UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA POLÍTICA

VIRGINIA ROCHA DA SILVA

THE FAMILY CONNECTION: political dynasties and government transparency in Brazilian municipalities.

Recife
2022

VIRGINIA ROCHA DA SILVA

THE FAMILY CONNECTION: political dynasties and government transparency in Brazilian municipalities.

Tese apresentada ao Programa de Pós-Graduação em Ciência Política da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de doutor em Ciência Política. Área de concentração: Democracia e Instituições.

Orientadora: Mariana Batista da Silva

Recife

2022

Catalogação na fonte
Bibliotecária Valdicea Alves Silva, CRB4-1260

S586f Silva, Virginia Rocha da
The family connection: political dynasties and government transparency in Brazilian municipalities. / Virginia Rocha da Silva – Recife, 2022.
159 f.: il.; 30 cm.

Orientadora: Profª. Drª. Mariana Batista da Silva.
Tese (Doutorado) - Universidade Federal de Pernambuco, CFCH. Programa de Pós-Graduação em Ciências Políticas, Recife, 2022.
Inclui referências apêndices
Idioma Inglês.

1. Ciências Políticas. 2. Catalogação por assunto. 3. Sociedades comerciais. 4. Imposto - Custo. 5. Diferentes organizações - Exercício do poder. 6. Políticos dinásticos. 7.Brasil. I. Silva, Mariana Batista da (Orientadora). II. Título.

320 CDD (22. ed.) UFPE (BCFCH2023-001)

VIRGINIA ROCHA DA SILVA

THE FAMILY CONNECTION: political dynasties and government transparency in Brazilian municipalities.

Tese apresentada ao Programa de Pós-Graduação em Ciência Política da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de doutor em Ciência Política. Área de concentração: Democracia e Instituições.

Data de aprovação: 30 de agosto de 2022.

Profa. Dra. Mariana Batista da Silva (orientadora)

Universidade Federal de Pernambuco

Prof. Dr. Gabriela da Silva Tarouco (examinadora interna)

Universidade Federal de Pernambuco

Prof. Dr. Timothy J. Power (examinador externo)

Universidade de Oxford

Prof. Dr. Robert Gregory Michener (examinador externo)

FGV-EBAPE

Prof. Dr. Yuri Kasahara (examinador externo)

Universidade Metropolitana de Oslo

Ao meu pai, Mario, e às minhas tias Iraci e Maria José (in memoriam), com amor e saudade.

À minha mãe, Maria, e à minha irmã, Rebeca, com todo o meu amor e gratidão.

AGRADECIMENTOS

"E como vai a tese?". Essa pergunta, que as pessoas faziam por curiosidade ou para tentar oferecer ajuda, me trouxe ansiedade por meses. Só na prática percebi que escrever uma dissertação de doutorado é tão difícil quanto as pessoas dizem. Mas também é muito enriquecedor terminar o trabalho e perceber o quanto você aprendeu ao longo dos anos. Por isso, sou muito grata a todos que me ajudaram neste processo.

Em primeiro lugar, agradeço à minha mãe por todo o amor e apoio e por ser o meu exemplo mais notável de dedicação, perseverança, gentileza e humanidade. Agradeço à minha maravilhosa irmã por sempre me apoiar e me inspirar. Seu amor e palavras de conselho e encorajamento estão sempre comigo. Como não poderei citar todos os meus parentes e amigos queridos pelo nome, quero que saibam que agradeço por todo a força e carinho. Sou grata a vocês por serem compreensivos e gentis durante este processo.

Agradeço à minha orientadora, professora Mariana Batista, por sua orientação, seu imenso apoio e compreensão, que foram cruciais para minha pesquisa. Mariana tem sido um exemplo para mim desde que comecei minha pós-graduação e sempre me incentivou a dar o meu melhor. Ela é um grande exemplo de professora e uma pesquisadora brilhante. Tenho muita sorte de ter tido uma pessoa tão talentosa, competente e gentil como orientadora.

Também sou grata ao professor Timothy Power, que me orientou durante o meu doutorado sanduíche em Oxford. Além de um pesquisador notório, o Prof. Power é uma das pessoas mais generosas e gentis que conheci na academia. Agradeço seus comentários didáticos e seus conselhos em relação à minha tese. Também quero agradecer aos professores Gabriela Tarouco e Flavio Rezende, que me orientaram na graduação e no mestrado. Obrigada por me ajudarem a ser uma pesquisadora mais perseverante e curiosa.

Agradeço também aos professores Nara Pavão e Gregory Michener por seus excelentes e detalhados comentários sobre a minha pesquisa de doutorado, feitos desde a qualificação. Muito obrigada também à professora Gabriela Tarouco e aos professores Timothy Power, Yuri Kasahara e Gregory Michener por aceitarem fazer parte da minha banca e pela leitura cuidadosa da tese e seus comentários agregadores.

Eu apresentei e debati minha tese muitas vezes. Nesse sentido, agradeço os comentários das Professoras Andreza de Souza Santos, Marta Rocha e Allyson Benton e ao professor Ezequiel Ocantos, feitos em conferências e reuniões. Agradeço a todos os colegas do grupo de pesquisa Instituições, Política e Governo, coordenado pela professora Mariana Batista.

Agradeço também a Bernardo Medeiros, Cristiana Maglia, Leonardo Gill e Sara Picheth pelos comentários e amizade durante o Doutorado Sanduíche.

Aos meus queridos amigos da pós-graduação que ainda não foram mencionados, obrigada pelos comentários, apoio e amizade. Vocês tornaram suportáveis os momentos mais cansativos e estressantes. Palloma Marciano, Haína Coelho, Heloá Landim, Manuela Pereira e Amanda Domingos, não tenho palavras para o apoio que vocês me ofereceram durante esses anos, principalmente durante a pandemia. Obrigado pelas conversas e sugestões e por me ouvir quando eu estava ansiosa ou cansada. Fico imensamente feliz por termos nos apoiado e gostaria que cada aluna e aluno de pós-graduação pudesse ter essa rede de apoio. Palloma e Haína, um viva e um agradecimento especial a nossas reuniões do TGCP! Agradeço também a Marcus Torres, Anderson Henrique, Jorge Oliveira Gomes, Matheus Cunha, Bhreno Vieira, Rodrigo Lins, Everton Lira e Willber Nascimento. Agradeço a Cláudio Monteiro pela ajuda considerável com Python e ao apoio de Djair Lourenço durante esses últimos anos do doutorado.

No doutorado, eu tive a oportunidade de trabalhar como assistente de pesquisa em projetos dos professores Marcus Melo, Taylor Boas, Daniel Hidalgo, Yuri Kasahara, Timothy Power e Cesar Zucco. Aprendi muito com essas experiências e agradeço a confiança em meu trabalho e a oportunidade que me deram. As pesquisas dos professores Melo, Boas, Hidalgo e Kasahara me inspiraram a estudar dinastias políticas e por isso sou também muito grata.

Gostaria de agradecer a todos os professores, funcionários e alunos do Programa de Pós-Graduação em Ciência Política da Universidade Federal de Pernambuco. Aos meus professores, obrigada pelo conhecimento que me transmitiram. Aos membros da equipe, agradeço o tempo, atenção e ajuda. A todos os meus colegas, compartilhar esta jornada com vocês foi um prazer. Agradeço também aos funcionários do Centro de Filosofia e Ciências Humanas e ao Comitê de Ética em Pesquisa.

Apesar de desafiador, ingressar e concluir a pós-graduação também é um privilégio, especialmente num país desigual como o Brasil. Agradeço à CAPES pela bolsa de doutorado que me permitiu desenvolver pesquisas acadêmicas. À CAPES agradeço também pela bolsa do Programa de Doutorado Sanduíche, que me possibilitou estudar por um semestre em Oxford. Na época, a Universidade de Oxford cobrava uma taxa de 2 mil libras por trimestre, que não era coberta pela bolsa do PDSE. Sou, portanto, muito grata ao *Latin American Centre* de Oxford, por ter dispensado a taxa de um dos trimestres, às várias pessoas, próximas ou não, que, por meio de doações, que me ajudaram a pagar a segunda taxa. Muito obrigada pela confiança e apoio.

"(...) Secrecy is corrosive: it is antithetical to democratic values, and it undermines democratic processes (…)." (STIGLITZ, 1999, p.1).

RESUMO

Por que alguns governos locais cumprem o acesso à informação enquanto outros não? O Brasil aprovou a Lei de Acesso à Informação (LAI), um marco legal da transparência pública, em 2011. No entanto, a regulamentação da LAI ainda é baixa: em 2014, apenas 20% dos municípios haviam regulamentado a lei. Evidências da literatura indicam que variáveis políticas, como margem de vitória e insegurança eleitoral, podem afetar os incentivos que os políticos têm para cumprir a transparência. Ainda assim, existem também fatores externos como o tamanho da população e o PIB per capita que podem afetar conformidade. Proponho a inclusão das dinastias políticas nesta discussão para dar conta das diferentes formas de organização e exercício do poder. Sugiro que os políticos dinásticos possuem incentivos diferentes em relação à transparência. Argumento que prefeitos dinásticos estão negativamente associados ao acesso à informação, e esse efeito está condicionado a outras variáveis políticas, quais sejam, margem de vitória, alinhamento partidário com o governador e alinhamento partidário com o presidente. Usando dados observacionais do TSE, do IBGE e da CGU para 2015 e 2016, testo este argumento com um modelo de regressão binomial negativa de zero inflado para três especificações diferentes da variável dependente (pontuação de regulamentação, pontuação da qualidade da informação e pontuação de transparência passiva). Quanto aos prefeitos dinásticos, proponho uma nova abordagem para identificar dinastias políticas: usar os resultados de pesquisa do Google para identificar indícios de parentesco dinástico nas notícias locais. Eu uso a API JSON de busca customizada da *Google* e o *Beautiful Soup* do *Python* para coletar resultados do Google e extrair texto das páginas da *web*. Em seguida, uso *ensemble learning* para classificação binária (isto é, se os textos indicam parentesco dinástico, 1, ou não, 0). Os resultados indicam que prefeitos dinásticos, quando alinhados com o presidente, são importantes para entender a variação local no cumprimento da transparência em cidades que já cumprem em alguma medida a lei. Para cidades que ainda não cumprem, dinastias podem influenciar negativamente suas chances de *compliance*, quanto o prefeito é do mesmo partido do governador. Como contribuição, a pesquisa sugere uma nova variável política para o debate sobre transparência local, com uma nova forma de identificar políticos dinásticos. Essa nova medida pode ser usada para identificar tais políticos e para coletar de forma automatizada informações mais detalhadas sobre as dinastias locais.

