearly see that the frequency of the polls gets higher when the election gets closer, which is the expected. In order to solve this issue, we are going to use an interpolation method based on a ARIMA process that is automatically calculated. All this process is well described in Moritz and Bartz-Beielstein (2017).

4.2 Frequentist Model Results

The first step in this frequentist analysis is the interpolation. It allows us to estimate interpolated data for the dates in which we don‘t have a poll. In the following figures 3 and 4 it is possible to see that we have a lot of missing data for the first days of campaing and, as the election day approaches, the frequency gets much higher. This happens for both Portuguese and Brazilian cases.

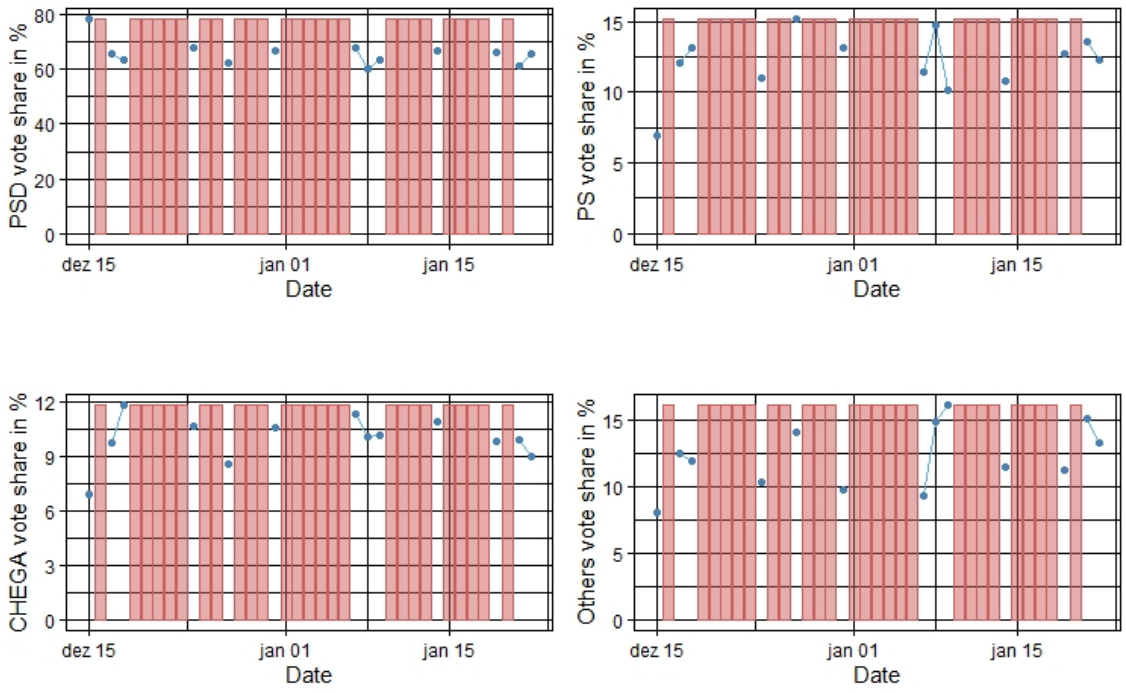

FIGURE 3: Missing data for Portuguese parties poll vote share of 2021 election

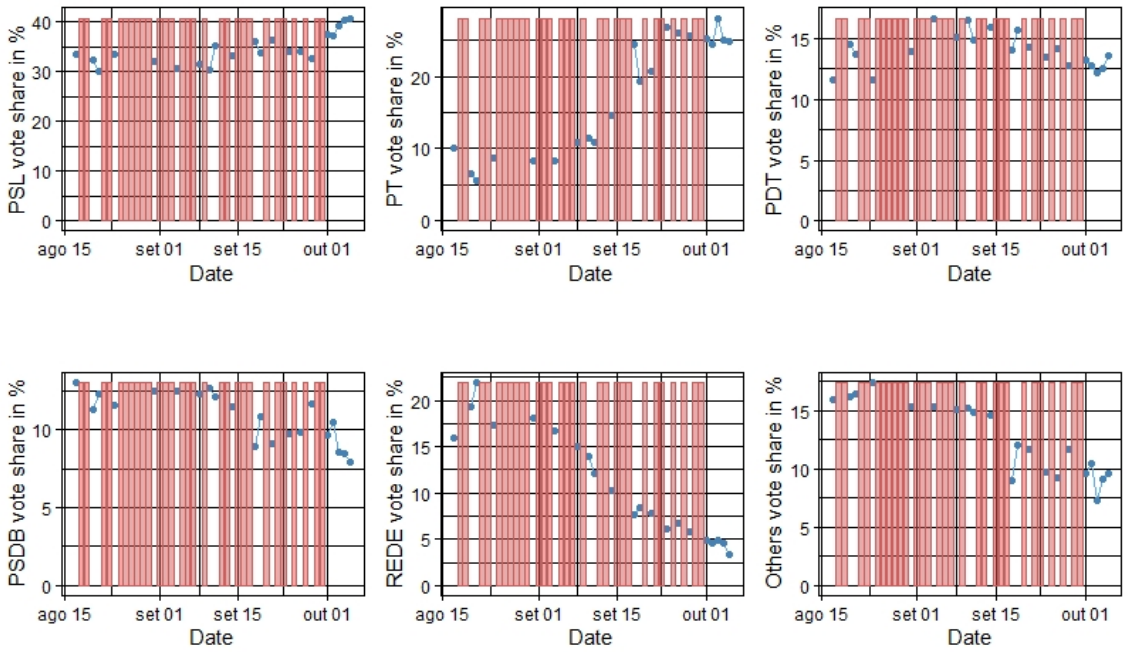

FIGURE 4: Missing data for Brazilian parties poll vote share of 2018 election

After, in the figures 5 and 6 we can see how interpolation works for both cases. It predicts a simple ARIMA process which clearly do not make huge assumptions on the processes and tries not to affect the final modelling heavily. As we are going to see further, the processes chosen are also not complex ones.

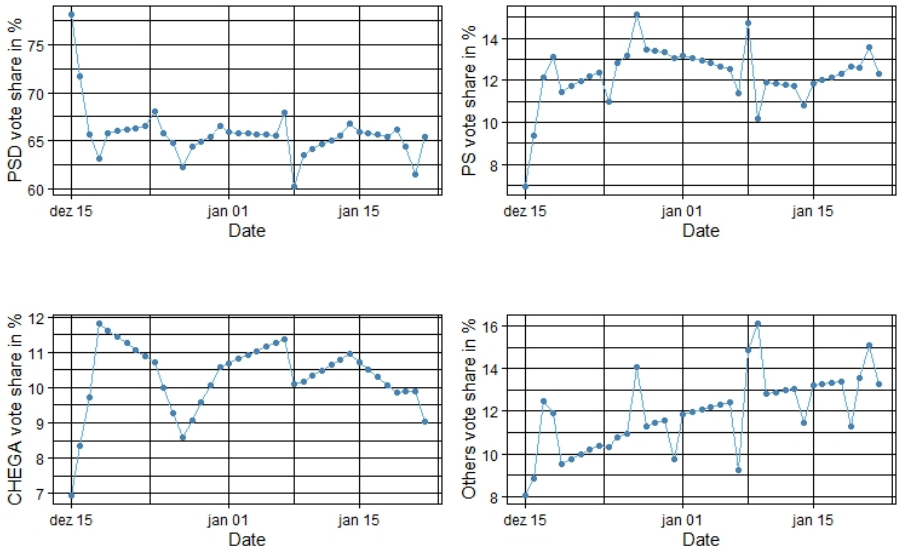

FIGURE 5: Polling data with ARIMA interpolation for portuguese polling

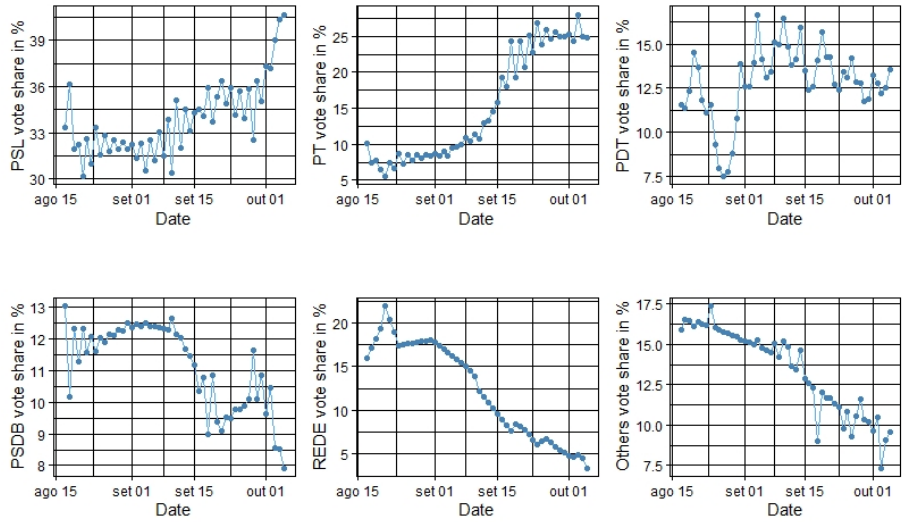

FIGURE 6: Polling data with ARIMA interpolation for brazilian polling

With these interpolated series we start the analysis for each country. First we will check which models came out from the ARIMA modeling and after we are going to forecast them.

Portugal

As we can see in table I and in the figure 7, all the process are very simple ones. But we are going to see in the section 5 that even with these simples models, we have good prediction power.

Party	Model
PSD	ARIMA(1,0,0)
PS	ARIMA(1,0,0)
CHEGA	ARIMA(2,0,0)
Others	ARIMA(0,1,0)

TABLE I: Automatic ARIMA results for the each of the corrected series in the portuguese 2021 election

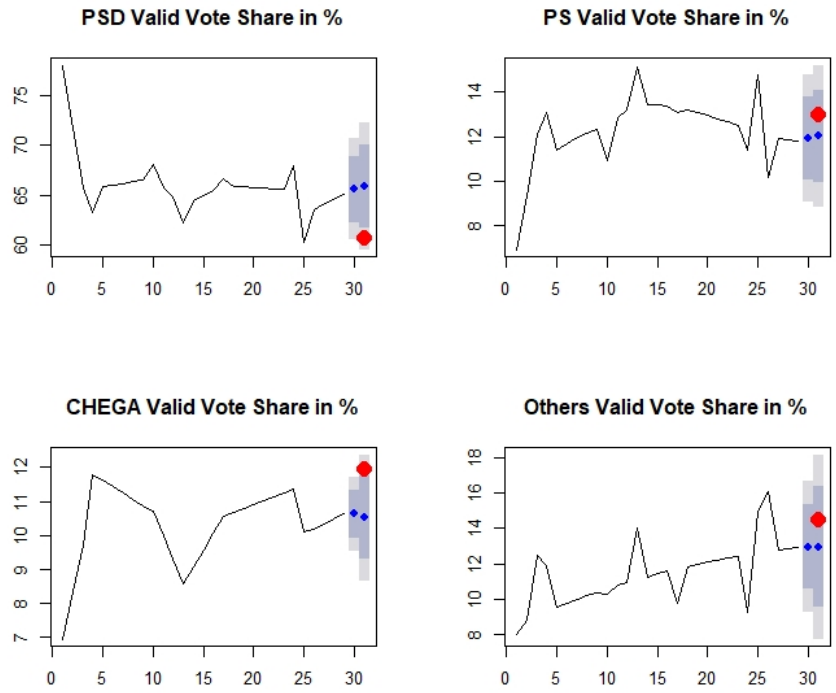

FIGURE 7: ARIMA modeling for the Portuguese 2021 elections

Here the lines represents each paty‘s time series and the red dot is the true election result.

Brazil

Now for the Brazilian case, we can see that the models are usually more complex than the Portuguese ones. This happens because the 2018 elections had some greater movements and it is indeed more difficult to understand how the vote share intention is behaving. Also, there is the bigger difference that all the models here are I(1), while they were I(0) for the Portuguese elections. The results are presented in the table II and in the Figure 8. Also, we can see that PSL and PT candidates had a fair growth during the campaing, followed a little by PDT's Gomes, while all the other parties faced a downward turn.