Palavras-chave: acesso à informação; transparência; conformidade; dinastias políticas; políticos dinásticos; Brasil.

ABSTRACT

Why do some local governments comply with access to information while others do not? Brazil passed its Access to Information Law (AIA), a legal milestone of public transparency, in 2011. Nonetheless, in 2014 only 20% of the cities had passed the law. Evidence from the literature indicates that political variables, such as the margin of victory and electoral insecurity, can affect politicians' incentives to comply with transparency. Still, there are also external factors, such as the size of the population and the GDP per capita, that can affect compliance. I propose the inclusion of political dynasties in this discussion to account for different forms of organizing and exerting power. I suggest that dynastic politicians have various incentives regarding transparency. I argue that dynastic mayors are negatively associated with access to information. This effect is conditioned to other political variables, i.e., the margin of victory, party alignment with the governor, and party alignment with the President. Using observational data from the TSE, IBGE, and CGU for 2015 and 2016, I test this argument with a zero-inflated negative binomial regression model for three different specifications of the dependent variable (regulation score, quality of information score, and passive transparency score). As for dynastic mayors, I propose a novel approach to identifying political dynasties: using Google search results to identify indications of dynastic kinship in local news. I use Google's Custom Search JSON API and Python's Beautiful Soup to collect Google results and extract text from the web pages. I then use ensemble learning for binary classification (that is, whether the texts indicate dynastic kinship, 1 or not, 0). The results suggest that when aligned with the president, dynastic mayors are essential to understand the local variation in compliance with transparency in cities that already comply to some extent with the law. For cities that do not yet comply, dynasties can negatively influence their chances of compliance when the mayor is from the same party as the governor. As a contribution, the research suggests a new political variable for the debate on local transparency, with a new way of identifying dynastic politicians. This new measure can be used to identify such politicians and to collect more detailed information about local dynasties automatically.

Keywords: access to information; transparency; compliance; political dynasties; dynastic politicians; Brazil.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABREVIATIONS

Absent information	NA
Access to Information	AIA
Alagoas	AL
Antônio Carlos Magalhães	ACM
Application Programming Interface	API
Beautiful Soup	Bs4
Ceará	CE
Controladoria Geral da União	CGU
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior	CAPES
Escala Brasil Transparente	EBT
Extensible Markup Language	XML
Freedom of Information	FOI
Goiás	GO
Google Search Engine	GSE
Gross Domestic Product	GDP
Hypertext Markup Language	HTML
Identification	ID
Instituto Brasileiro de Geografia e Estatística	IBGE
JavaScript Object Notation	JSON
Maranhão	MA
Minas Gerais	MG
Ministério Público do Estado do Rio de Janeiro	MPRJ
Ministério Público Eleitoral	MPE
Movimento Democrático Brasileiro	MDB
Not confirmed	NC
Partido Comunista do Brasil	PC do B
Partido da Social Democracia Brasileira	PSDB
Partido do Movimento Democrático Brasileiro	PMDB
Partido dos Trabalhadores	PT
Partido Socialista Brasileiro	PSB
Partido Trabalhista Brasileiro	PTB
Pernambuco	PE
Piauí	PI
Random Forest	RF
Receiver Operating Characteristic	ROC
Rio de Janeiro	RJ
São Paulo	SP
Support Vector Machine	SVM
Tribunal Superior Eleitoral	TSE
United Kingdom	UK
Zero-Inflated Negative Binomial Regression	ZINB

TABLE OF CONTENTS

1 INTRODUCTION

"Of course, he is my child; I intend to benefit my child...yes, I do (...)" (GLOBO NEWS, 2019, free translation). That could be a regular sentence coming from a supporting father if it had not been said by the President of Brazil, Jair Bolsonaro, in a press conference about the decision to choose his son, Eduardo Bolsonaro, as the Ambassador of Brazil to the United States. Eduardo had no technical credentials for the position. President Bolsonaro is also a record holder in the use of the Access to Information Act (*Lei de Acesso à Informação*) - not to strengthen the law, but to justify the creation of 100 years of secrecy about information that is in the public interest, such as the record of how many times his children, also politicians, visited the President in his official residence. The presidency claims that this information is personal in nature (ALESSI, 2021).

There are several anecdotes regarding political dynasties, not only in Brazil but worldwide. In recent years, studies on political dynasties - i.e., when you have two or more elected officials from the same family in consecutive elections or not (DAL BÓ *et al*., 2009, BRAGANÇA *et al*., 2015) - have started to flourish in the political science and the political economy. Topics such as quality of government (BESLEY; REYNAL-QUEROL, 2017; DANIELE; GEYS, 2015; TUSALEM; PE-AGUIRRE, 2013), corruption (OLIVEIRA; SOUZA, 2022), representation (CROWLEY; REECE, 2013), and effects on the party system (AMUNDSEN, 2016; CHHIBBER, 2013), for instance, have included political dynasties as explanatory variables.

Perissinoto (2019) claims that the study of political and social phenomena could benefit from looking beyond the institutional arrangements and the economic structure and examining political elites and how they act in certain scenarios. I argue that Political dynasties represent, to a certain level, an informal organization of power that interacts with formal and modernized public realms. Thus, including dynastic politicians in the analysis of public transparency could help to understand the variation in compliance with access to information in Brazilian municipalities.

Brazil passed its Access to Information Act (AIA), a legal milestone of public transparency, in 2011. Nonetheless, according to data from the Brazilian Institute of Geography and Statistics (2014), the regulation of AIA is still low: only 20% of the cities have passed the law. *Why do some local governments comply with transparency while others do not?* That is the question I try to answer in this dissertation.

Evidence from the literature indicates that political variables, such as margin of victory and electoral insecurity, can affect the incentives politicians have to comply with transparency. Still, there are also external factors such as the size of the population and the GDP per capita that can affect compliance (BATISTA *et al*., 2022; BERLINER; ERLICH, 2015; MICHENER; NICHTER, 2022).

I propose that in cities where the chief of the Executive is dynastic, the level of compliance, both *de jure* and *de facto*, will be lower. I argue that dynastic politicians act in the constant tension between the public and the private realm. Hence, to better understand if and how they are linked to compliance, it is important to consider other aspects that denote the level of consolidation of a given dynasty. I suggest three additional hypotheses focused on the conditional effect of dynastic mayors on compliance considering the margin of victory, the party identification between the mayor and the governor, and the party identification between the mayor and the President.

Methodologically, I focus on compliance with access to information. *De jure* compliance refers to actions to implement legal measures, and *de facto* compliance regards how these measures are delivered to the public (MICHENER; NICHTER, 2022). The national regulation of AIA in 2012 made the law valid for all Federal entities, yet municipalities and states need to regulate their legal device (BATISTA, 2017). Regulation concerns whether the municipal government passed its legal device or not and the quality of the law - which refers to *how* the law was regulated. Passive transparency regards how well governments respond to access requests (Access to Information Act [AIA], 2011).

Using observational data from the Superior Electoral Court (TSE, from the initials in Portuguese), the Brazilian Institute of Geography and Statistics (IBGE, from the initials in Portuguese), and the Comptroller General of Brazil (CGU, from the initials in Portuguese) for 2015 and 2016, I test my argument with a zero-inflated negative binomial regression model. This is a more suitable approach for the zero-inflated distributions of the data from the Brazilian Transparency Scale (in Portuguese, *Escala Brasil Transparente* – EBT), which I use to measure *de jure* and *de facto* compliance (LONG, 1997). It means that EBT data contains several 0 values for three different specifications of my dependent variable (regulation score, quality of information score, and passive transparency score).

I use two measures of political dynasties: automated classification and manual coding. For the former, I focus on the 2012 municipal elections, when the mayors governing the cities evaluated by CGU from 2015 to 2017 were elected. For the latter, I use data from 2000 to 2016 since the manual search started before the 2020 municipal elections. For the automated classification, I collected data on the elected mayors with Custom Search JSON API, which is an application that allows you to create a costumed search engine on Google. Then, you can collect google search results delivered by your engine. In this way, I could automatize the manual collection of information about the mayors.

Yet, Google's API only deliver the title, link, and snippet of google results. Since each result is a web page, I used a Python library for web scraping called Beautiful Soup. With this, I extracted the page content of each result. I then used supervised learning to classify mayors as dynastic or not. I manually coded more than 800 texts according to the presence (1) or absence (0) of information indicating a dynastic linkage for the elected mayor. I then tested several classification algorithms, such as random forest and support vector machine. I combined the algorithms with the best performances to arrive at a final classification via ensemble learning.

Regression results indicate that dynastic politicians' effect on compliance with access to information may be significant when conditioned to whether the mayor and the President, and the mayor and the governor are from the same party. Also, different from what I argued, dynastic mayors are positively associated with transparency, not the contrary, for cities that have already started to comply with the law. For cities that have not complied yet, but can do it in the future, dynastic politicians, when aligned with the governor, may reduce the chances of the municipalities start to comply. Beyond the interaction between dynastic mayors and party identification with the President and the governor, the log of the population, the log of GDP per capita, and reelection are also important parameters when it comes to compliance *de jure* and *de facto* with the access to information.

In that way, this research offers two main contributions: testing a new variable in the debate of compliance and suggesting a novel approach to the automated identification of dynastic politicians. This is the first version of this measure, which contains important limitations that can be improved in future research, as I discuss in the following chapters. Yet, it has the potential to help in the identification of dynastic politicians in the sense that it can be combined with other measures, such as name matching; it helps to find relevant information about politicians, which can aid qualitative research in the future, and, because it gathers a

diverse set of information about each politician, could be used to build other measures related to political dynasties, beyond identifying the politicians who have dynastic ties.

This dissertation is organized as follows. In the next chapter, I discuss the definition of compliance and the primary theoretical debate on compliance with access to information at the local level in Brazil. In the following chapter, I discuss the definition of political dynasties and the discussion regarding this topic. I then present my argument and its empirical implications. In the fourth chapter, I present my proposal for a new measure of political dynasties. Subsequently, in chapter 5, I test the main hypothesis of this thesis using zero-inflated negative binomial regression; and discuss the main analysis results. Finally, I conclude this research by discussing its main findings, contributions, and research agenda.

2 COMPLIANCE WITH ACCESS TO INFORMATION: A BRIEF REVIEW

Among the transparency initiatives examined by scholars, especially in political science, the Freedom of Information (FOI) laws has received the most attention. These instruments consist of legal provisions allowing access to public information as a government duty. Thus, they differ from the Right to Information, which has the objective of establishing access to information as a citizen's right (MEIJER, 2015).

FOIs are relevant for institutionalizing transparency in ways that make it difficult to weaken them later (BERLINER, 2014). Three main factors characterize them. The first is the guarantee of the right of access to information for voters, civil society organizations, businesses, etc. All of these groups rely on government information. Thus, information assumes the role of a commodity, and the government can profit illegally from its scarcity (STIGLITZ, 1999). Therefore, FOIs limit politicians' control over information and, consequently, their ability to profit from its availability (bribery, information leakage to privileged groups, etc.) (*ibidem*). The second is to enable access to information as a monitoring tool in the logic of the principal-agent theory. The third is the possibility of limiting future governments, guaranteeing that the government will maintain the law (BERLINER, 2014)

After the 1970s, a wave of adoptions of the so-called Freedom of Information Laws came to fruition. The dissemination of FOI regimes around the world was one of the consequences of greater demand for the visibility of information, especially for institutions such as the World Bank and the International Monetary Fund, as loans started to be conditioned to the implementation of transparency mechanisms (MICHENER; BERSCH, 2013).

Although Sweden was the first country to adopt a transparency law - with the Freedom of the Press Act in 1766, the diffusion of such laws into their modern conception came about through the implementation of the Freedom of Information Act in 1966 United States. Initially, the process occurred slowly and then more intensively from the 2000s (MEIJER, 2015). In Latin America, the first country to adopt a transparency law was Colombia in 1985. Mexico, which enacted its law in 2002, has one of the strongest FOIs in Latin America. Brazil, however, may be considered a late adopter for having published the AIA only in 2011 (FOI, 2017).

In fact, the Brazilian FOI was hardly passed (MICHENER, 2011). Senator Fernando Collor de Melo, a former President who resigned his term to avoid impeachment after a corruption scandal, held the project in the Foreign Relations Commission for as long as possible. By that time, the Foreign Relations Ministry and the Armed Forces had already stated

that they had no opposition regarding the establishment of a deadline for the secrecy of public information. The delay to vote on the bill created an embarrassment to then President Dilma Rousseff when she took part in the launch of the Open Government Partnership, led by then President of the United States Barack Obama (MICHENER, 2011). Finally, in November 2011, the law was sanctioned.

The Brazilian FOI law, called the Access to Information Act (AIA) (*Lei de Acesso à Informação*, in Portuguese), establishes procedures that must be adopted to ensure the disclosure of information, and it is applied to entities directly or indirectly controlled by the Federal Government, States, Federal District, and Municipalities, as well as private non-profit entities which received public funds. Its general assumption is that secrecy should be an exception, while the disclosure is the rule. The law encompasses two types of transparency: active and passive. The first relates to the information spontaneously made public by the government, while the latter reflects information disclosed by request (BRASIL, 2011; BATISTA, 2017).

From May 2012, the law had immediate application at the federal leve; however, at the subnational level, it still needs to be regulated by states and municipalities. According to AIA's articles 9 and 45, states and cities are responsible in each locality for creating a service to provide information to citizens. This service's structure should enable i) the welcoming and instructions provisions to citizens on how to access information; ii) informing the flow of documents in its organization, and iii) the filing of documents and the requirement of information by citizens.

Federal entities should also hold public hearings and consultations to stimulate public participation and other forms of access to information dissemination (BRASIL, 2011). In that way, AIA implementation depends on an excellent level of the quality of its local regulation. Local political actors have the autonomy to decide how and when to regulate their transparency dispositive (BRASIL, 2011; BATISTA, 2017).

The main object of this chapter is compliance with access to information in Brazilian municipalities. I approach this theme through different lenses. First, I discuss the definition of transparency and the focus on the compliance with access to information in its both aspects, *de jure* and *de facto*. Then, I briefly present data and studies on the level of compliance in Brazilian municipalities. Subsequently, I present the primary evidence of the literature regarding these two types of compliance. Finally, I conclude this chapter with a concise discussion of a new

variable that could be considered in the studies of compliance with transparency: political dynasties.

2.1 DEFINING TRANSPARENCY

Stiglitz (1999) argues that there should be a solid normative predisposition for transparency. The reason is that "(...) secrecy is corrosive: it is antithetical to democratic values, and it undermines democratic processes" (STIGLITZ, 1999, p.1). Since public officials produce information, it becomes a public good. Thus, secrecy means the suppression of information and theft of a public good (STIGLITZ, 1999). This argument is aligned with the advocacy of transparency as an end. On the other hand, one can also think of public transparency as a means to fight corruption, increase political participation, and improve governance (BERLINER, 2015, MEIJER, 2015; BATISTA, 2017).

Despite the dissemination of transparency around the world, its definition is ambiguous. For Meijer (2013), transparency consists of an institutional relationship of information exchange on an actor's performance. Thus, the availability and accessibility of information over time and the fact that external actors can use it are critical features of transparency. In this sense, it has complex dynamics, being a disputed domain among several political actors, government stakeholders, citizens, scientists, and journalists (MEIJER, 2015). For Heald (2006), to clarify the concept of transparency it is necessary to contemplate its distinct directions and varieties. He suggests four types of transparency directions that vary as to what can be observed (either inwards or outwards) and whether vertical or horizontal transparency (RODRIGUES, 2019).

The first direction is upwards (vertical). It refers to a hierarchical relationship in a principal-agent model, where the principal, who would be superior hierarchically, can observe the behavior and results of his subordinate (agent). In the second type of direction, downwards (vertical), those who are governed can observe the behavior and results of their leaders, being a direction of transparency linked to the concept of accountability and democracy. The third direction, outward (horizontal), concerns the possibility for the agent to observe what happens outside the organization of which she is a part. It is important because it allows the understanding of the environment in which the organization is inserted. It also allows you to observe your competitors.

Finally, the fourth direction is inward (horizontal), that is, the ability to observe what happens within the organization. This type of transparency is mainly linked to the Freedom of Information (FOI) legislation and social control mechanisms (HEALD, 2006). FOI defines

access to information as a government obligation and a citizen's right (KOSACK; FUNG, 2014). Thus, such direction is associated with surveillance and monitoring among peers. This is valid for organizations and citizens, as they can monitor and report deviations to the competent authorities. However, privacy issues define limits to this type of transparency (HEALD, 2006). Furthermore, the monitoring aspect of this object is also related to Meijer's (2013) definition of transparency as the disclosure of government information to enable its monitoring by external actors.

Nonetheless, adopting and implementing FOI regimes does not automatically result in effective transparency. Comparing nominal to effective transparency, Heald (2006) argues that an example of the issues that can affect the effectiveness of transparency would be the various problems in the implementation of fiscal transparency in the United Kingdom, e.g., the privileged dissemination of information by the government to politically aligned newspapers. Another challenge was the capacity of the users of transparency to process the information released. When those aspects are present, we have an illusion of transparency.

Also, considering the constitutive elements of transparency, i.e., visibility and inferability (MICHENER; BERSCH, 2013), Rodrigues (2019) explains that in nominal transparency inferability is limited, but visibility is present. Visibility regards to what extant an information is complete *and* easily found. Inferability, on the other hand, denotes the quality of information. Inferability is present when information is disaggregated *or* simplified, *or* verifiable (MICHENER; BERSCH, 2013). In nominal transparency, information is made available, but there is no guarantee that individuals will be able to process and use it, that is, transparency is not effective (HEALD, 2006).

That does not mean that nominal transparency implies non-compliance with a country's legislation (RODRIGUES, 2019). Instead, this means that nominal transparency, which can be generated by a freedom of information law such as AIA, does not guarantee that transparency will be effective. In that way, the movement from nominal to full transparency indicates an institutional effort to overcome the limitations to achieve effective transparency (*ibidem*).

This connection between nominal transparency and compliance could be classified as what Michener and Nichter (2022) call *de jure* compliance. The extent to which governments follow legal guidelines, through regulation, active and passive transparency represents the level of compliance with the law. In respect to compliance at the local level in Brazil, the monitoring is based on the federal law. It means to say that active and passive transparency, for instance,

regards different aspects of the law and how the governments follow and deliver what was legally determined.

Michener and Nichter (2022) draw attention to two types of compliance with transparency: *de jure* compliance and *de facto* compliance. The first type covers the actions made by public officials regarding the implementation of transparency legal measures. One example is the regulation of the access to information law at the local level. The second type of compliance relates to how legal measures translate into service delivery. The authors exemplify this concept with passive transparency. That is, how well governments respond when citizens request information.

Based on this reasoning, Michener and Nichter (2022) argue the distinction in the types of compliance is important to achieve a more accurate measurement of what concerns the local level of transparency. In line with this discussion, the present research is focused on the level of both types of compliance - *de jure* and *de facto* - with access to information in Brazilian municipalities.

2.2 COMPLIANCE WITH ACCESS TO INFORMATION IN BRAZIL

Brazil has been experiencing expressive setbacks with transparency in the last years. On January 24, 2019, current Vice-president (then acting President), Gen. Hamilton Mourão, attempted to make AIA more flexible by issuing a decree to expand the range of authorities who would be able to classify information as confidential (FOLHA DE SP, 2019). Activists repudiated the change (FOLHA DE SP, 2019A). On February 19 of this year, the Chamber of Deputies approved the urgency of a bill that repeals the effects of the decree signed by the Vice-president (FOLHA DE SP, 2019B) and on February 26, 2019, President Jair Bolsonaro announced the withdrawal of the decree (FOLHA DE SP, 2019c).