Party	Model
PSL	ARIMA(3,1,0)
PT	ARIMA(3,1,0)
PDT	ARIMA(0,1,0)
PSDB	ARIMA(2,1,2)
REDE	ARIMA(1,1,0)
Others	ARIMA(0,1,1) with drift

TABLE II: Automatic ARIMA results for the each of the corrected series in the brazilian 2018 election

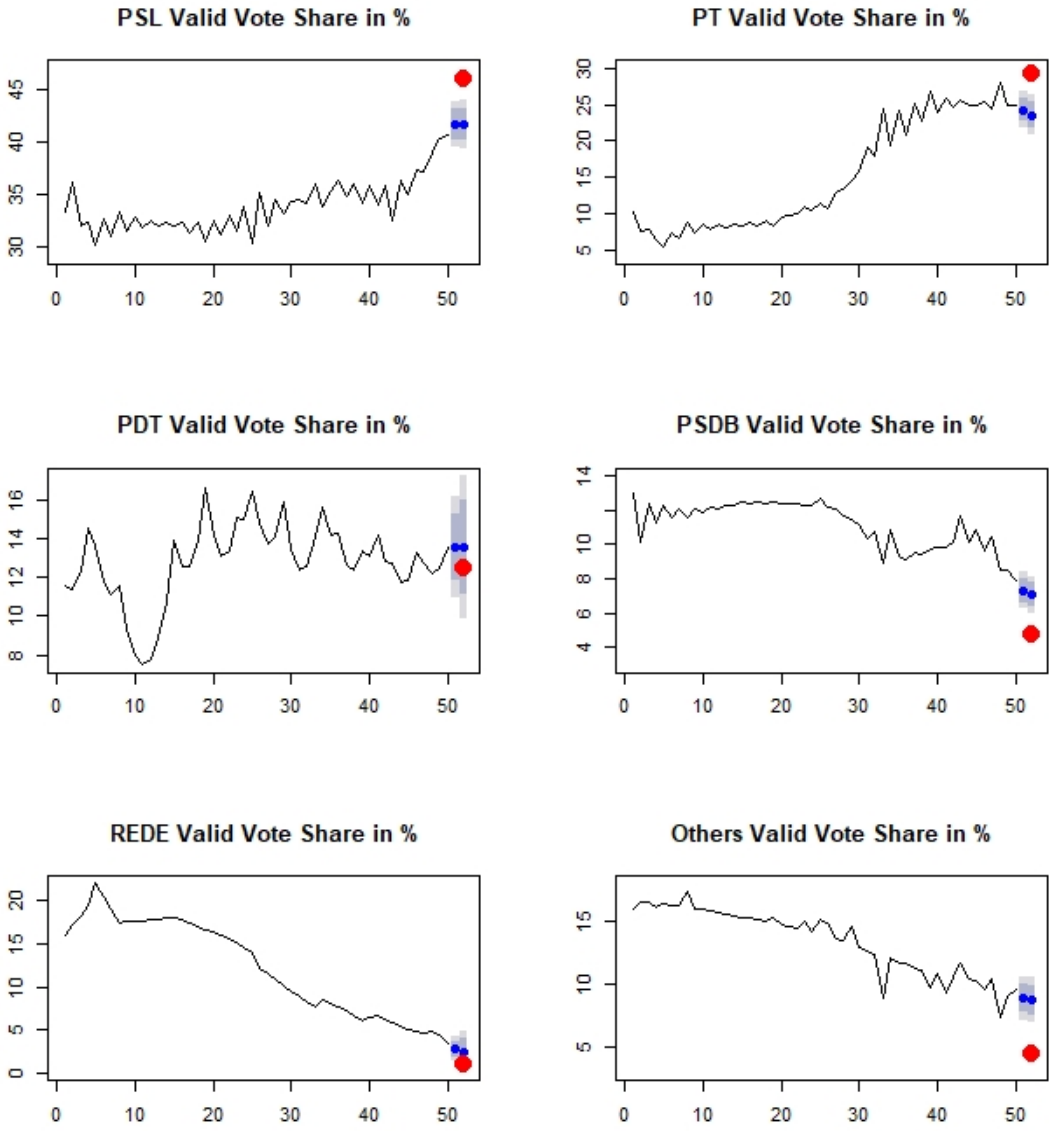

FIGURE 8: ARIMA modeling for the Brazilian elections

Again, the lines represent each time series while the red dot represents the final result.

As a general conclusion for this section, it is possible to see that the Brazilian 2018 elections had more twists and turns than the 2021 Portuguese ones. This is expected, since in Brazil we had a new candidate, a polarization which was unseen and a change in the intentions along the way, while in Portugal it was quieter, with Marcelo Rebelo de Sousa always in the lead by far.

4.3 Structural Model

According to (Reiss & Wolak, 2007), structural econometric models are the ones which combines explicit economic theories with statistical models. That is, how a set of endogenous variables are related to a set of explanatory variables, both if observable. However, the literature in forecasting elections uses this nomenclature when referring to models similar to the regression performed by (Abramowitz, 2008). This nomenclature is followed in the famous (Linzer, 2013) and is called "fundamentals" model in the aready quoted (Stoetzer et al., 2019). Throughout this work, it will be always referred as "Structural Model" even though in the end we will show that this model is not a usefull tool neither for Portuguese nor for Brazilian presidential elections.

(Abramowitz, 2008) refers his model as Time-for-change, which is based on three indicators: the incumbent‘s approval rating at mid-year (POP, from popularity); the GNP growth from te fourth quarter of the previou year to the fourth quarter of the election year (GNP); and a dummy variable which takes the value 0 if the incumbent party has controlled the executive power for four years or 1 if it as controlled for eight years of more (TERM). Then a simple OLS is performed with y_p (a little notation change from the original article) being the percentage of votereceived by the incumbent party.

$$y_p = \beta_0 + \beta_1 POP + \beta_2 GNP + \beta_3 TERM + \epsilon \tag{1}$$

Despite its simplicity, this model exhibits a good forecast performance in some elections, the American one, for example. This model can and is used in systems which are more similar to a bipartisanship, which is not the case of the countries we are trying to model here. Turgeon and Rennó (2012) tried to make this exercise for brazilian elections, although it was good to forecast the results for the incumbent of some elections, the model failed incredibly for 2010 and stated how difficult it is given the high amount of parties and low amount of elections held, which makes the sample size of the regression really small.

Given the structural model‘s good performance on the forecast of the american elections, it is used as prior in a lot of exercises, in the usual form with the three variables

described or enhanced with some more. For example, Linzer (2013) uses the simple form as the mean of the prior in his work. Stoetzer et al. (2019) on other hand uses other features which includes *Institutional features* and *Short-term campaign effects* and applies it to a dirichlet regression. We will show that in our cases these results are not statistically significant and we are going to incorporate another prior on our conjugated model.

Also, in our model we can only use macroeconomic variables, since for many presidents we don't have a popularity research for both countries which makes it impossible to have the POP variable like in Abramovitz' model. The variable TERM is also difficult to include since few elections were held and we have a lot of parties in both countries and the elections have a high variability. For this reason, we will try instead to add two more variables which reflects the macroeconomic situation: inflation and unemployment. The full models will be fully discussed in the next subsection.

4.4 Structural Model Results

As we said in section 4.3, the results here are rather disapointing. The disapointment comes both for brazilian and portuguese cases, which we can se there is no statistical significance or little statistical significance for any variables. We also have the problem of small sample size in both countries, which highly affects the regression's results. Besides, for the brazilian case we have the difficulties of having a new party PSL.

Portuguese Case

For the Portuguese case there are lots of specificities that makes this analysis more difficult. First of all there is no constancy on which parties are running or winning.For example, in the first three elections, PSD did not have an official candidate, nor a supported candidate. Another example is that in 2016 the same happened for PS. And there were even cases when these two parties made agreements to run together. It all makes this analysis much harder to be estimated. Also, usually there is a candidate which wins by far on the first round.

Another difficulty found is regarding the quality of the data. While in (Abramowitz, 2008) we had easy access to data which could measure the popularity of the incumbent, for the Portuguese case (and also for the Brazilian) this is not available. Also, the low sample size and the high variability of the elections make it a little difficult to add the dummy variable for the time term as Abramovitz did. Therefore, we are just going to use macroeconomic variables.

For this analysis, we will perform two regressions, one regression for PS and the other for the most voted party besides PS, regardless if it is the winner or not. The variables we are going to use are Unemployement rate, GDP Growth rate and inflation rate, all of them for the year before the election was held, all of them annually also. All the

macroeconomic data is available at the PorData website [5].

$$y_p = \beta_0 + \beta_1 Inflation + \beta_2 GDP + \beta_3 Unemp + \epsilon \tag{2}$$

The results for both regressions are below at table III.

TABLE III: Structural regression for the portuguese case

	Dependent variable:	
	PS	Other
	(1)	(2)
Inflation_previous_year	1.039	-1.186^{**}
	(0.871)	(0.453)
GDP_Growth	-0.357	1.280
	(1.970)	(1.058)
Unemployment	-6.616	2.993
	(3.585)	(1.517)
Constant	83.543^{**}	20.442
	(29.559)	(14.612)
Observations	8	9
R^2	0.557	0.719
Adjusted R^2	0.225	0.550
Residual Std. Error	18.833 (df = 4)	10.517 (df = 5)
F Statistic	1.677 (df = 3; 4)	4.256^{*} (df = 3; 5)
Note:		*p<0.1; **p<0.05; ***p<0.01

As it is possible to see, only one of the variables are significant in each regression, with the constant one in the PS being the significant one. One more reason not to use these results as our prior. If we were to use the results of the forecast of this model for the prior of our joint model – which was the original plan, the results would be: 41.3% for PS and 30.9% for PSD. Those results of obviously far from the true ones, 60,66% for Marcelo de Sousa from PSD and 12,96% for Ana Gomes, from PSD.

When running also a bayesian regression with defaulted parameters, the results also show a low probability of these predictors of being significative. This was only an exercise runned from a bayesian point of view to compare with the usually employed frequentist one. This was done using the default glm by the R package described in the

[5]The PorData website is availabe at https://www.pordata.pt (n.d.)

article Goodrich, Gabry, Ali, and Brilleman (2020). The results of the regression are at the appendix D.

Brazilian Case

In Brazilian electoral history, as we have shown in section 2, usually there were three major forces running for presidency, PT, PSDB, and another varying candidate. Therefore we will run three regressions, one for each candidate's final vote share with the dependent variables: a dummy if the country had a hyperinflation or not, since Brazil had a large period of hyperinflation (higher than 100%) and this would make the regression too noisy; the GDP growth on the year before the election and the unemployment rate on the month before. We would like to use the incumbent's approval rate, but just like in the Portuguese case, that information is not available for all the periods.

$$y_p = \beta_0 + \beta_1 InflationDummy + \beta_2 GDP + \beta_3 Unemp + \epsilon \tag{3}$$

As it is possible to see on the table IV, none of the variables are significant for PT and PSDB. The significance is that lower that the adjusted R squared is negative, another effect of the really small sample size for the regression. We have indeed some significance for the third forces, but we still are not going to use these results as a forecast.

TABLE IV: Structural regression for the brazilian case

	Dependent variable:		
	PT	PSDB	Third
	(1)	(2)	(3)
Inflation_Dummy	−0.173	0.194	−0.159*
	(0.121)	(0.191)	(0.042)
GDP_Growth	0.012	0.010	−0.001
	(0.027)	(0.043)	(0.010)
Unemployment	0.006	0.021	−0.027*
	(0.022)	(0.035)	(0.008)
Constant	0.351	0.177	0.367**
	(0.201)	(0.317)	(0.070)
Observations	6	6	6
R^2	0.580	0.443	0.916
Adjusted R^2	−0.051	−0.392	0.791
Residual Std. Error (df = 2)	0.092	0.145	0.032
F Statistic (df = 3; 2)	0.920	0.531	7.292
Note:	*p<0.1; **p<0.05; ***p<0.01		

The results for the forecasted values using these regressions and the data for 2018 would be 45% for PT, 45% for PSDB and 4% for the third force, which would be PSL's Bolsonaro.

When performing a simple bayesian regression, the same we did for the portuguese case in this same as it ws done here. For the three cases, we had only one variable - the inflation in the previous year for the Third party – in which zero lies outside the 95% HPDI.

We can see that the R^2 is negative for two of the models, the reason is that we have only a few elections held, which shows as result a very poor result. Since we do not have more elections to test, there is not much to do besides calling the results absolutely inconclusive.