Following this episode, the Federal Government classified several information with the higher level of secrecy according to the Access to Information Act: 100 years of secrecy. Information ranged from the President vaccination card to the registering of the meetings between the President and church pastors involved in a scandal arising from the accusation that the Ministry of Education favored some municipalities in exchange for bribes in gold bars to the then minister (ANGELO, 2022).

These examples send an alert that even a strong law, as considered by experts (MICHENER, 2011), that have been in force for a couple of years can be significantly undermined if political leaders are not committed to complying with it. In this sense, Brazil still

has a long way to go in terms of compliance. In an evaluation of Brazil's Access to Information, Michener et al. (2018) find that less than one request in their sample was answered by Brazilian agencies. Half of the requests exceeded the time limits established by law. The results indicate a low compliance to access to information in Brazil.

Similarly, the compliance with the Access to Information Act (AIA) in Brazilian municipalities is still low (CONTROLADORIA GERAL DA UNIÃO, 2020). In 2016, of the 2,328 municipalities randomly selected for evaluation of the Brazil Transparent Scale (*Escala Brasil Transparente*), only 9% of municipalities had a high score (9 - 10), while 21.3% had a low score (0 - 0.99). Over half of the cities analyzed (69.8%) were evaluated with scores below 7.

Michener et al. (2021) evaluate compliance with transparency federal laws in Latin America and show that the compliance rates at the federal level are 63% and 57% for passive and active transparency in Brazil, respectively. At the local level, the rates are 20 and 53%, for passive and active transparency, in that order. The authors draw attention to two gaps revealed by these results: 1) compliance rates to the federal law at the local level are lower than those at the national level, especially for passive transparency, and 2) compliance rates to the federal law are much smaller for passive transparency than for the active one.

In terms of active transparency, Figueiredo e Gazoni (2016) presents descriptive and qualitative data of Rio de Janeiro's municipalities and concludes there is still a long way to go in terms of the real-time release of relevant information to the public. Lunkes et al. (2015) also find that capital cities need to improve the availability of information in their so-called transparency portals.

Regarding passive transparency, Raupp and Pinho (2016) investigate compliance in cities with populations greater than 300,000 people using a qualitative approach. Their results show that only 15,19% of the city councils gave satisfactory responses to the author's information requests. Also, 84,81% of the municipal legislative institutions did not fully comply with the passive transparency requirements determined by AIA.

This bleak scenario for access to information indicates transparency is not effective as, in general, the government does not correctly provide information that can be used by external actors, indicating there is an illusion of transparency, in the terms of Heald (2006). This panorama is aggravated by the practical absence of sanctions on localities that do not comply with AIA. The law establishes consequences for public servants who disrespect the law but not for federal entities (BRASIL, 2011).

Among the institutions that monitor AIA's application at the local level little can be done to enforce the law. The *Controladoria Geral da União* created, in 2013, the Brazil Transparent Program (*Programa Brasil Transparente*), for which the Brazil Transparent Scale is a sub-product. This program provides training, supporting material, and a source code so local governments can offer an online platform for transparency. Municipalities that do not regulate the law, however, do not have to face any sanctions (CGU, 2020). In the next section, I will discuss the main factors explored in the literature about compliance with transparency, centering the discussion on the case of Brazil.

2.3 WHAT EXPLAINS THE VARIATION IN COMPLIANCE THE AIA?

Research about local compliance with transparency has been focused on what is called active or proactive transparency, that is, how governments release information and fulfill their legal responsibility with disclosure. On the other hand, reactive or passive transparency is also crucial, especially with the increase in the platforms that enable online request of information (MICHENER ET AL., 2018; MICHENER; NICHTER, 2022). With that in mind, I present in the next subsections the main explanations mobilized by the literature to understand *de facto* and *de jure* compliance.

2.3.1 *De jure* compliance

Since *de jure* compliance involves the regulation of FOI laws and the active or proactive transparency of governments, one starting point to understand what drives this compliance is to focus on its adoption. It means to say that it is important to know the motivations of politicians who have passed an FOI law. While some studies that investigate the dissemination of freedom of information laws use the policy diffusion framework (BATISTA, 2017, BATISTA *et al*., 2022), others dive into the effects of internal variables such as political competition, for instance.

In what concerns policymakers' motivations to adopt disclosure initiatives there is an assumption that the costs of adopting a Freedom of Information Law and likewise initiatives are greater than its benefits, and decision-makers would not have clear incentives to promote transparency (BERLINER, 2014, BERLINER, 2017, SCHNELL, 2017). Opacity, as opposed to a transparent government, gives politicians the ability to accumulate private gains, either because they control information that is a scarce good, or because misbehavior would be harder to detect and punish (FEREJOHN, 1986; STIGLITZ, 1999; BERLINER, 2014).

Consequently, a central puzzle about the decision of regulating transparency initiatives is: why do politicians adopt a policy that will constrain themselves? In general, incumbents who seek reelection in a competitive dispute would benefit from adopting transparency policies as a signal of credible commitment to fighting corruption and improving governance (BERLINER, 2014). Regulation of transparency could also be only a "cheap signal". That is, policymakers adopt an FOI law, but expect to control its implementation process and its consequences (SCHNELL, 2017). In other words, they can pass weak mechanisms or even not provide the structure necessary to guarantee the right to access information.

However, Schnell (2017) argues that decision-makers do not have that control – even if transparency mechanisms were initially weakly implemented, the attention on this issue would move from its adoption to its implementation and policymakers could be pressured to enforce the law (SCHNELL, 2017). Also, reversing a transparency dispositive becomes harder as it would entail reputational costs. That would be true when there is electoral competitiveness (*ibid.*). This argument is also related to the importance of Freedom of Information Laws as an instrument to institutionalize transparency and minimize the risks to revert to opacity (BERLINER, 2014), as FOI would be an institutional measure of transparency (HOLLYER et al. AL., 2014).

In some cases, the benefits of regulating access to information would overcome the political costs assumed to be associated with transparency. Michener (2015a) argues that transparency policies can be strategic as a means to monitor coalitions' allies. This would be stronger for more ideologically heterogeneous coalitions (MICHENER, 2015b). The researcher states that when it comes to the multiparty coalition, the higher the number of parties, the greater the benefits of adopting the law, as transparency dwindles the information effect, i.e. difficulties in achieving loyalty and avoiding opportunism. In that way, the challenges for coordination in coalitions with few parties are lower and the benefits of secrecy prevailed.

Conversely, as the number of parties grows, dilemmas also increase, making the benefits of transparency overcome those of secrecy. Leaders motivated by the monitoring benefits of transparency and willing to pass a strong law need both legislative control (majority in Parliament), to deal with defiance, and a good justification to persuade partisan allies of the need for transparency (MICHENER, 2015b). Monitoring it is more attractive not only because of the number of parties but also due to ideological heterogeneity. To put it differently, when the coalition is more homogenous in terms of ideology, incentives can lead to opacity to protect allies from scrutiny (MICHENER, 2015b; MICHENER *et al*., 2018).

For Berliner (2014), when the incumbent faces electoral insecurity and predicts she is going to leave office, the enactment of FOI laws becomes strategic as a way of tying the successor's hand as she transfers the costs of transparency to the next incumbent while obtaining reputational benefits for the decision. Berliner and Erlich (2015) find evidence of this argument for the diffusion of Mexican access-to-information law at the subnational level. Michener (2015a) discusses how, in Panama, the president's decision to adopt FOI was influenced, among other factors, by the possibility of tying the hand of his successor (MICHENER, 2015a).

Batista et al. (2022) investigate the mechanisms of the diffusion of AIA in Brazilian municipalities and find that reelected mayors have greater chances of regulating AIA at the local level. In this case, they aim to understand the decision to adhere to AIA regulation at the federal level, which is slightly different from the idea of FOI adoption, used to understand regulation at the national level. Based on the insurance mechanism, they argue that mayors in their second terms establish access to information legally as a way of tying their successor's hand.

While Berliner (2014) focuses on the lame-duck period, i.e. the period between the politicians' defeat and the end of their term, Batista et al. (2022) shed light on the motivations of mayors who are institutionally constrained to leave power, as in Brazil the Executive branch is bound to one reelection. In both cases, there is evidence that mayors who are about leave office may have a greater likelihood to regulate FOI laws. Michener and Nichter (2022) also find that mayors in their second term are more likely to enact AIA locally, which is aligned with Batista et al. (2022) results.

Concerning the political variables, Michener and Nichter (2022) also discuss results indicating that the greater the margin of victory of the elected mayor, the greater the city's regulation score. They also find that if the majority of the city councilors are affiliated with the mayor's party, there is a 70-point increase in the regulation score. A positive association is also found for the probability of passing local legal dispositions for access to information, as well as for the time the law has been in effect.

Batista et al. (2022) also find supporting evidence for the learning mechanism, for which the members of a given social system can learn from one another and emulate policies from members who are geographically closer to them. The author's results signal that the greater the proportion of municipalities adopting AIA in a certain state, the greater the chances of a given

municipality in that state will adhere to the law. This outcome suggests that reducing informational costs may influence *de jure* compliance.

The authors also identify that structural variables such as GDP, population size, and state capacity can influence whether and when a municipality will regulate access to information. An interesting finding they show is that external factors lessen their influence on regulation when moderated by the proportion of municipalities that adopted AIA in a state. It means to say that the reduction in the informational cost may help local officials to overcome difficulties related to structural variables, which are normally limited to incremental changes over time.

Likewise, Michener and Nichter (2022) find evidence that income per capita is positively associated with the regulation score measured by the *Controladoria Geral da União.* Population size has a positive association with the probability of regulating access to information, as well as with the time since regulation (i.e., for how long the law has been operative). Moreover, municipalities for which neighbors have been audited by CGU were also associated with greater regulation scores and a higher probability to enact an FOI law. They also had an expected increase in the duration of the law. On the other hand, taking part in a training program from the CGU had no significant association with *de jure* compliance.