As a general conslusion for both Portuguese and Brazilian cases, it is possible to say that the structural model is not a good tool for forecasting the election results, at least using the traditional variables used by Abramowitz (2008). The effect of political and economical instability, of having a lot of parties, the fact we still had few presidential elections held and the creation and destructions of parties are factors which can explain

the results we had in this chapter. All these factors are in the end translated statistically to a very small sample size and very noisy data. This makes this structural analysis proofless for these two countries.

4.5 Polling company bias Model

There is a common belief, or at least a common discussion that polling companies are - or at least may be - biased towards a party or an ideology. This leads to several big discussion, as one of the best example worldwide being the 2016 American Presidential Elections, in which almost all the polls had Hillary‘s victory as outcome but, in the end, Trump was the winner. This discussion is well described at Cohn (2017).

In Portugal, the polls are more challenged usually in the legislative elections, which may be the most important elections in the country. In 2019, Rui Rio, the leader of the center-right party PSD criticised severely the polls during the campaign and even after elections, as shown in ECO and Lusa (2019) and Azenha (2019)

Also, in 2018 this was a discussion in Brazil, with several enterpreneurs writing public letters agains the polling companies, as stated by one of the most importat newspapers of the countrym Folha de São Paulo, NA (2018b). Even the candidate who was elected, Jair Bolsonaro, said that the Datafolha polls were "shameful" [6] .

In the articles about election forecasting, polling company bias is usually treated as a random variable with mean zero and a low variance. This is the case of, for example, Traugott and Wlezien (2010). Other case is Linzer (2013) which uses what he calls "house effects" to make the fine tunning in the prior sensitivity analysis. Stoetzer et al. (2019) on other hand assumes the bias as a normal process with zero mean and low variance.

Here we will perform Bayeian ANOVAs to confirm whether or not there is a sistematic bias toward specific parties in specific polling houses.

In order to perform this analysis, we will compare the poll results which are closer to the elections with the actual result of the election and see if there is any kind of bias for any party in any polling company. In Portugal‘s case we will use polls from "Intercampos", "Eurosondagem" e "Pitagórica" for every possible election where we can find results and compare, like the *autárquicas* or the regional ones from Ilha da Madeira and Azores.

In the Brazilian case we will use the results from the two most famous and traditional polling companies, "Ibope" and "Datafolha". The elections compared will be for Presidential, State Governor, Mayor and Senator.

Usually the models will be a common bayesian ANOVA where each observation i for the predicted variable y_i will be the observed bias, that is the actual result minus the last forecasted result (Actual value - last polled value).

[6] NA (2018a) discuss the issue, but no author is assigned for the story.

With that for each election we will have several observations - one for each party - where we can identify the bias for every party in every election for every company in which we have polls. The choice of using ANOVA models is that this is the perfect case where we have a numerical depended variable and only nominal predictors.

The rationale behind this is to check if a company has a sistematic bias toward specific parties or if the errors are just errors towards a zero mean. This zero mean will be our prior, that means we are going to assume there are no sistematic errors, just errors which comes from a normal distribution.

Our model is based on the ANOVA from the book (Kruschke, 2014). For each observation i we will estimate y_i based on the estimation of its mean μ_i. The nominal predictors are going to be denoted by the vectors $\vec{x}_1 = (x_{1[1]}, ..., x_{1[P]})$ where P is the number of parties and $\vec{x}_2 = (x_{2[1]}, ..., x_{2[C]})$ where C. When an observation falls in the group p of the first vector, this is represented by $x_{1[p]} = 1$ and $x_{1[i \neq p]} = 0$. The same applies for the other vector. We are also going to have a vector of the combination of both vectors. In the end we will have the bias been modeled like:

$$y_i \sim Normal(\mu_i, \sigma^2)$$

with its mean being:

$$\begin{aligned} \mu_i &= \beta_0 + \vec{\beta}_1 \vec{x}_1 + \vec{beta}_2 \vec{x}_2 + \vec{\beta}_{1x2} \vec{x}_{1x2} \\ &= \beta_0 + \sum_p \beta_{1[p]} x_{1[p]} + \sum_c \beta_{2[c]} x_{2[c]} + \sum_{[p,c} \beta_{1x2[p,c]} x_{1x2[p,c]} \end{aligned}$$

And the deflecations within each factor and within each interaction are constrained to sum to zero, which means all the following conditions are applied at the same time:

$$\begin{aligned} \sum_p \beta_{1[p]} &= 0 \\ \sum_c \beta_{2[c]} &= 0 \\ \sum_p \beta_{1x2[c,p]} &= 0 \text{ for all } c \\ \sum_c \beta_{1x2[c,p]} &= 0 \text{ for all } p \end{aligned}$$

The necessity and the proof of these constraints are well explained in the chapter 15 of

the book (Kruschke, 2014). We also have the following parameter as input to the models:

$$\beta_0 \sim Normal(\mu_0, \sigma^2)$$
$$\beta_{1[p]} \sim Normal(0, \sigma_1^2)$$
$$\beta_{2[c]} \sim Normal(0, \sigma_2^2)$$
$$\beta_{1x2[p,c]} \sim Normal(0, \sigma_{1x2}^2)$$
$$1/\sigma \sim Gamma(1.5, 6)$$
$$1/\sigma_1 \sim Gamma(1.5, 6)$$
$$1/\sigma_2 \sim Gamma(1.5, 6)$$
$$1/\sigma_{1x2} \sim Gamma(1.5, 6)$$

The justification for the choice of priors is that we will first assume the bias is toward 0 with a high variance which is supported mainly by the fact some of the results on other elections have some pretty high outliers. Even if this is not the case for the presidential elections. The hyperprior for the standard deviation comes from a gamma distribution, since the values must be positive. More information can be found in the appendix A.

We will run 3 MCMC chains with 5000 iterations and burn the first 1000 ones.

4.6 Polling company bias Results

Here we present the results for the Polling company bias model. We will see that the results were quite good in the sense we can see that the bias is not statistically relevant for any party in any of the companies in our analysis.

Portugal

For the portuguese case of polling company bias we are using the five most important polling companies - Aximage (AXI), Eurosondagem (EURO), Intercampus (Inter), Pitagórica (Pita) and Universidade Católica de Portugal (UCP) - for all the parties with a candidate in 2021's election. As it is shown in table V, the bias is very low in all the cases and zero is always in the HDI, therefore we can say that none of the coefficients are significant.

Brazil

For Brazil we are going to model for two polling companies, Datafolha and Ibope. Another company, Ipespe, will be used on the conjugated model, but since there are few observations of the company before the 2018 election, we are going to assume mean zero for it. We will see that zero is a good prior for every company. Also we are only using the parties which are running for the 2018 presidential election.

In the table VI, we can see that the the bias is also not statistically significant for any of the coefficients calculated.

TABLE V: bayesian anova for portuguese bias model

	Mean	SD	Lower 95	Upper 95	Lower 90	Upper 90
b0	-0.86	0.32	-1.49	-0.23	-1.38	-0.33
b1[BE]	0.34	0.36	-0.28	1.09	-0.20	0.93
b1[CHE]	0.09	0.40	-0.72	0.95	-0.54	0.79
b1[IL]	0.12	0.42	-0.68	1.04	-0.57	0.78
b1[PCP]	0.24	0.43	-0.56	1.14	-0.41	0.94
b1[PS]	-0.52	0.44	-1.40	0.20	-1.20	0.14
b1[PSD]	-0.28	0.36	-1.06	0.35	-0.88	0.27
b2[Axi]	0.06	0.26	-0.45	0.64	-0.35	0.50
b2[Euro]	-0.00	0.25	-0.52	0.54	-0.39	0.44
b2[Inter]	0.02	0.26	-0.53	0.57	-0.42	0.43
b2[Pita]	0.18	0.30	-0.35	0.84	-0.28	0.67
b2[UCP]	-0.24	0.29	-0.86	0.24	-0.71	0.17
b1b2[BE,Axi]	-0.06	0.32	-0.82	0.56	-0.61	0.42
b1b2[CHE,Axi]	-0.01	0.33	-0.74	0.71	-0.55	0.52
b1b2[IL,Axi]	-0.01	0.33	-0.81	0.65	-0.55	0.51
b1b2[PCP,Axi]	-0.00	0.33	-0.70	0.75	-0.54	0.53
b1b2[PS,Axi]	-0.02	0.31	-0.70	0.67	-0.56	0.46
b1b2[PSD,Axi]	0.10	0.32	-0.52	0.84	-0.36	0.67
b1b2[BE,Euro]	-0.00	0.31	-0.67	0.68	-0.53	0.48
b1b2[CHE,Euro]	-0.01	0.34	-0.76	0.74	-0.56	0.52
b1b2[IL,Euro]	-0.02	0.34	-0.76	0.71	-0.54	0.55
b1b2[PCP,Euro]	0.05	0.33	-0.65	0.81	-0.48	0.59
b1b2[PS,Euro]	-0.03	0.30	-0.71	0.60	-0.53	0.46
b1b2[PSD,Euro]	0.02	0.30	-0.65	0.67	-0.47	0.52
b1b2[BE,Inter]	0.11	0.35	-0.55	0.89	-0.40	0.68
b1b2[CHE,Inter]	-0.02	0.35	-0.77	0.74	-0.60	0.51
b1b2[IL,Inter]	-0.03	0.34	-0.79	0.67	-0.59	0.50
b1b2[PCP,Inter]	-0.05	0.34	-0.77	0.70	-0.65	0.45
b1b2[PS,Inter]	-0.09	0.33	-0.88	0.51	-0.66	0.39
b1b2[PSD,Inter]	0.08	0.32	-0.57	0.76	-0.38	0.64
b1b2[BE,Pita]	-0.13	0.35	-1.00	0.47	-0.69	0.40
b1b2[CHE,Pita]	-0.06	0.34	-0.81	0.66	-0.64	0.46
b1b2[IL,Pita]	-0.03	0.33	-0.77	0.69	-0.59	0.48
b1b2[PCP,Pita]	-0.09	0.36	-0.89	0.64	-0.68	0.46
b1b2[PS,Pita]	0.18	0.38	-0.42	1.09	-0.33	0.80
b1b2[PSD,Pita]	0.13	0.35	-0.50	0.97	-0.38	0.73
b1b2[BE,UCP]	0.08	0.30	-0.51	0.77	-0.40	0.57
b1b2[CHE,UCP]	0.09	0.34	-0.55	0.91	-0.41	0.67
b1b2[IL,UCP]	0.08	0.34	-0.52	0.93	-0.44	0.64
b1b2[PCP,UCP]	0.09	0.34	-0.55	0.90	-0.45	0.63
b1b2[PS,UCP]	-0.03	0.28	-0.66	0.56	-0.51	0.41
b1b2[PSD,UCP]	-0.32	0.43	-1.27	0.27	-1.01	0.20