2.3.2 *De facto* compliance

When it comes to passive transparency, politicians can implement FOI laws as a way to spoil opponents' advantages too. According to this reasoning, higher competition and party alternation would be positively related to higher compliance with freedom of information (BERLINER, 2014; MICHENER, 2011; MICHENER ET AL., 2018).

Berliner's (2017) analysis of the access to information law's compliance at South Africa's subnational level sheds light on other relevant factors to compliance at the local level. Considering that localities have low state capacity and low enforcement provided by national authorities, one could expect low compliance with transparency. Yet, some localities enforced access to information. Berliner (2017) argues that local political competition affects compliance through three distinct mechanisms: reelection responsiveness, intraelite monitoring, and intertemporal benefits/costs. He finds evidence of political competition's association with a higher propensity to comply with the law (BERLINER, 2017).

That is, competition increases the likelihood to comply with the law and build an institutional arrangement able to stimulate actual implementation. Results demonstrate that

greater political competitiveness increases local compliance with access to information. This evidence contrasts with the argument that greater competition would harm compliance, as resources employed to implement transparency could be applied to deliver other public goods with possible greater electoral return. Also, transparency could enable the opposition to gather negative information on the incumbent (BERLINER, 2016a). Political competition is also found to be positively associated with higher levels of fiscal transparency in the United States (ALT; LASSEN, 2006). Additionally, it is positively associated with more internet-enabled transparency at the local level in the USA (LOWATCHARIN; MENIFLELD, 2015).

In Brazil's subnational context, Michener *et al.* (2018) find that, at the local level, neither symbolic legitimacy, monitoring, nor the level of political competition has led to greater compliance with transparency. In addition, fragile oversight can worsen this situation (Michener *et al.*, 2018). Other studies, specifically in the case of Brazilian municipalities, find no significant effect of different measures of electoral competition – such as the margin of victory (BATISTA, 2017).

In a more recent study, Michener and Nichter (2022) found that only party alignment between the mayor and the city councilors had a positive association with the passive transparency score, the probability of having a fiscal system of information for the citizens, and the quality of information score. The margin of victory and whether the mayor is in her first or second term had no significant association with *de facto* compliance proxies. Finally, the audition of neighboring cities had no significant association with *de facto* compliance measures. Conversely, cities that received training from CGU had a significant expected increase in the quality of information score.

Besides political factors, Michener and Nichter (2022) examine how socioeconomic variables are associated with *de facto* transparency, measured by CGU's passive transparency scores, whether the city has a system of information in which citizens can request information and the quality of government responses to specific requests of information made by CGU. The results highlight that socioeconomic factors, such as income per capita and population size, are positively related to compliance. The former increases passive transparency, while the latter increases the passive transparency score and the quality of information score.

Batista et al. (2022) and Michener and Nichter (2022) also explore ideology, observing the mayor's party. Michener and Nichter (2022) tested four dummies: if the mayor were affiliated to DEM, PT, PMDB, and PSDB. They find no significant effect for none of the proxies used to measure *de jure* compliance. On the other hand, Batista et al. (2022) find that

right wing mayors are negatively and significantly associated with the propensity to regulate access to information.

Silva e Bruni (2019) also explore EBTs passive transparency score for 1,113 municipalities in Brazil and find that the level of education of the population and of public officials (percentage of the population who has graduated in high school or undergraduation), the income per capita, the municipal per capita revenue, the average age of the population and the age of public officials are positively related to passive transparency. Only the average age of the population, education, and municipal per capita revenue presented significant results. Both income and education followed the expected direction, but age was expected to be negatively associated with passive transparency. The reasoning presented by the researchers is that the use of passive transparency mechanisms is usually linked to the younger extracts of society. Silva and Bruni's results indicate these expectations do not find support in the empirical outcomes, neither for the population in general nor for the public officials.

Drehmer and Raupp (2020) also evaluate passive transparency for the platforms to request information for the Executive, Legislative, and Judiciary branches in Brazil. They perform a qualitative evaluation of online platforms according to three aspects: the communication on the platform (e.g.: the platform allows filling an information request, receiving responses and appealing responses), login, and receipts (e.g.: if the platform has a specific login), and barriers to access (including whether the platform requires personal identification or not). The results indicate that the three branches vary among the three aspects of information platforms. While the Executive branch had a better performance relative to communication and login and receipts, Judiciary and Legislative branches had a higher score when it came to barriers.

The following table summarizes the main explanations presented by literature regarding *de jure* and *de facto* compliance. Only income seemed to be consistently associated with both types of compliance. Political variables are mostly related to *de jure* compliance, which make sense since regulation substantively depends on political will. For *de facto* compliance, political competition is a variable that presents a significant effect for other countries' local context, such as Mexico and South Africa, but not for Brazil.

Table 1 - A summary of the evidence on de jure and de facto compliance at the local level (continues).

Object	**Variables**	**Evidence**	**Authors**
De jure compliance	Margin of Victory	significantly positive	Batista et al. (2022), Michener and Nichter (2022)
	Electoral insecurity/Second-term mayors	significantly positive	Batista et al. (2022), Berliner and Erlich (2015), Michener and Nichter (2022)
	Party identification between mayor and city councilors	significantly positive	Michener and Nichter (2022)
	Party identification between mayor and the President	significantly positive	Batista et al. (2022)
	Ideology	significantly negative (for right wing mayors)	Batista et al. (2022)
	Proportion of municipalities in a state that have regulated AIA	significantly positive	Batista et al. (2022)
	Neighboring cities audited in CGU's evaluation	significantly positive	Michener and Nichter (2022)
	GDP per capita	significantly positive	Batista et al. (2022), Michener and Nichter (2022)
	Population size	significantly positive	Batista et al. (2022), Michener and Nichter (2022)

Table 1 - A summary of the evidence on *de jure* and *de facto* compliance at the local level (ends here).

Objects	Variables	Evidence	Authors
De facto compliance	Political competition	Positive and significant relationship in South Africa and the United States. No significant evidence in Brazil.	Alt and Lassen (2006), Berliner (2017), Batista (2017), Lowatcharin; Menifleld (2015), Michener et al. (2018), Michener and Nichter (2022), Batista et al. (2022)
	Participating in CGU training	significantly positive	Michener and Nichter (2022)
	Age of population	significantly positive	Silva and Bruni (2019)
	Level of education	significantly positive	Silva and Bruni (2019)
	Municipal per capita revenue	significantly positive	Silva and Bruni (2019)
	GDP per capita	significantly positive	Michener and Nichter (2022)
	Party identification between mayor and city councilors	significantly positive	Michener and Nichter (2022)

Source: the author.

2.4 A BRIEF REFLECTION ON THE LITERATURE

The arguments discussed in this chapter demonstrate that different variables, among political and socioeconomic factors, can influence compliance with access to information at the local level. Previous work suggests that reelected mayors and party alignment, for instance, can be associated with positive changes in regulation, and active and passive transparency. However, it seems that those variables have more influence on regulation, or *de jure* compliance than on passive transparency, for example. As argued by Michener et al. (2018), neither symbolic legitimacy nor monitoring and competition has increased commitment to transparency in Brazil at the subnational level.

Evidence signals that structural variables, such as per capita income and the size of the population, are relevant when trying to understand compliance. However, as highlighted by Batista et al. (2022), those variables change at a gradual pace. As structural factors are consistently relevant for transparency, one could expect the variation in AIA's regulation and passive transparency measures, such as the Brazil Transparent Scale (EBT, in Portuguese), to be gradual, as capacity would not radically change from one year to another. However, data shows that some municipalities can substantially improve year by year and even worsen transparency scores in the same length of time.

For instance, according to the third edition of the EBT, the city of Florianópolis, which is the capital of Santa Catarina state, had a decrease of -6.67 in its score compared to the second edition of EBT's evaluation. Meanwhile, Américo dos Campos, a small city in the state of São Paulo, with 5.706 habitants (IBGE, 2020), had an increase of +8,61 in its score (CONTROLADORA GERAL DA UNIÃO, 2020a). The time difference between both evaluations is nearly one year – the second edition was performed from July to August 2015, while the third one was executed from June to October 2016 (CONTROLADORA GERAL DA UNIÃO, 2020b).

It is important to explore which other variables may lead to short-term changes in compliance. That is not to say capacity does not matter, but rather that other factors should be also considered when one tries to understand the variation in compliance among the municipalities. Hence, it is important to shed light on issues such as intentional interference with information and political variables and the administrative dilemmas it can generate, as underlined by Rodrigues (2019). Similarly, one of the difficulties of AIA's implementation highlighted by public servants at the federal and municipal levels is the low interest and lack of commitment of public managers in promoting transparency (BATISTA, 2017). This signals that political will may be a central aspect of promoting access to information by municipal governments, and so are the political incentives affecting decision-making.

While the debate has already discussed important political variables such as electoral competition, electoral insecurity, party alignment, among others, there is an important political factor that is yet to be explored by this literature: the presence of dynastic mayors. Recent studies have flourished in political science and political economy exploring the effects of political dynasties in a myriad of variables (GEYS; SMITH, 2017).

Among them, there is evidence that dynastic politicians can be more inclined to engage in clientelism and corruption (MENDONZA *et al*., 2012; OLIVEIRA AND SOUZA, 2022;

SETYANINGRUM; SARAGIH, 2019). There is also evidence that political dynasties can lead to worse quality of public spending, even when dynastic politicians raise resource distribution (BRAGANÇA *et al*., 2015; DANIELE; GEYS, 2015) and have null effects on substantive representation (LABONNE *et al*., 2019). In this sense, mayors that are more inclined to wrongdoing may as well be less willing to comply with transparency.

On the other hand, positive outcomes related to the dynastic leaders suggest that these politicians may be more inclined to implement long-term policies, as they value building reputation and legacy not only for their careers, but also for their family members that are in politics (or might enter it in the future) (TUSALEM; PE-AGUIRRE, 2013). Access to information is a long-term policy in the sense that its endurance requires constant commitment to work, and its results depend on its proper functioning to appear in the long run.

Families have had a central role in local level politics since Brazil's first republic (ARRUDA; KERBAUY, 2016). Even though nowadays the presence of political dynasties is not necessarily related to practices such as clientelism, patronage and vote-buying (OLIVEIRA et al., 2017), these phenomena is still present in local politics, but within a different dynamic (BORGES, 2018). Political dynasties are still a relevant part of Brazil's politics, at the national and local level (BOAS *et al.*, 2018; OLIVEIRA, 2012).

With that in mind, I suggest exploring the relationship between political dynasties and the compliance with access to information in Brazilian municipalities. I believe this investigation may contribute with novel evidence of a new political factor that seems to be relevant in local politics and can help to deepen our knowledge on the political incentives that decision-makers may have to comply with access to information.

3 POLITICAL DYNASTIES AND THE COMPLIANCE WITH ACCESS TO INFORMATION

To the best of my knowledge, the relationship between political dynasties and public transparency is yet to be explored. For that reason, it is not possible to build an argument based on what previous studies have found when it comes to this specific relationship I aim to investigate. Rather, I intend to present a broad approach on the discussion on political dynasties in general, and, more specifically, in Brazil, and connect it with the debate on compliance with AIA, discussed in the previous chapter.

To do so, I first present a general view on the definition of political dynasties. Then, I discuss the agenda of the studies about political dynasties and the main issues they are focusing on. Based on that, I center the debate on the effects of political dynasties on different variables, including factors that are usually associated with transparency, such as corruption. After reflecting more broadly on political dynasties, I focus on the presence of political dynasties in Brazil. I discuss canonical studies such as Victor Nunes Leal's (2012 [1948]), up to more recent investigations. Finally, I try to connect the contributions of previous research to argue how political dynasties are relevant to the study of the local level compliance with access to information in Brazil.

3.1 DEFINING POLITICAL DYNASTIES

There are many ways to define political dynasties. Either in a more general definition, such as the transmission of power to relatives (ROSSI, 2017), or in a more restrictive and detailed definition, such as the transmission of power by politicians belonging to a small group of families who benefit from electoral advantages (QUERUBIN, 2016), or who monopolize power over time (QUERUBIN, 2011). This classification can expand from the individual level, i.e., dynastic politicians, to the partisan level, i.e., dynastic parties. These are parties in which leadership comes from a family in power (CHHIBBER, 2013; AMUNDSEN, 2016).

The definitions may also be either more restrictive and limit dynasty to successions within the same jurisdiction (e.g., the same city) or broader and include the possibility of not immediate succession at different levels of government (e.g., mayor and governor). Rulers who initiated the political dynasty created advantages for their successors but did not benefit from dynastic advantages in their first term. Thus, in general, they are not included in the definition of a political dynasty (GEYS; SMITH, 2017).

Other studies define dynastic politicians solely as those whose relatives were previously elected. Kinship, in this case, includes blood and marital relations (CLUBOK *et al.* 1969; DAL BÓ *et al.*, 2009; FEINSTEIN, 2010; ASAKO *et al.* 2015; BRAGANÇA *et al.* 2015; FIVA; SMITH, 2018; FOLKE *et al*., 2016; GEYS, 2017; SMITH; MARTIN, 2017). Ishibashi and Reed (1992) describe dynastic politicians as those who inherited the power of relatives.

This latest definition has the advantage of being straightforward and precise, facilitating the operationalization of the concept. As a drawback, this thin concept of political dynasties may be too broad. It can encompass families in which several members participated in politics throughout the years and families in which only two individuals took office a few times. In this case, both of them can be defined as political dynasties, but they can have distinct levels of concentration of power and political influence.

In this dissertation, as the focus relies on the variation of compliance with access to information in recent years, the goal is to observe if and how ruling political dynasties are associated with differences in compliance across municipalities. Unlike studies that focus on a historical approach and intend to track families who concentrated power in certain cities in the past (e.g., FERRAZ ET AL., 2020), the aim in this research is to examine dynastic politicians who were in office during the evaluations of compliance with access to information.

While families continue to enter politics in Brazil at the local and national levels, the social and political landscape in which they exist has significantly changed. When trying to define what are political dynasties for research centered in the Brazilian case, it is crucial to ponder what families were and what they might be now in local politics.

In Brazil's first republic, families were a central part of the formation of the state. Through the *coronelismo* system, and even after that, during the dictatorship, families greatly influenced the public sphere (HAGOPIAN, 1986; OLIVEIRA *et al*., 2017). Families were at the core of *coronelismo* (ARRUDA; KERBAUY, 2016) and enjoyed power concentration and different political influence over economically dependent individuals in rural areas. Currently, even in small municipalities in Brazil's countryside, political family groups exist in a considerably more modernized context, competing in regular elections under the rules of Brazilian electoral justice.

In this sense, the jurisprudence of the Brazilian electoral justice represents how the set of legal provisions related to elections have been interpreted and how cases have been judged. It sets the rules under which candidates, including the dynastic ones, can compete. The ineligibility law (*Lei de Inelegibilidades*) (BRAZIL, 1990) states in what conditions candidates

can lose eligibility. One of them is related to what they call *inelegibilidade reflexa* (reflex ineligibility). Article 14, 7th paragraph of the Federal Constitution defines that relatives - by marriage and by blood - of politicians who are in office in the Executive branch (i.e., mayor, governors, and president) cannot run for elections in the same jurisdiction. Candidates who have been elected before or are running for reelection are not included in this rule.

Another concept, the itinerant mayor (*prefeitos itinerantes*), aims to prevent mayors from running indefinitely for elections by alternating the jurisdictions and, in that way, remaining in power. Regarding political dynasties, what is particularly interesting in this case is that the jurisprudence of this legal dispositive also takes politicians' relatives into account. In 2019, for a case that happened in the city of Barra de Santo Antônio (AL), Judge Luis Roberto Barroso explained that politicians' relatives are free to run for elections in other municipalities, unless the constituency is a result of a dismemberment or merging from another jurisdiction (TSE, 2019).

Taken together, both concepts suggest that having two members of the same family running for elections, in some situations, can be already considered a political dynasty. Excerpts from decisions made by the Superior Electoral Court and other instances of the Brazilian Electoral Justice reinforce this idea. One example is the case of Nivaldo Santos and his wife, Rosiane Santos, in São Miguel dos Campos (AL). Justice understood that Rosiane's election would represent a perpetuation of the same family in power, since Rosiane and Nivaldo had a common-law marriage, and both have been elected two times as mayors of the city. A judge justified the decision stating that: "there was the intention to start a family, the intention to deceive the Justice, *the decision to keep a dynasty in power*. The evidence collected proves what was happening in São Miguel was a huge fraud" (JORNAL EXTRA, 2012, free translation, emphasis added).

In this sense, Brazil's electoral justice understanding of what represents a political dynasty seems closer to a minimal definition of it. As the focus in this research lies in the modern aspects of political dynasties, that is, what it means to have a dynastic mayor in office in today's time, I follow the simple and most common definition of political dynasties used in studies on this theme around the world. *I define dynastic politicians as those whose at least one relative has been previously elected, in consecutive elections or not, and regardless of being in the same jurisdiction or not*.

3.2 STUDIES ON POLITICAL DYNASTIES

The main arguments of the theoretical debate about political dynasties explore this phenomenon either as a dependent variable or as an independent one. In the first case, a central issue researchers examine is self-perpetuation. That is, what variables facilitate the continuation of dynastic politicians in power. Different studies present evidence that the longer the time a politician stays in government, the greater the likelihood of self-perpetuation. Dal Bó et al. (2009) were the first to empirically identify the causality between term length and dynasty persistence in the United States. Abramson and Rivera (2016) find that this relationship also holds in Europe. Likewise, dynastic politicians whose predecessors were longer in power were less likely to experience depositions and parliamentary constraints.

Conversely, Van Coppenolle (2017) finds no effect of serving longer on the likelihood of establishing or maintaining a political dynasty. However, other variables that are found to raise the probability of having future relatives in the office are incumbency (Querubin 2016) and winning the close election for mayor (Bragança *et al*., 2015), in the Philippines and Brazil, respectively. Similarly, Querubin (2011) finds no evidence of a statistically significant impact of term limits on the decline of families' persistence in power. This means that even with a reform that limits the time a politician can stay in power in subsequent terms, they are still able to raise the chances of self-perpetuation. That is, the chances of having a relative elected.

3.2.1 Political dynasties as independent variables

Other studies have focused on the possible consequences of political dynasties (GEYS; SMITH, 2017), i.e., dynasties as independent variables. The evidence they present is mixed in the sense that some results suggest a positive effect of political dynasties, while others indicate negative association with dynastic politicians. The goal here is to present the many areas in which political dynasties were found to be significant independent variables. To do so, I present each area separately.

3.2.1.1 Policy continuity

As Tusalem and Pe-Aguirre (2013) explain, some scholars argue that the perpetuation of families in power would lead to policy continuity. In this logic voters would have expectations for the effective provision of public goods, leading to incentives for accountable

and responsible leadership. In addition, politicians who do not come from political dynasties could have less incentives to promote public goods and carry out progressive reforms because they would not have to maintain the family legacy or the goal of perpetuating themselves in power. In summary, this argument highlights the incentives dynastic politicians can have to preserve legacy and reputation.

3.2.1.2 Political representation

In terms of representation, Crowley and Reece (2013) find evidence that political dynasties increase incumbent accountability. This is because, according to the authors, families avoid opportunistic behavior during their term so as not to undermine prospective family incumbents in the future. On gender representation, the dynamics of political dynasties can help women to overcome a voting disadvantage in elections, as the quality of predecessors may be used to recruit female dynastic successors. It is important to highlight that dynastic women have higher observable qualifications than dynastic men (Folke *et al.*, 2016). This is not necessarily translated into substantive feminine representation. Labonne *et al*. (2017) find that dynastic female mayors in the Philippines, in comparison to their male counterparts, have no differential effect on policy, economic or electoral outcomes during their first term.

3.2.1.3 Party System and Elections

In a broader sense, Amundsen's (2016) results indicate that dynastic parties led to a very restricted degree of internal democracy in Bangladesh's political parties. Additionally, Chhibber (2013) verifies that dynastic parties are related to greater volatility in party systems. Moreover, dynastic parties are associated with a less representative political system as voters are less likely to see dynastic parties as representing their interests. What's more, researchers identify differences in background and effort between dynastic and non-dynastic politicians. In that way, a political selection process led by politicians and not by the electorate produces dynastic candidates with relatively lower levels of education (GEYS, 2017). There is also a negative correlation between being a political dynasty and political effort (ROSSI, 2017).