TABLE VI: bayesian anova for brazilian bias model

	Mean	SD	Lower 95	Upper 95	Lower 90	Upper 90
b0	-0.02	0.20	-0.40	0.37	-0.33	0.31
b1[NOVO]	0.00	0.39	-0.82	0.81	-0.65	0.63
b1[PDT]	-0.03	0.33	-0.74	0.64	-0.59	0.52
b1[PMDB]	0.22	0.31	-0.36	0.86	-0.28	0.73
b1[PPL]	-0.01	0.35	-0.75	0.71	-0.59	0.56
b1[PSC]	-0.08	0.35	-0.82	0.65	-0.68	0.48
b1[PSDB]	-0.26	0.30	-0.91	0.26	-0.78	0.19
b1[PSL]	-0.04	0.38	-0.84	0.73	-0.66	0.58
b1[PSOL]	0.05	0.26	-0.47	0.58	-0.36	0.50
b1[PSTU]	0.01	0.27	-0.53	0.54	-0.42	0.46
b1[PT]	0.15	0.28	-0.37	0.74	-0.29	0.61
b1[REDE]	-0.01	0.38	-0.78	0.81	-0.64	0.61
b2[DF]	0.14	0.15	-0.12	0.44	-0.08	0.38
b2[IB]	-0.14	0.15	-0.44	0.12	-0.38	0.08
b1b2[NOVO,DF]	-0.04	0.33	-0.76	0.64	-0.59	0.50
b1b2[PDT,DF]	0.25	0.35	-0.33	1.03	-0.26	0.83
b1b2[PMDB,DF]	-0.05	0.26	-0.59	0.47	-0.49	0.37
b1b2[PPL,DF]	-0.05	0.30	-0.70	0.56	-0.57	0.43
b1b2[PSC,DF]	-0.10	0.32	-0.81	0.50	-0.65	0.38
b1b2[PSDB,DF]	0.58	0.43	-0.07	1.42	-0.07	1.22
b1b2[PSL,DF]	-0.06	0.32	-0.77	0.58	-0.60	0.46
b1b2[PSOL,DF]	-0.08	0.23	-0.58	0.37	-0.46	0.31
b1b2[PSTU,DF]	-0.18	0.25	-0.72	0.26	-0.60	0.21
b1b2[PT,DF]	-0.21	0.27	-0.80	0.26	-0.67	0.19
b1b2[REDE,DF]	-0.06	0.33	-0.81	0.59	-0.61	0.48
b1b2[NOVO,IB]	0.04	0.33	-0.64	0.76	-0.50	0.59
b1b2[PDT,IB]	-0.25	0.35	-1.03	0.33	-0.83	0.26
b1b2[PMDB,IB2]	0.05	0.26	-0.47	0.59	-0.37	0.49
b1b2[PPL,IB]	0.05	0.30	-0.56	0.70	-0.43	0.57
b1b2[PSC,IB]	0.10	0.32	-0.50	0.81	-0.38	0.65
b1b2[PSDB,IB]	-0.58	0.43	-1.42	0.07	-1.22	0.07
b1b2[PSL,IB]	0.06	0.32	-0.58	0.77	-0.46	0.60
b1b2[PSOL,IB]	0.08	0.23	-0.37	0.58	-0.31	0.46
b1b2[PSTU,IB]	0.18	0.25	-0.26	0.72	-0.21	0.60
b1b2[PT,IB]	0.21	0.27	-0.26	0.80	-0.19	0.67
b1b2[REDE,IB]	0.06	0.33	-0.59	0.81	-0.48	0.61

More information about the sensitivity analysis of the ANOVA model for polling company bias can be found at appendix B. There it is possible to see that the model is not sensible to changes in the prior variance.

4.7 Marginal Dynamic Linear Model

The marginal models in the bayesian perspective will be just simple dynamic linear models, which we will make them the usual univariate space state model.

The Dynamic Linear Models are usually well described in a lot of bayesian books. Here we will use West and Harrison (1997) for the basic theory and Holmes, Scheuerell, and Ward (2021) as some good examples for the coding in JAGS for R.

Both of the brazilian and the portuguese models will follow this model for each of the series for each party. For Brazil we will set this model six times, five for the following parties: PSL, PT, PDT, PSDB, REDE and one more for the sum of the other small candidates. For Portugal we will have PSD, PS, CHEGA and the sum of the others. It will be, as always, for valid votes, which means we are not counting undecided votes or answers not given.

In this model we assume that y_t is the poll preference for each candidate at the time t. The expected value for y_t is:

$$E[y_t] = x_t \tag{4}$$

We assume that the candidate‘s percentage preference follows a univariate space state model, which is the usual stochastic level model. The model is described as below:

$$y_t = x_t + v_t, v_t \sim N(0, \sigma^2)$$
$$x_t = x_{t-1} + w_t, w_t \sim N(0, \sigma_1^2)$$

With the hyperparameters:

$$1/\sigma^2 \sim Gamma(1.5, 6)$$
$$1/\sigma_1^2 \sim Gamma(1, 3)$$

With this modelling, we have the true level of the series being the results of each available poll y_t and the obeserved level of vote share being x_t. As we will see in section 4.6, the polls have little sistematic bias, therefore equation 4 shows a fair assumption.

The justification of the choice of the priors in better explained in the appendix A.

4.8 Marginal Dynamic Linear Model Results

In the marginal Dynamic Linear model, every vote share is considered separatedly and modelled as a Single DLM where there is an observed level of intended vote share.

Portuguese Case

In the portuguese case we are going to analyze four different parties: PSD, PS, CHEGA and the sum of all others as one. We can see first that the model estimated for the Portuguese case seem very simple, in the figure 9, it is possible to see that all the four models do not have abrupt changes. The darker area is the 95% HDI. The red dot represents the true election result.

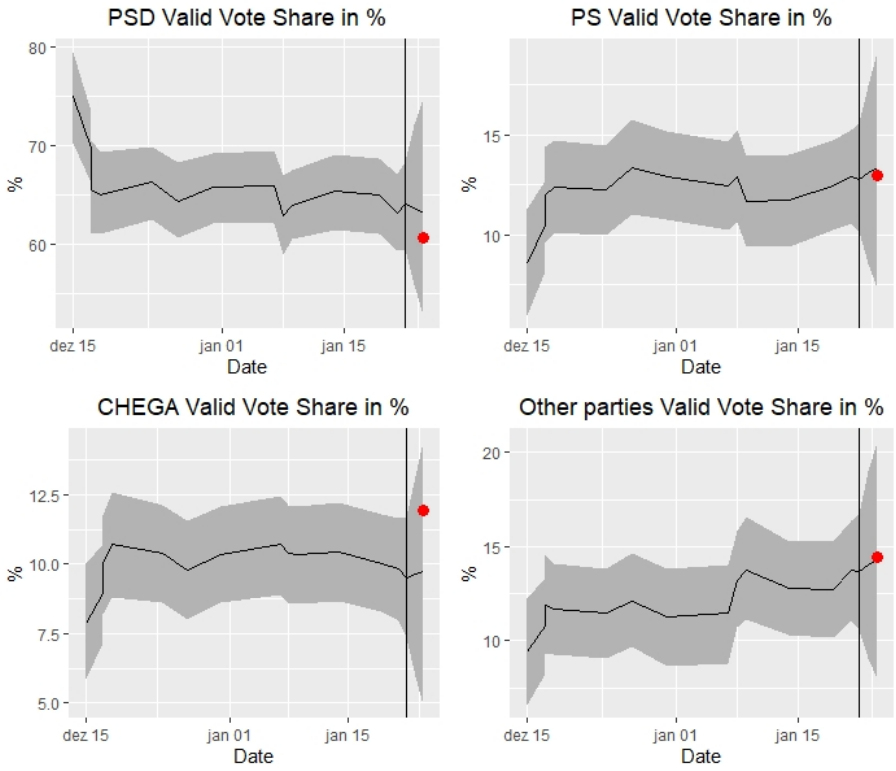

FIGURE 9: Marginal Dynamic Linear Model results for Portuguese elections

Brazilian Case

In the brazilian case the analysis uses four parties: PSL, PT, PDT, PSDB, REDE and the sum of all others as one. It is different from the Portuguese case in the sense that the movements are much more abrupt here, since you can see a big rise in the vote share intention for PSL's Bolsonaro, PT's Haddad and a fair rise for PDT's Gomes. Meanwhile, all the other parties faced a big decline during the campaign. The darker area is the 95% HDI. The red dot is again the true election result.

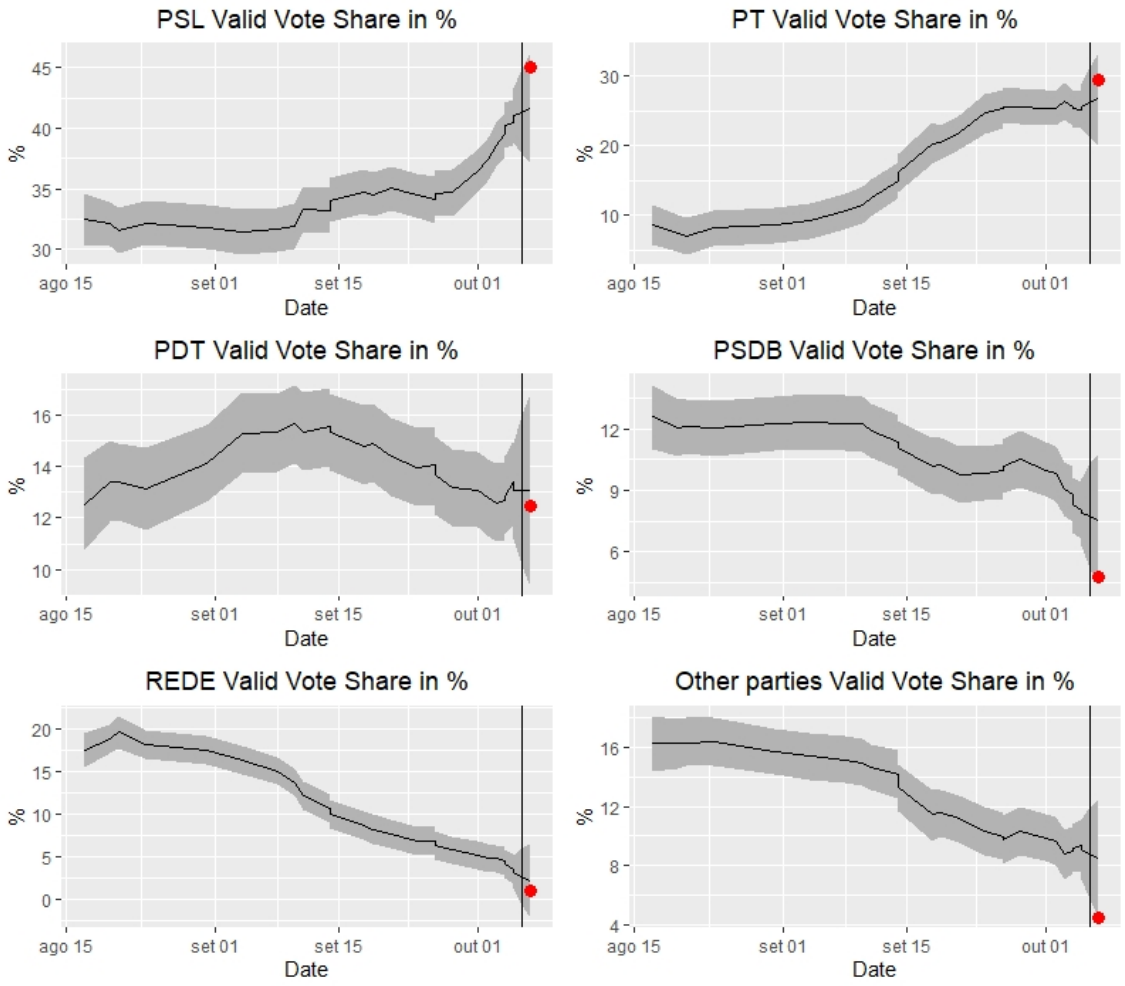

FIGURE 10: Marginal Dynamic Linear Model results for Brazilian elections

The small conclusion for this section is similar to the one from the section 4.2, the Brazilian 2018 elections were far more complicated to analyze than the Portuguese 2021 ones, with an easy measure of who was going to win.

4.9 Joint Dynamic Linear Model

The Joint model is the one which takes the most information amongst all the models discussed here. It is based on the Stoetzer et al. (2019). This article combines the prior result given by the stuctural model with an inverse dynamic linear model from a Multinomial Distribution.

Our model will be similar, with the addition of our polling company bias results and with a difference in our prior results, which in here will be the first poll result. The justification for that is, as we are going to further discuss in section 4.6, the bias for the companies is usually very small and the results can be used as a first source of information. Another reason is the results we are going to see at section 4.4, the results from the

structural model are not usefull neither from portuguese, nor from brazilian elections.

The model is as follows, y_{ctp} is the reported vote share of the candidate from party $p = 1, ..., P$, at time $t = 1, ..., T$ from the results of company $c = 1, ..., C$. Where P is the total number of parties, C is the total number of polling companies. The time t is expressed in terms of the first poll, where the first poll is at time 1 and the last at time T. Each poll has a sample size N_{ct}. The published results are P-dimensional with $\vec{y}_{ct} = (y_{ct1}, ..., y_{ctP})$ which is assumed to come from a multinomial process from the vector of expected daily support of $\vec{\pi^*_{ct}} = (\pi_{ct1}, ..., \pi_{ctP})$.

$$\vec{y_{ct}} \sim Multinomial(\vec{\pi^*_{ct}}, N_{ct}) \tag{5}$$

After that we employ a log ratio transformation in the expected support share vector $\vec{\pi^*_{ct}}$:

$$\boldsymbol{\pi_{ct}} = alr(\vec{\pi^*_{ct}}) \tag{6}$$

$$alr(\vec{\pi^*_{ct}}) = (log(\frac{\pi^*_{ct1}}{\pi^*_{cP}}), ..., log(\frac{\pi^*_{ct(P-1)}}{\pi^*_{cP}})) = (\pi_{ct1}, ..., \pi_{ct(P-1)}) \tag{7}$$

Where alr is the log ratio transformation. This is necessary to map the proportions into o a vector of unbounded, real-valued quantities.