3.2.1.4 Political competition

Political competition can be affected by political dynasties as well. As we already saw, there is plenty of evidence of families' endurance in power. This generates a disincentive for challengers. Querubin (2011) identifies that the introduction of term limits besides not having an effect on self-perpetuation can also stimulate challengers to prefer to wait for the dynastic politician to run out of terms to compete. That means that dynastic politicians will take advantage of an even less competitive environment in their initial terms.

Moreover, Fiva and Smith's (2018) outcomes point to incumbent advantage in party-centered and closed proportional representation systems, as well as Smith and Martin (2017) find that dynastic politicians benefit from an advantage in cabinet selection. Still, this advantage cannot be attributed simply to greater electoral popularity. Likewise, dynastic candidates can benefit from advantages in campaign funding. As they have among their greatest donors' family and friends, this closer relationship increases the candidate's credibility vis-à-vis those who finance their campaign (SAMUELS, 2001). Dynastic politicians have more access to resources and to a wider political network (GEYS; SMITH, 2017). Nonetheless, the literature finds a null effect of "brand name advantages" (FEINSTEIN, 2010) and electoral advantage for second-generation candidates (ISHIBASHI; REED, 1992).

Still regarding electoral advantages of dynastic politicians, Oliveira et al. (2017) affirm that traditional families can affect political competition via candidate selection (OLIVEIRA *et al.*, 2017). In this realm, Besley (2005) discusses four types of political selection: i) by lottery, as occurred in Athens; ii) through hereditary aristocracies, which are based on the transfer of power through kinship; iii) power by force, related to autocratic governments; iv) and elections. Ideally, elections have the role of disciplining politicians and ensuring that they act in the public interest (thereby reducing incentives for self-interest) and, of course, selecting representatives who are capable and committed to fulfilling this mission (BESLEY, 2005).

Classic models of electoral control argue voters select politicians based on their performance, and not on their platforms, as voters cannot be certain that politicians will not deviate from what they have promised, once they get elected (BARRO, 1973; FEREJOHN, 1986). In an electoral model of political dynasties, the evaluation of dynastic politicians would be based on the performance of their ancestors. That is, dynastic status affects voters' beliefs regarding the candidates' type (ZHENG *et al.*, 2017).

3.2.1.5 Quality of government

The studies about the association between political dynasties and quality of government include phenomena such as public spending, accountability, corruption, clientelism, and so on. Regarding public spending, there is evidence of a positive effect of the dynasties on public spending and the distribution of benefits. In that sense, dynastic politicians are found to be associated with higher economic growth, but only when executive constraints are weak (BESLEY; REYNAL-QUEROL, 2017). They are also related to more distributions to the district, higher electoral success, and a bigger number of distributive benefits. However**,** their spending can harm the economic performance of districts (ASAKO *et al.*, 2015).

This is similar to the findings of Bragança *et al.* (2015) in Brazil, where dynastic leaders spend more resources, especially on urban infrastructure, health, and sanitation, but do not generate improvements in economic growth and changes in the quality of public services. Furthermore, Daniele and Geys (2015) find that dynastic mayors spend more and obtain higher transfers in the year preceding elections. However, there is no effect of dynastic mayors on average spending, revenues, and transfers.

So far, the evidence of a positive relationship between political dynasties and the quality of government is not definitive. As discussed, dynastic politicians raise and spend more money, but in general, do not increase the quality of government. The domination of families established in power has, therefore, a negative effect on the allocation of public goods, even if their presence leads to greater representation in the legislature (TUSALEM; PE-AGUIRRE, 2013). Similarly, the presence of dynasties can increase the descriptive representation of women in the legislature, without this leading to a substantive feminine representation (LABONNE *et al.*, 2019).

The assumption that political dynasties negatively affect the quality of government is a close link to the argument, thoroughly explored in political economy, that the concentration of power favors the establishment of extractive institutions and clientelism (FERRAZ *et al.*, 2020). The classic debate on elites postulates that elite-dominated political systems favor the perpetuation of ruling families, which should not face much accountability coming from the disorganized masses (TUSALEM; PE-AGUIRRE, 2013). Acemoglu and Robinson (2008) find that the domination and persistence of elites in power lead to the weakening of political and economic institutions. Tusalem and Pe-Aguirre (2013) add to evidence indicating that systems

dominated by elites must lead to laws that favor them, such as the laws of rural workers in Latin America and Central America.

Aligned with these results, some scholars argue that the domination of power of dynastic politicians harms electoral accountability and then precludes the rise of high-quality, reform-oriented candidates. The underlying logic in this argument is that political dynasties tend to favor patronage and clientelism and are focused on maintaining the status quo. In this context, voters' evaluation is based not on performance, but on client-patron relations and the candidate's family ties. So, incumbents have less incentive to implement long-term infrastructure projects and policy reforms that will deliver incremental benefits rather than immediate ones. In this scenario, it is possible that politicians would prefer to intensify the delivery of public goods immediately before elections occur instead of providing public goods throughout their term in office (TUSALEM; PE-AGUIRRE, 2013).

In the Philippines, one of the countries with the biggest prevalence of political dynasties, Mendonza *et al.* (2012) identifies that the client-patron relation developed by political dynasties in Congress is related to corruption. Moreover, the election of inexperienced politicians, respect for the power accumulated by families, and the greatest difficulty citizens can face to present their demands to the government are some of the effects of political dynasties on the quality of government (TUSALEM; PE-AGUIRRE, 2013). Setyaningrum and Saragih (2019) find that political dynasties in Indonesian local governments have a negative effect on government performance. However, this impact is moderated by the quality of local governance. In that way, good public governance may be a way to overcome the negative effects of having dynastic politicians in power.

Coupled with this, the mechanisms of exchange and monitoring linked to patronage and clientelism relations guarantee the perpetuation of families in power, and family recognition overcomes ideology, preventing dynastic candidates from facing stronger setbacks due to political scandals. In this context, political parties are weakened by the domination of families in power. They fail to represent cleavages, issues, or ideology and have merely the function of formalizing the candidacy of dynastic individuals (TUSALEM; PE-AGUIRRE, 2013).

Boas *et al*. (2018) identify the presence of political dynasties in the Brazilian municipalities and how voters show loyalty to these families. Attitudes toward dynasties would often be tantamount to strong partisan identification. This mechanism would affect the role of access to information in the voting decision. Thus, the authors verify that there is no normative indifference of the voter regarding poor performance, but rather that other factors such as

employment and income concerns and loyalty to political dynasties are more important in their decision.

3.3 POLITICAL DYNASTIES IN BRAZIL

"I quit and leave Junior here. *It would be the same as if I were occupying the seat*" (FOLHA DE SP, 2001, emphasis added). This phrase echoed in the backs of the crisis that reached the then Senator Antonio Carlos Magalhães (ACM) in 2001, after the disclosure of what became known in Brazil as the "electronic panel fraud". The senator referred to his son, ACM Filho, who was his substitute in the Senate. The Magalhães, led by ACM, is one of the largest political dynasties in Brazil and dominates the political scene in the state of Bahia (just as the Collor de Mello do in Alagoas and the Sarney do in Maranhão, among others).

The scandal that led to the fall of ACM and made him stay an unusual period out of power (2001-2003) revealed the violation of the Senate electronic panel in the secret voting for Senator Luiz Estevao's cassation. ACM would have had access to the votes of his colleagues through a violation commanded by the former governor of the Federal District, José Roberto Arruda, who was convicted in 2016 (REVISTA ISTO É, 2016). On his resignation, ACM stated: "although without a mandate, I feel freer than ever to work for what interests me: Brazil and Bahia. Do not think that you are deciding my destiny. *Who decides my destiny is Bahia*" (PORTAL TERRA, free translation, emphasis added). Bahia elected him senator again in 2002 with 2.9 million votes.

The existence of political dynasties in Brazil is widely known (SAMUELS, 2001). Historical and sociological studies show the family's central role in forming the Brazilian social structure and its influence on politics. The family assumes the role of organizing the country, attenuating tensions and contradictions between the groups that initiated the Brazilian colonial society. With the rise of the modern state in the country, there would be tension between the traditional and personalist society and the state sphere. Some argue the result of this process was the prevalence of private interests over public interests and the victory of the patrimonial state (OLIVEIRA *et al*., 2017).

Oliveira Viana (1999), in describing the political institutions in Brazil in the first republic, highlights the absence of common services. In Maranhão, for instance, he says: "(...) there is no butchery, no riverside, no garden, no tent ..." (p. 131). Anyone who needed any service would have to look for it in the private sphere; that is, to move around the city, they would need their own means of transport; to eat, their own vegetable garden, and so on. And

so, Viana concludes: “(...) each family becomes a republic” (p. 132). The mills represented self-sufficient “agrarian autarchies”, representing the family and patriarchal individualism pointed out by Oliveira Viana.

The absence of a collective spirit associated with these “feudal clans” is one of the author's critiques of democracy in Brazil at that time, specifically to the federalization that, for him, was an obstacle to the formation of public interest (OLIVEIRA *et al.* 2017). There is consequently an amalgamation of the public and the private in the face of the central role of patriarchal families and their influence on the public sphere. According to Willens (1953), the interests of these families controlled the Brazilian system through practices such as "familism" and nepotism once they ascended to power.

Arruda and Kerbauy (2016) also call attention to the fact that *coronelismo,* a historical phenomenon that marked the political system of Brazil’s first republic, was organized as a group of core influential families. One of the most canonical writings on this topic is Victor Nunes Leal’s *Coronelismo, Enxada e Voto* (1948). His object of investigation in this book is the municipality and how representation took place at this level. Leal approaches *Coronelismo* as a phenomenon that involves all levels of the state, i.e., municipalities, state, and central government, as well as *coronéis* at the local level, governors, and the president. As explained by José Murilo de Carvalho in the preface to the book's seventh edition, this is one of Leal’s innovations. In this sense, Leal focuses on the tension between private and public realms associated with *coronéis* and governors. The central issue in his view was not specifically the coronéis but the system they were part of.