Now we decompose the transformed vector $\vec{\pi_{ct}}$ into the latent support vector, $\vec{\alpha_t} = (\alpha_1, ..., \alpha_t)$ and a vector of the polling company bias, $\vec{\delta_p} = (\delta_1, ..., \delta_P)$. Since the is no sistematic bias, there will be only one vector, if there was a sistematic bias, there would be C vectors of size P. In the end, we have:

$$\vec{\pi_{ct}} = \vec{\alpha_t} + \vec{\delta_p} \tag{8}$$

With that, we model $\vec{\alpha_t}$ as a random walk, which starts at the first day of campaign and moves forward until the Election Day. In the Stoetzer et al. (2019), this modelling is performed through a backward random walk to incorporate the prior dirichlet regression. Since in our case we are not going to use it, the normal forward random walk was chosen.

$$\vec{\alpha_t} = \vec{\alpha_{t-1}} + \omega_t, \omega_t \sim Normal(0, \boldsymbol{W}) \tag{9}$$

With this specification, we can estimate the vote shares level even if there is no poll at a specific day or if there is more than one poll for the same day. The variance is based on the evolution variance $\boldsymbol{W}$, which is best described in (West & Harrison, 1997). This allows $\boldsymbol{W}$ to be constant over time, but allows the latent states to covary. The full matrix is specified in the Appendix E

$$\delta_c \sim Normal(\mu_c, \sigma_c^2) \tag{10}$$

Where $\mu_{cp}$and the σ_{cp^2} are 0 and 0.02, which is the same as the one specified in the section 4.6.

4.10 Joint Dynamic Linear Model Results

The joint linear model, as explained in the section 4.9, is built to incorporate priors for both the historical data based on a linear regression using socioeconomic factors and also for systematic polling company bias in which we calculated using a two-way bayesian anova. The linear regression showed an absence of significance for every variable for both countries. The two-way bayesian anova also showed there is no systematic bias for any party in any polling company. With those results, we are going to use the results for the first poll as the prior which would arise from the historical data and zero as the mean prior for bias.

As it was said in the model description, this model is heavily used nowadays, specially for the American case, but also for other elections. But we will see that it captures the movements for the elections differently for the brazilian case.

Portuguese Case

For the portuguese case it is possible to see that the model fitted in the end is very similar to the other ones, with no hard movements whatsoever. The scale in the figure 11 is much smaller than in the figure 12 because of the variance of the forecasted coefficients. The darker area is the 95% HDI.

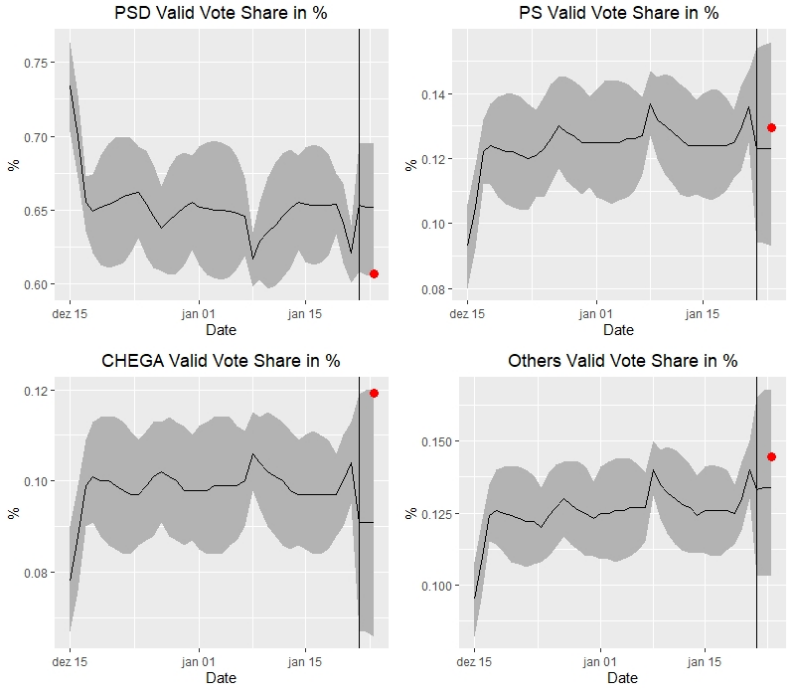

FIGURE 11: Joint Dynamic Linear Model results for Portuguese elections

Brazilian Case

In the Brazilian case we have in this model some differences when comparing to the Frequentist and the Marginal Dynamic. First of all it fits a much smaller growth for PSL than the other models - even though the growth fro PT is also important here - it also fits a downward curve for PDT and a somewhat stable curve for PSDB. In the end we can see that the uncertainty of this model for the greater variance. The darker area is the 95% HDI.

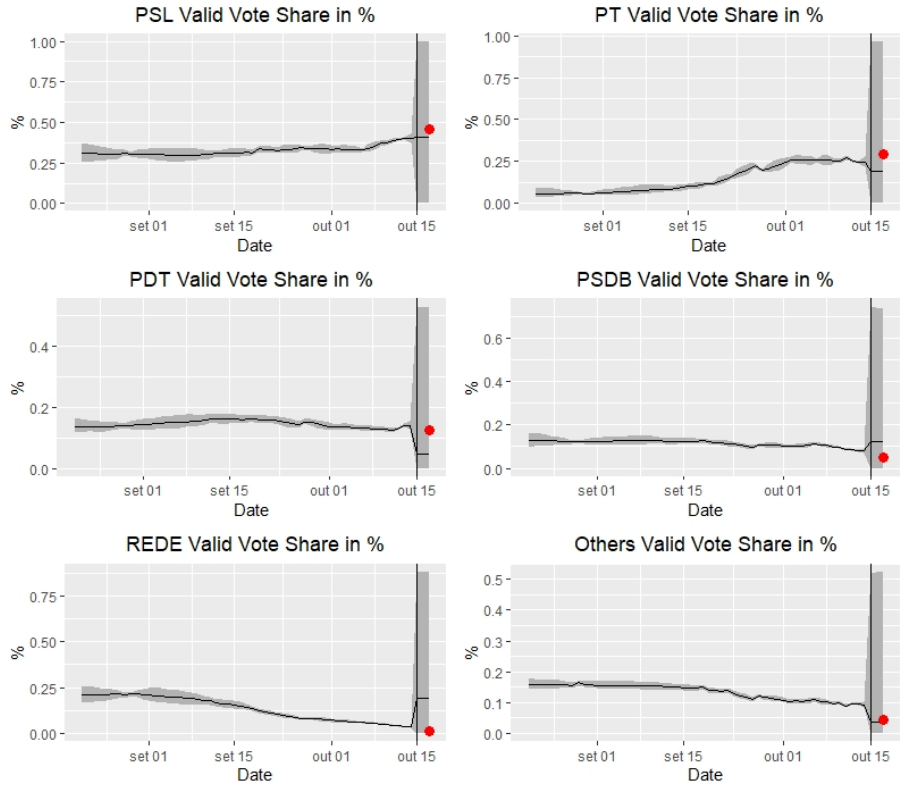

FIGURE 12: Joint Dynamic Linear Model results for Brazilian elections

As the conclusion for this section we have that the results for the portuguese election are not that different from the other models, while we have different results for the nrazilian example. We are going to see how it affects the final results in the section 5.

5 COMPARISON

In this section we will compare the results of the three main models we have: the frequentist model based on the arima modeling, the Marginal Dynamic model in which each candidate is modeled simultaneously. We will compare only the mean of the estimate, since we care about the final result. We will compare the estimations at two points in time, two days before the elections, that means the last poll in both Portuguese and Brazilian cases, in which we have the polls on Friday and the elections on Sunday. The other point in time is 10 days before the elections. A little back in time, in order to see the effect of the uncertainty, but not that much in the way we don't have enough data to estimate, since even a 15 day before forecast could be too biased due to low amount of data, specially in the portuguese case where fewer polls are made.

Starting with the analysis of the forecast two days before the elections - which are presented in the tables VII and VIII, we can see that the model which gets results closer to the actual one for the Portuguese 2021 election is the Marginal one, followed by the

Joint model and the Frequentist ARIMA has the worse results. It is important to remember that the joint model must sum to 100, since all the series are estimated as a set. It will not sum only if there is a rounding problem. The other two may not sum to 100.

For the Brazilian 2018 elections we have mixed results. While the Marginal model predicted better the results for the winner and runner's up, the joint model had closer results for PDT and PSDB, but terrible results for the others. The Frequentist model kept being closer to the Marginal, but always a little further from the actual result.

TABLE VII: Model comparison for the Portuguese case 2 days before election in %

Party	True Result	Joint	Marginal	Frequentist
PSD	60.66	65.20	63.12	65.89
PS	12.96	12.30	13.31	12.17
CHEGA	11.93	9.10	9.70	8.83
Others	14.45	13.40	14.25	13.26

TABLE VIII: Model comparison for the Brazilian case 2 days before election in %

Party	True Result	Joint	Marginal	Frequentist
PSL	46.03	40.90	41.69	41.60
PT	29.28	19.40	26.74	23.46
PDT	12.47	12.30	13.10	13.54
PSDB	4.76	4.60	7.57	7.06
REDE	1.00	3.80	2.17	2.57
Others	4.46	19.00	8.58	8.74

Now for the analysis 10 days before the election - available in the Tables IX and X - we can see that for the portuguese case, the forecasted results are close to the actual results. This is true specially for the marginal model, which is the bet performing model for the Portuguese case 10 days before the election, with the other two performing more or less similarly depending on the party.

For the brazilian case, on other hand, the results, specially for PSL, are much further than the actual ones. This shows how great the growth on Bolsonaro's voting intention was in the last days of campaign. It is even difficult to see which model was better for this case, with a lot of mixed results.

TABLE IX: Model comparison for the Portuguese case 10 days before election in %

Party	True Result	Joint	Marginal	Frequentist
PSD	60.66	63.50	63.10	65.89
PS	12.96	10.10	13.35	12.00
CHEGA	11.93	10.20	9.68	10.51
Others	14.45	16.20	14.21	12.97

TABLE X: Model comparison for the Brazilian case 10 days before election in %

Party	True Result	Joint	Marginal	Frequentist
PSL	46.03	34.20	36.40	35.53
PT	29.28	26.80	34.44	25.46
PDT	12.47	13.40	15.34	13.12
PSDB	4.76	9.70	7.24	7.06
REDE	1.00	6.10	0.00	2.57
Others	4.46	9.80	5.74	8.74

6 CONCLUSION

There are some nice conclusions which can be presented here. The first one is that the usual source of prior information, the model used by Abramowitz (2008) to forecast elections using historical data, is not usefull for the cases analysed here. The reason is that in opposition to the american elections, there are many parties, fewer elections held since the two countries are recent democracies and also because the economy and political background is more unstable, with parties being created or dismissed with a high frequency. This is better analysed in the chapter 4.4.

The second conclusion is about the second source of historical prior proposed in this work is about the systematical in house bias. As we could see in the chapter 4.6 which uses bayesian ANOVA for the bias analysis, there are no systematic parties for any parties in any of the polling houses analysed for both Brazil and Portugal. This also means that since the polls are systematically unbiased, they can be used as the best source of information for forecasting elections.

The third conclusion is that the portuguese 2021 elections were far more quieter and predictable than the brazilian 2018. This is explained by a simple analysis, Portugal had a favorite from the start to reelection and the campaigns were smooth without great turnarounds. Meanwhile, for the brazilian case we had a newcomer rising, a hugh polarization, an atempted murdering in Bolsonaro and a lot of ups and downs and a huge rise of PSL's candidate specially in the last days of campaing.

Finally, it is difficult to point which model is better, which means here that the complexity applied in the Joint model may be too much to forecast these specific elections. Also we can see that there is not a single good model to solve the issues pointed here. Elections in recent democracies are too noisy and it is still difficult to forecast future movements in public's opinion.

With all that being said, there are some further advancements that can be made for the future. In the bayesian framework, there is the possibility of incorporating priors for each candidate's growth. That can even be done, for example, taking specialists opinions or

something like that. Also, other methods of explaining the opinion before the campaign could be studied, since the usual structural model is not significant for both elections.

A further disclaimer could also be done when forecasting semi-presidentialist systems like Portugal and Presidentialist ones like Brazil.

In the Frequentist framework, a further advancement could be a way to incorporate the the information of the evolution of the other candidates into one's evolution.

REFERENCES

Abramowitz, A. I. (1988). An improved model for predicting presidential election outcomes. *Political Science and Politics*, *21*(4), 843-847.

Abramowitz, A. I. (1996, October). Bill and al's excellent adventure: Forecasting the 1996 presidential election. *American Politics Quarterly*, *24*(4), 4334-442.

Abramowitz, A. I. (2008, October). Forecasting the 2008 presidential election with the time-for-change model. *Political Science and Politics*, *41*(4), 691-695.

Azenha, A. S. (2019). Rui rio atribui derrota nas eleições a críticos e sondagens. *Correio da Manhã*. Retrieved from https://www.cmjornal.pt/politica/detalhe/rui-rio-atribui-derrota-nas-eleicoes-a-criticos-e-sondagens

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1994). *Time series analysis, forecasting and control* (P. Hall, Ed.). Prentice Hall.

Cohn, N. (2017). A 2016 review: Why key state polls were wrong about trump. *The New York Time*. Retrieved from https://www.nytimes.com/2017/05/31/upshot/a-2016-review-why-key-state-polls-were-wrong-about-trump.html

ECO, & Lusa. (2019). Rui rio critica sondagens. pergunta se "ainda vale a pena fazer eleições?". *Eco*. Retrieved from https://eco.sapo.pt/2019/07/22/rui-rio-critica-sondagens-e-pergunta-ainda-vale-a-pena-fazer-eleicoes/

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). *rstanarm: Bayesian applied regression modeling via Stan.* Retrieved from https://mc-stan.org/rstanarm (R package version 2.21.1)

Holbrook, T. M., & DeSart, J. A. (1999, April). Using state polls to forecast presidential election outcomes in the american states. *International Journal of Forecasting*, *127*(1/2), 207-223.

Holmes, E. E., Scheuerell, M. D., & Ward, E. J. (2021). *Applied time series analysis for fisheries and environmental sciences*. Northwest Fisheries Science Center.

https://www.pordata.pt. (n.d.). *Pordata.* Retrieved from https://www.pordata.pt (Accessed: 2021-04-25)

https://www.whitehouse.gov/. (n.d.). *White house.* Retrieved from https://www.whitehouse.gov/ (Accessed: 2021-03-05)

Kruschke, J. K. (2014). *Doing bayesian data analysis: A tutorial with r, jags and stan*. Elsevier.

Lewis-Beck, M. S., & Rice, T. W. (1993). Forecasting elections. *International Journal of Forecasting*, *9*(2), 274-275.

Linzer, D. A. (2013). Dynamic bayesian forecasting of presidentialelections in the states. *Journal of the American Statistical Association*, *108*(501), 124-134.

Mata, M. E., & Valério, N. (2011). *The concise economic history of portugal: a comprehensive guide* (E. A. SA, Ed.). Almedina.

Mellow, N., & Trubowitz, P. (2005). Red versus blue: American electoral geography and congressional bipartisanship, 1898-2002. *Political Geography*, *24*, 659-677.

Moritz, S., & Bartz-Beielstein, T. (2017). imputets: Time series missing value imputation in r. *The R Journal*, *9*(1), 207–218.

NA. (2018a). Bolsonaro chama pesquisa datafolha de 'vexame'. *PODER360*. Retrieved from https://www.poder360.com.br/eleicoes/bolsonaro-chama-pesquisa-datafolha-de-vexame/

NA. (2018b). Corrente que circula entre investidores questiona metodologia adotada por pesquisas eleitorais. *Folha De São Paulo*. Retrieved from https://www1.folha.uol.com.br/poder/2018/09/corrente-que-circula-entre-investidores-questiona-metodologia-adotada-por-pesquisas-eleitorais.shtml

Norpoth, H. (1996). Of time and candidates: A forecast for 1996. *American Politics Quarterly*, *24*(4), 443-467.

Plummer, M. (2019). rjags: Bayesian graphical models using mcmc [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=rjags (R package version 4-10)

Rawicz, F. (2021). *Frawicz/election.* GitHub. Retrieved from https://github.com/frawicz/Elections

Reiss, P. C., & Wolak, F. A. (2007). Handbook of econometrics, volume 6a. In Elsevier (Ed.), (p. 4281-4317). Elsevier.

Stan Development Team. (2020). RStan: the R interface to Stan [Computer software manual]. Retrieved from http://mc-stan.org/ (R package version 2.21.2)

Stoetzer, L. F., Neunhoeffer, M., Gschwend, T., Munzert, S., & Sternberg, S. (2019). Forecasting elections in multiparty systems:a bayesian approach combining polls and fundamentals. *Political Analysis*, *27*(2), 255-262.

Team, R. (2020). Rstudio: Integrated development environment for r [Computer software manual]. Boston, MA. Retrieved from http://www.rstudio.com/

Team, R. C. (2017). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Traugott, M., & Wlezien, C. (2010). The dynamics of poll performance during the 2008 presidential nomination contest. *Public Opinion quaterly*, *73*(5), 866-894.

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

West, M., & Harrison, J. (1997). *Bayesian forecastingand dynamic models*. Springer.

GLOSSARY

FHC Fernando Henrique Cardoso. 48

PDT Partido Democrático Trabalhista. 12, 48

PRN Partido da Reconstrução Nacional. 12, 48

PS Partido Socialista. 48

PSD Partido Social Democrata. 48

PSDB Partido da Social Democracia Brasileira. 12, 48

PT Partido dos Trabalhadores. 12, 48

A PRIOR CHOICES

Formulas for finding the prior for the variance in the hierarchical models using a inverse gamma distribution:

$$1/\sigma^2 \sim Ga(\frac{\nu_0}{2}, \frac{\nu_0}{2}\sigma_0^2) \tag{11}$$

For both hierarchical models for polling company bias we will assume variance goes around 3 p.p., that means $\sigma_0^2 = 4$, and to quantify our uncertainty, ν_0 will be 3. That leaves us with:

$$1/\sigma^2 \sim Ga(1.5, 6) \tag{12}$$

B SENSITIVITY ANALYSIS

For the Sensitivity analysis we will test the models with a much higher variance and uncertaity for variance to test if the results change a lot.

In the polling company bias, for example, we will assume that $\sigma_0^2 = 5$ and $\nu_0 = 5$ and after even higher values of $\sigma_0^2 = 10$ and $\nu_0 = 10$. These values will be the same for both Portugal and Brazilian bias.

B.1 Polling Company Bias

For the both tables exposed here, the parties and polling houses order is the same as in the section 4.6.

TABLE XI: Sensitivity Aalysis ANOVA bias PT

	Mean model 1	SD model 1	Mean model 2	SD model 2
b0	-0.91	0.29	-1.02	0.21
b1[1]	0.28	0.30	0.12	0.15
b1[2]	0.07	0.32	0.02	0.14
b1[3]	0.09	0.32	0.03	0.14
b1[4]	0.18	0.33	0.06	0.14
b1[5]	-0.40	0.35	-0.15	0.17
b1[6]	-0.22	0.29	-0.09	0.14
b2[1]	0.05	0.22	0.02	0.11
b2[2]	0.00	0.22	0.01	0.11
b2[3]	0.02	0.22	0.01	0.11
b2[4]	0.14	0.25	0.06	0.12
b2[5]	-0.21	0.24	-0.10	0.13
b1b2[1,1]	-0.05	0.26	-0.02	0.11
b1b2[2,1]	-0.00	0.26	-0.00	0.11
b1b2[3,1]	-0.01	0.26	-0.00	0.11
b1b2[4,1]	-0.00	0.26	-0.00	0.11
b1b2[5,1]	-0.01	0.25	-0.00	0.11
b1b2[6,1]	0.08	0.26	0.02	0.11
b1b2[1,2]	0.00	0.25	-0.00	0.11
b1b2[2,2]	-0.01	0.26	-0.01	0.11
b1b2[3,2]	-0.02	0.27	-0.01	0.11
b1b2[4,2]	0.03	0.26	0.01	0.11
b1b2[5,2]	-0.02	0.25	-0.01	0.10
b1b2[6,2]	0.02	0.25	0.01	0.11
b1b2[1,3]	0.09	0.28	0.02	0.11
b1b2[2,3]	-0.01	0.27	-0.00	0.11
b1b2[3,3]	-0.02	0.27	-0.01	0.11
b1b2[4,3]	-0.04	0.27	-0.01	0.11
b1b2[5,3]	-0.07	0.26	-0.02	0.11
b1b2[6,3]	0.06	0.26	0.02	0.11
b1b2[1,4]	-0.11	0.28	-0.03	0.11
b1b2[2,4]	-0.04	0.27	-0.01	0.11
b1b2[3,4]	-0.02	0.26	-0.01	0.11
b1b2[4,4]	-0.06	0.28	-0.02	0.11
b1b2[5,4]	0.13	0.30	0.04	0.11
b1b2[6,4]	0.10	0.27	0.03	0.11
b1b2[1,5]	0.08	0.25	0.03	0.11
b1b2[2,5]	0.07	0.27	0.02	0.11
b1b2[3,5]	0.07	0.27	0.02	0.11
b1b2[4,5]	0.07	0.27	0.02	0.11
b1b2[5,5]	-0.03	0.24	-0.01	0.10
b1b2[6,5]	-0.25	0.34	-0.08	0.13

TABLE XII: Sensitivity Aalysis ANOVA bias Br

	Mean model 1	SD model 1	Mean model 2	SD model 2
b0	-0.01	0.18	0.01	0.15
b1[1]	0.00	0.31	-0.00	0.14
b1[2]	-0.02	0.28	-0.00	0.13
b1[3]	0.17	0.26	0.06	0.13
b1[4]	-0.00	0.29	-0.00	0.14
b1[5]	-0.06	0.30	-0.02	0.14
b1[6]	-0.21	0.26	-0.07	0.13
b1[7]	-0.03	0.30	-0.01	0.14
b1[8]	0.04	0.23	0.01	0.12
b1[9]	0.00	0.23	-0.01	0.12
b1[10]	0.12	0.24	0.04	0.13
b1[11]	-0.01	0.31	-0.00	0.14
b2[1]	0.12	0.13	0.06	0.08
b2[2]	-0.12	0.13	-0.06	0.08
b1b2[1,1]	-0.03	0.25	-0.01	0.09
b1b2[2,1]	0.16	0.27	0.03	0.09
b1b2[3,1]	-0.03	0.21	-0.01	0.09
b1b2[4,1]	-0.04	0.24	-0.01	0.09
b1b2[5,1]	-0.07	0.25	-0.02	0.09
b1b2[6,1]	0.42	0.35	0.10	0.13
b1b2[7,1]	-0.04	0.25	-0.01	0.09
b1b2[8,1]	-0.06	0.19	-0.01	0.09
b1b2[9,1]	-0.13	0.21	-0.03	0.09
b1b2[10,1]	-0.15	0.23	-0.03	0.09
b1b2[11,1]	-0.04	0.25	-0.01	0.09
b1b2[1,2]	0.03	0.25	0.01	0.09
b1b2[2,2]	-0.16	0.27	-0.03	0.09
b1b2[3,2]	0.03	0.21	0.01	0.09
b1b2[4,2]	0.04	0.24	0.01	0.09
b1b2[5,2]	0.07	0.25	0.02	0.09
b1b2[6,2]	-0.42	0.35	-0.10	0.13
b1b2[7,2]	0.04	0.25	0.01	0.09
b1b2[8,2]	0.06	0.19	0.01	0.09
b1b2[9,2]	0.13	0.21	0.03	0.09
b1b2[10,2]	0.15	0.23	0.03	0.09
b1b2[11,2]	0.04	0.25	0.01	0.09

B.2 Marginal Dynamic Linear Model

C POSTERIOR ANALYSIS

Polling company bias model

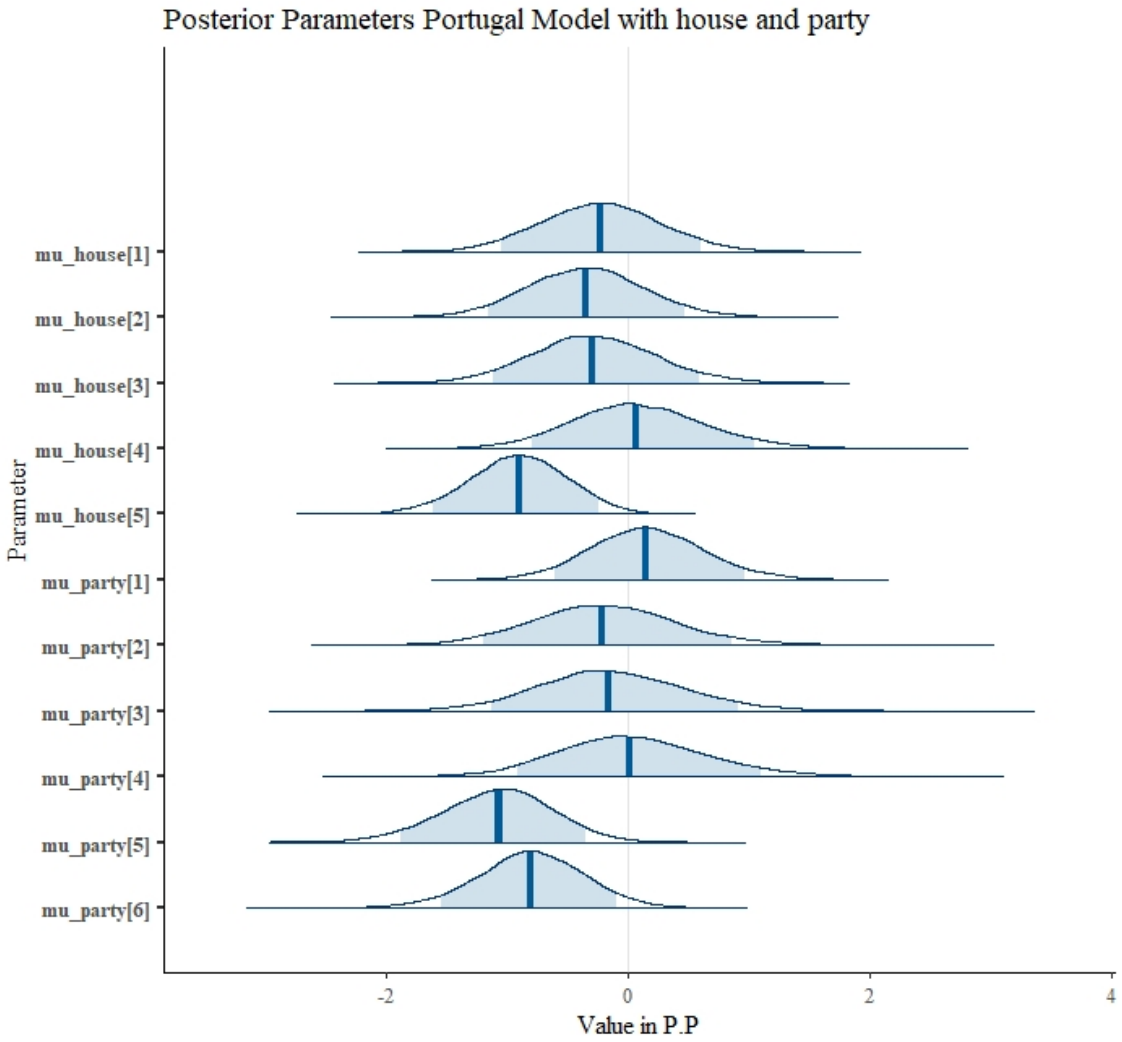

FIGURE 13: Posterior analysis for portuguese bias model with party and house

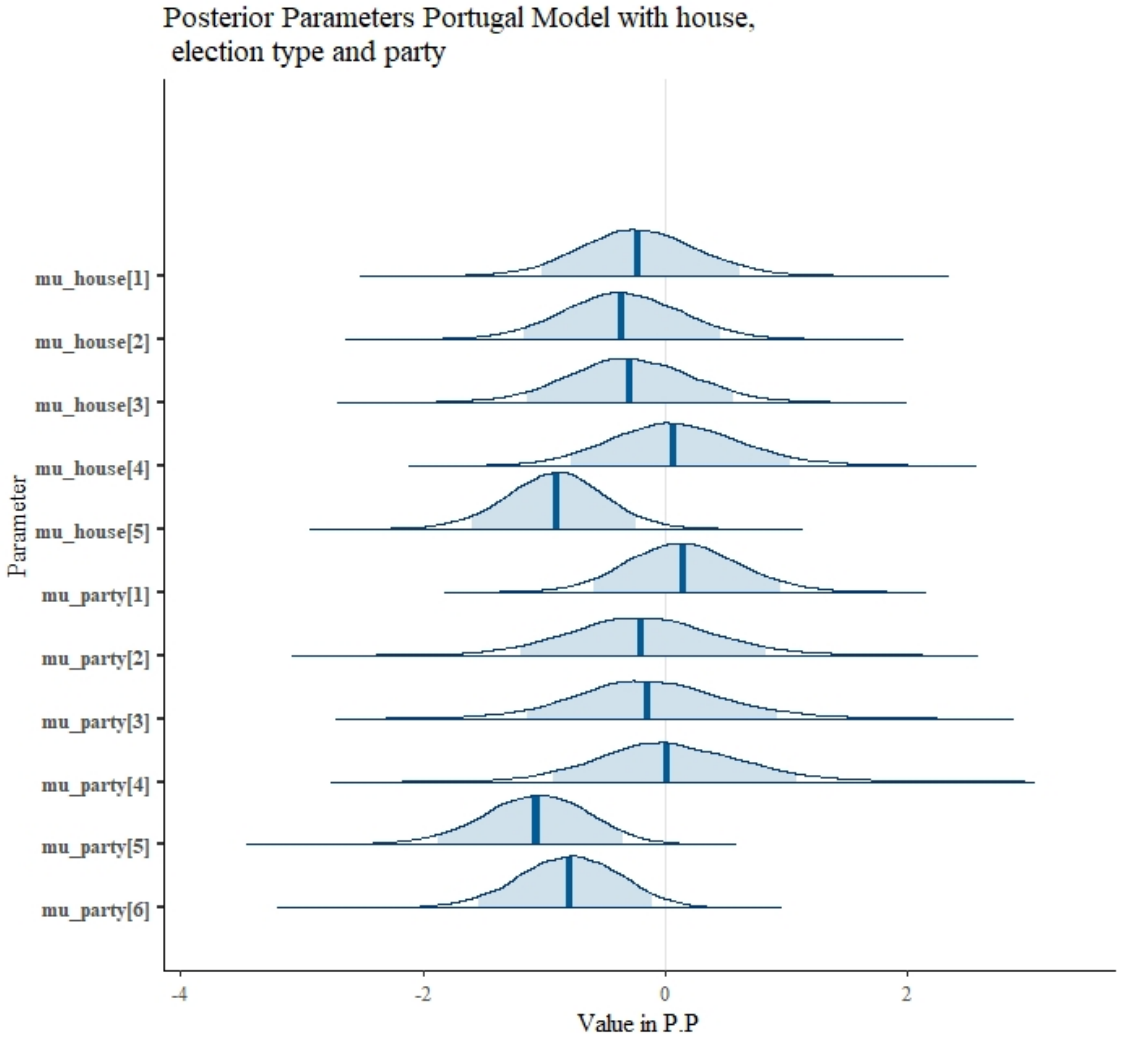

FIGURE 14: Posterior analysis for portuguese bias model with party, election type and house

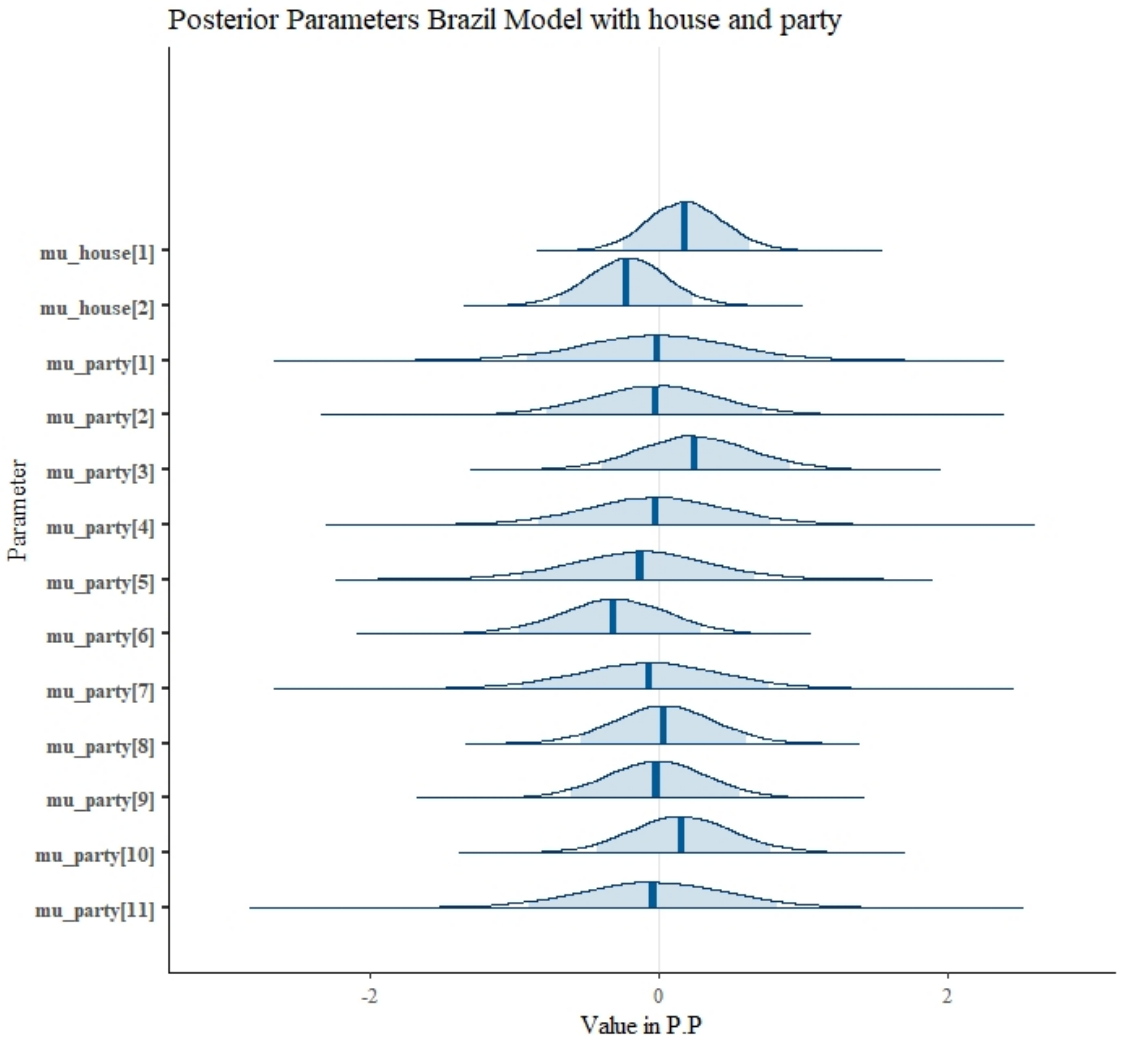

FIGURE 15: Posterior analysis for brazilian bias model with party and house

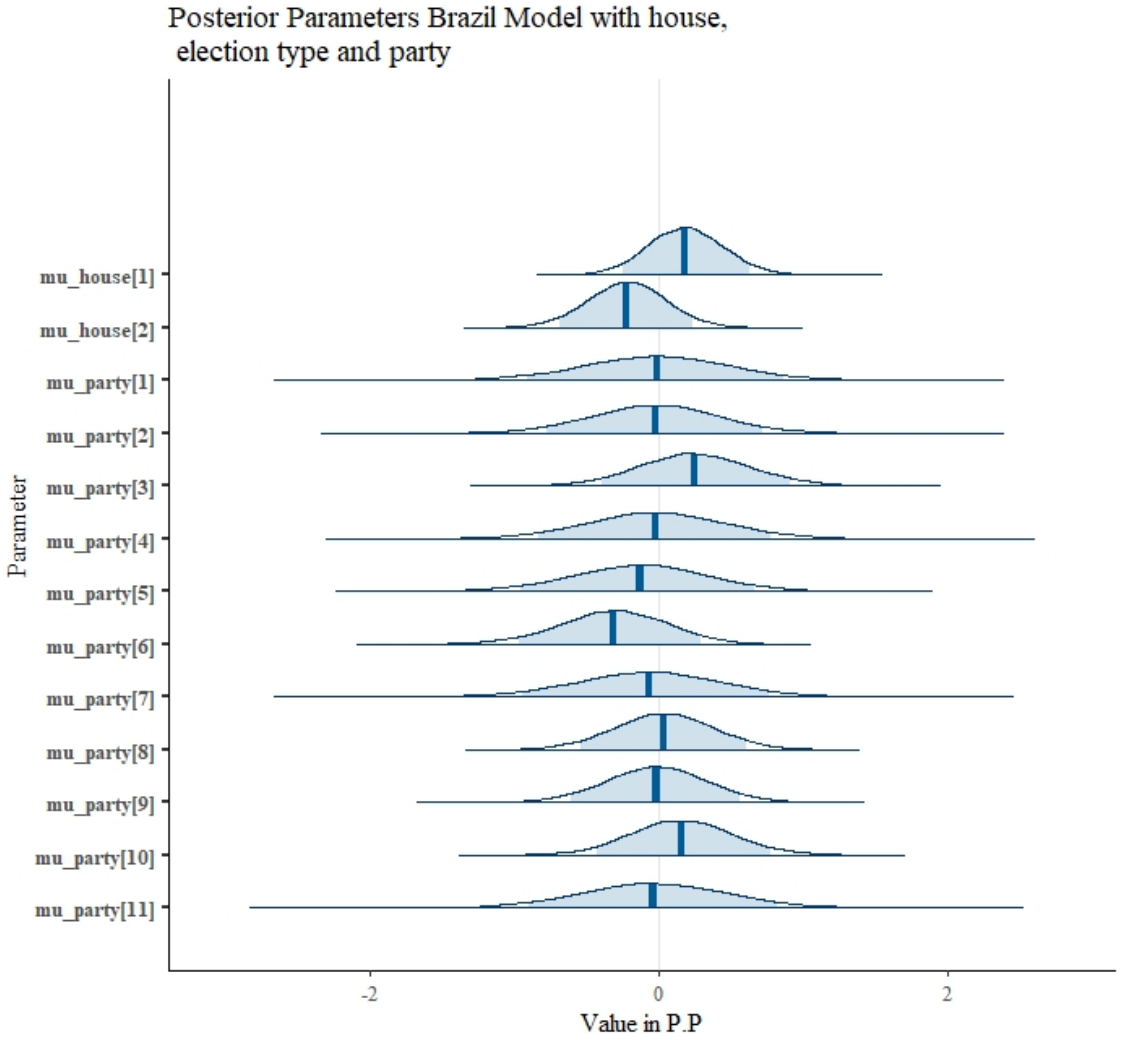

FIGURE 16: Posterior analysis for brazilian bias model with party, election type and house

Marginal model

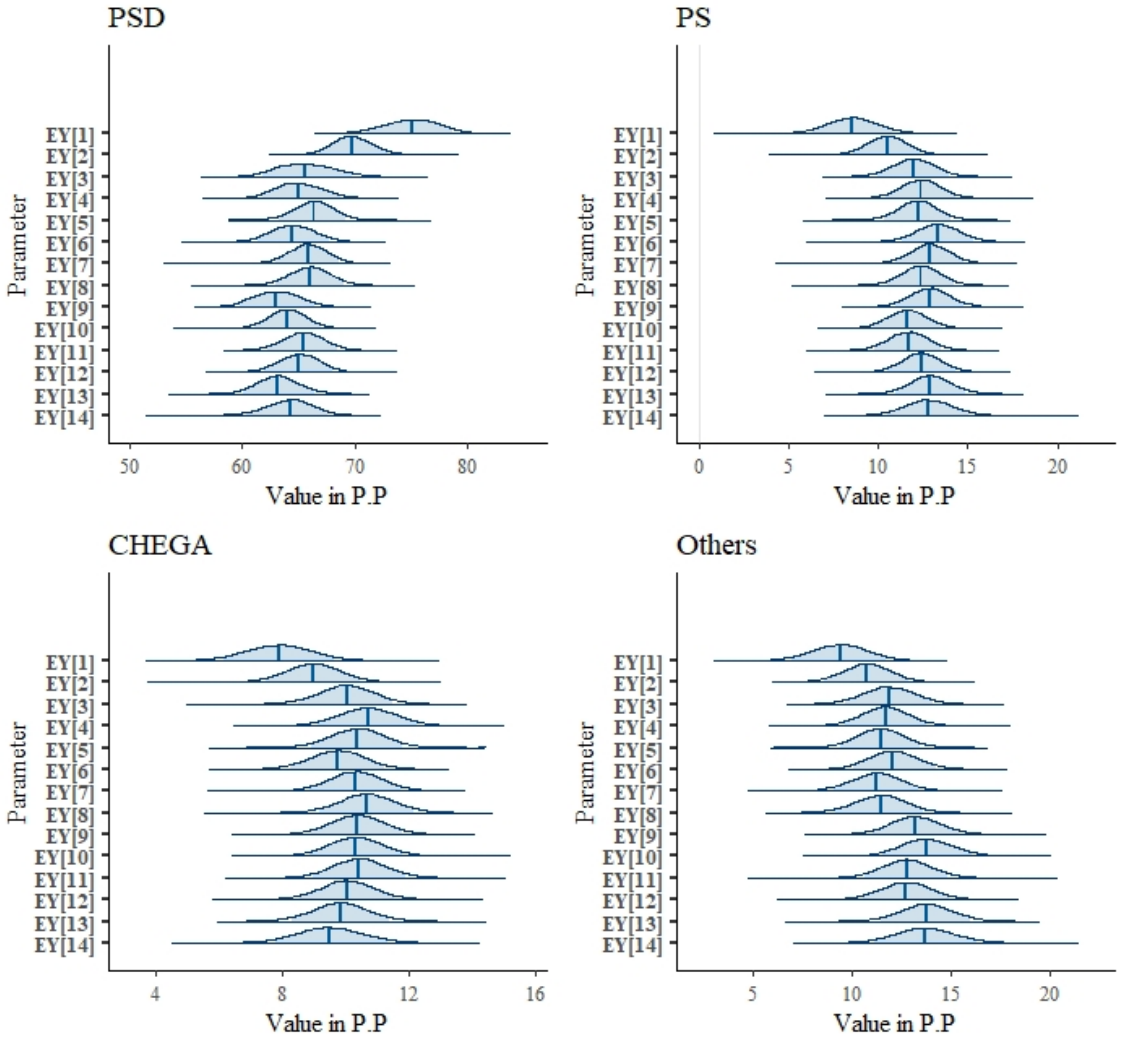

FIGURE 17: Posterior analysis for portuguese marginal model

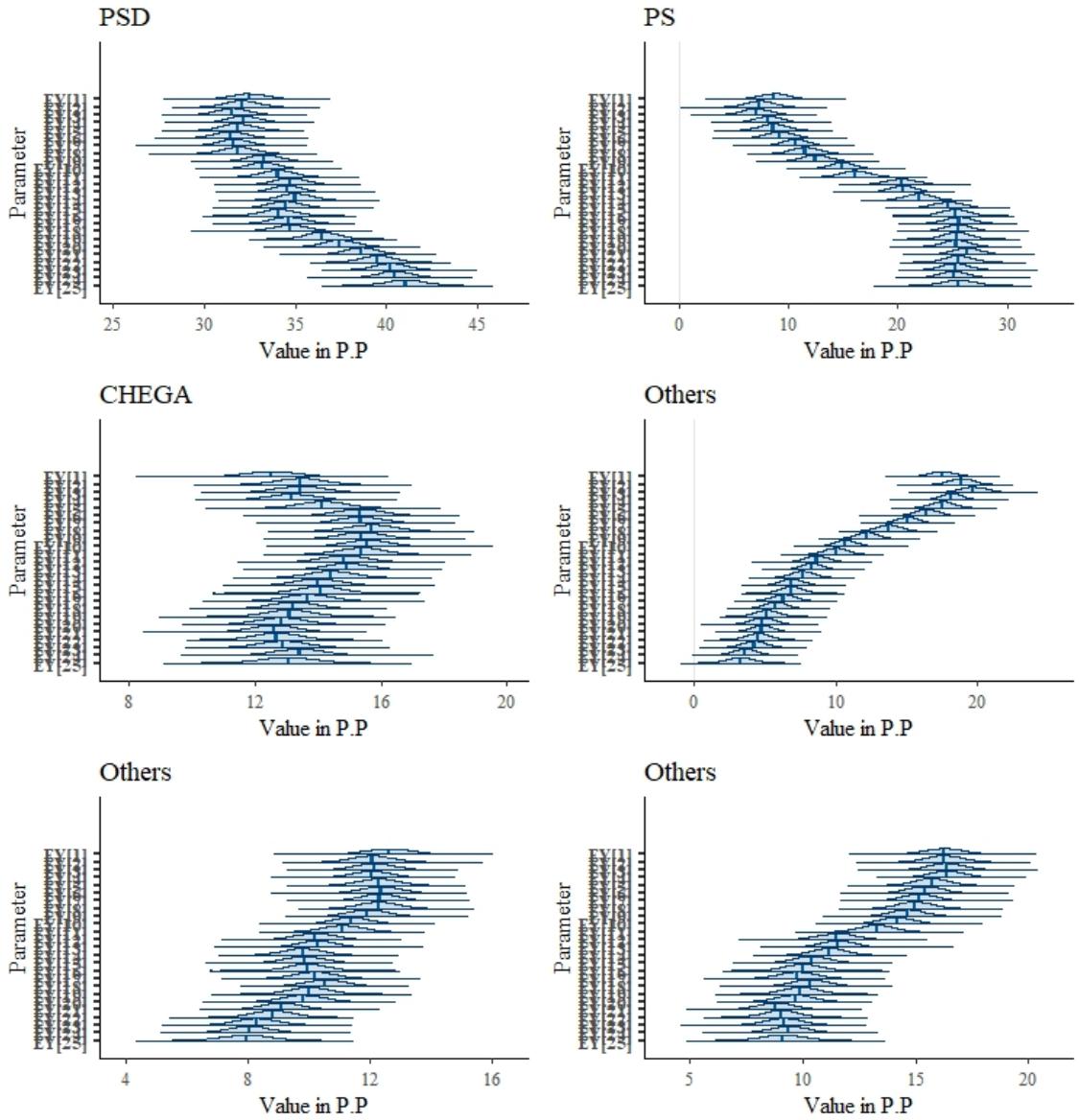

FIGURE 18: Posterior analysis for brazilian marginal model

D BAYESIAN STRUCTURAL

Portuguese bayesian structural regression

TABLE XIII: Bayesian structural models Portugal

	Model for PS	Model for others
Inflation_previous_year	1.03	−1.16
	[−1.21; 3.23]	[−2.29; 0.00]
GDP_Growth	−0.32	1.24
	[−5.11; 4.45]	[−1.39; 3.86]
Unemployment	−6.43	2.99
	[−14.66; 2.06]	[−0.85; 6.81]

* Null hypothesis value outside 95% credible interval.

TABLE XIV: Bayesian structural models Brazil

	Model 1	Model 2	Model 3
Inflation_previous_year	−0.16	0.18	−0.16*
	[−0.44; 0.14]	[−0.26; 0.60]	[−0.30; −0.00]
GDP_Growth	0.01	0.01	−0.00
	[−0.06; 0.07]	[−0.09; 0.11]	[−0.03; 0.03]
Unemployment	0.01	0.02	−0.03
	[−0.05; 0.06]	[−0.06; 0.10]	[−0.05; 0.00]

* Null hypothesis value outside 95% credible interval.

E EXTRA MATERIAL FOR THE JOINT DYNAMIC LINEAR MODEL

$$W = \begin{bmatrix} w_1^2 & \cdots & w_{1,(P-1)} \\ \vdots & \ddots & \vdots \\ w_{1,(P-1)} & \cdots & w_{(P-1)}^2 \end{bmatrix}.$$

FIGURE 19: Matrix W for the Joint Model, from (Stoetzer et al., 2019)