In 1831 the National Guard was created to take the place of colonial militias and *ordenanças*. The rank of colonel was granted to those individuals with great local influence, usually owners of large land properties. *Coronelismo* is how landowners, whose private power was decaying, could adapt to representation in the public sphere. On the other hand, the state could gain more control over the fiscal-dependent municipalities (LEAL, 2012 [1948]). The *coronel* gained local control over his municipality, and, in exchange, he should grant the electoral support of his dependents (family and individuals economically dependent who “followed” the local *coronel*) to the politicians at the regional level (i.e., the *oficialismo estadual*) (LEAL, 2012 [1948]; HAGOPIAN, 1986).

In this context, local administration was weak. The public sphere was fragile, facilitating that *coronéis* become responsible for delivering social goods. Because of that, practices such as paternalism or *filhotismo* were common, i.e., favoring allies through little benefits up to

offering jobs in public administration for them. On the other side, political persecution of the opponents, the so-called *mandonismo,* was also prevalent and notably increased in times preceding elections.

Leal observes in his studies that local leaders were not necessarily *coronéis,* that is, the image of a landowner. With the diffusion of higher education, doctors and lawyers were also prestigious enough to exert the role of local leaders. Yet, Leal underlies that these new types of leaders were either relatives or political allies of the *coronéis*. The agrarian context in this situation is crucial to understanding *coronelismo*. This system is based precisely on the concentration of land, limited access to education, and other social factors. Hence, Leal argued that modernization in Brazil depended on social changes such as urbanization, education improvement, and amplification of the labor market.

At this moment, the majority of the electorate was rural and had little or no economic autonomy, depending on landowners. Access to education was very limited, and individuals had no incentives to vote. Leal explains that it is no surprise the *voto de cabresto* was a common practice. Landowners and local leaders were the ones that paid for the costs of the election, including voter's electoral registration, transportation, meals, and even election day clothing. These costs were raised when changes in the electoral process required voters' presence to cast their ballots.

Yet, in the 1945 and 1947 elections, changes in this dynamic were already observed. Leal explains that remarks about how some rural workers betrayed landowners and voted against their orientation circulated. Radio was becoming more common, expanding the reach of radio advertising. With migration, some rural workers also started to move to urban areas. These changes gradually contributed to dwindling economic dependence.

Leal's analysis of local politics also challenges the idea that local political leaders lacked public spirit. The author argues that these leaders often invested in improvements in their districts or municipalities. They used influence with the public sphere to demand public services related to water access, schools, health, roads, and so on. Other than their access to public power, sometimes they resorted to personal resources and prestige. As *chefes municipais* (local leaders), these contributions were important to secure and reinforce their local leadership position.

Leal states that the literature at the time usually described *coronéis,* i.e., the local leaders, as plain self-interested leaders with no political ideals. However, he ponders that many politicians that acted at the state and national levels started in municipal politics. Therefore,

what influenced representation at the local level to be less republican or democratic was not the personality traits of the local leaders but rather the social and economic structure in force in the municipalities at the time. Again, Leal turns the focus to *coronelismo* as a system.

As Hagopian (1986) states, *coronelismo* varied in Brazil. In some states, *coronéis* were so influent that they did not depend on regional oligarchies to remain in power. This way, the arrangements between local leaders and governors took distinct shapes depending on regional variation around the country. Remarkably, *coronelismo* was not only present in Brazil's northeast but also in the center-south of the country, where state oligarchies were more powerful. For instance, São Paulo's *coronelismo* inspired Victor Nunes Leal's research (HAGOPIAN, 1986).

According to the classic argument about political oligarchies, traditional elites would be weakened in the face of the state's modernization process (HAGOPIAN, 1986). Indeed, the military dictatorship weakened traditional elites' power since they invested in raising political competition in cities with a higher concentration of power. Ferraz, Finan, and Martinez-Bravo (2020) find evidence that the military's strategy represented a shock in political dynasties' power, which led to greater economic development in those places in the long run. Nevertheless, in Brazil and other Latin American countries, such dynastic elites remained in power despite industrialization, most intensely developed in Brazil during the military dictatorship (HAGOPIAN, 1986).

Brazilian sociological studies attribute this to the ability of these political groups to change and create strategies to remain in power (OLIVEIRA *et al*., 2017). In that way, during the dictatorship, modernization did not transform Brazil into a liberal-capitalist society but into state capitalism, which eventually enabled the persistence of these elites in Brazilian local politics. The economic relationships previously restricted to the private realm were then inserted into the public and political sphere. As a result, the economy was politicized, and the nature of politics changed. The state becomes the sole power source, whereas, in liberal capitalist societies, power emanates from places other than the state (HAGOPIAN, 1986).

While Hagopian's (1985) research focuses on Brazil in its military regime, the most recent studies on family and their role in politics post-re-democratization also argue that the social changes that occurred in Brazil did not lead to a decrease in the presence of families in the political field. In 2017, according to Congresso em Foco's data, 62% of House Parliamentarians and 73% of Senate representatives had some kinship with other politicians (OLIVEIRA *et al*., 2017).

Oliveira (2012) states that families still matter in politics in Brazil. It addresses not only the control of families over political parties, executive and legislative branches, but also other spheres of power in the society, such as the Judiciary, with emphasis on the State Court of Audit (TCE), which he calls the "court of relatives", and the media. The media, however, despite having been dominated by families for a long time, has been undergoing a process of modernization and professionalization (OLIVEIRA, 2012). Thus, the presence of families in the public sphere, in general, persisted despite social changes, as the forms of insertion of these families also changed.

In this sense, Boas *et al*. (2018) discuss how mass partisanship in local politics is considered weak, while voter attitudes toward local political dynasties can be understood as strong partisan identification. Some dynastic candidates compete consistently for the same family, but others constantly change parties (BOAS *et al.*, 2018). In the words of Boas *et al*. (2018): "(...) as party affiliation, a candidate's membership in a local political family might serve as a cue of competence or policy positions or as a proxy for clientelistic networks" (p.12).

This party identification may limit, according to the authors, the ability of information to change voting behavior, as far as can lead to motivated reasoning. In other words, voters may interpret allegations of corruption against a family group as invented by another family that opposes that clan. Thus, the effect of information on voting behavior should be greater in places not dominated by family dynasties.

However, the presence of dynasties is no longer necessarily linked to an oligarchic history in which there is state inequality and inefficiency (OLIVEIRA *et al*., 2017). Although features such as patronage and clientelism are still present in Brazilian politics, the dynamics that involve those practices are not the same. While the scenario described by Victor Nunes Leal pointed to individuals who had no incentives to vote and deeply depended on local leaders, vote buying in current times seems to respond to a different logic, as shown by Borges (2018).

Before discussing Borges' findings for clientelism in Brazil, it is interesting to make a brief digression on this topic. Kitschelt (2000) sheds light on the different mechanisms of party linkage between voters and elites, which go beyond the programmatic linkage based on the distribution of costs and benefits in an equitable manner independently of the individual's vote. The first is the charismatic linkage. It refers to when citizens base their judgment on whether or not they like the candidates' gestures and styles. There is also the clientelistic linkage, when the relationship between voters and elites is based on the delivery of specific material advantages by politicians in exchange for electoral support.

These alternative forms of linkages are born of the need to facilitate collective action, which refers to resource challenges faced by candidates and voter's information asymmetry problems. Investing in the organizational-administrative infrastructure of the parties would facilitate collective action. When politicians do not invest in organizational infrastructure, they rely on their charismatic authority. When, in turn, representatives invest in technical administrative infrastructure, but not in forms of aggregation of interest and program formation, they create direct, personal, and material links with voters (KITSCHELT, 2000).

This connection takes place from two exchange circuits. The first involves politicians with money in exchange for material favors, leading to rent-seeking and market distortions. This happens when there are a lot of resources and few votes. On the other hand, when there are a lot of votes but few resources, voters receive selective material incentives before and after the election in exchange for voting (*ibidem*).

Thus, clientelism involves both reciprocity and voluntarism, and exploitation and domination. The exchanges, although mutual, are asymmetric (KITSCHELT, 2000). In this way, Kitschelt shows how the representation logic existing in clientelism is not based on performance evaluation, as it happens in the programmatic linkage, but in particularistic exchanges. The voter would have no interest in pushing the incumbent for better policies because he feels represented by receiving individual benefits rather than public goods that meet their needs.

Gottlieb and Kosec (2019) argue that political competition can be understood as a mediator for the provision of public goods insofar as it affects the electoral incentives of incumbents seeking reelection and reduces the efficiency of legislative bargaining. Evidence, however, does not always corroborate this relationship. The authors find robust evidence that political competition negatively affects the provision of public goods in the presence of weak political parties and low transparency. According to the authors, more competitive or fractionalized parties can lead to coalition formation, making legislative bargaining more difficult and thus improving the efficiency of policymaking.

However, if bargaining becomes harder in environments with poor accountability, the relationship between political competition and the provision of goods will be negative (GOTTLIEB; KOSEC, 2019). Melo *et al*. (2009), when analyzing the performance of audit agencies in Brazil, conclude: "(...) Political competition is only governance enhancing when it is programmatically oriented. Otherwise, it only creates electoral volatility based on personalistic ties" (p.23).

Through an ethnography in the *sertão* of Bahia, Borges (2018) presents a new approach to vote buying from the perspective of clientelism. The author identifies that politicians buy votes to influence voters' perception of the viability of their candidacy. This conclusion diverges from the literature's expectations, for which candidates' direct resources to voters with a high chance of converting the benefit received into votes.

According to Borges (2018), individuals vote for candidates who have political strength i.e., those they perceive to have real chances of winning. Vote buying is one way to demonstrate this strength. Thus, candidates distribute resources indiscriminately and link the efficiency of vote buying to the impression they can make on voters. Some candidates have greater socialization in the clientelistic practice and find it easier to buy votes in a way that generates the desired impression on the electorate. For other candidates, this practice is less natural. Borges' findings reveal an electorate much less indifferent to elections than rural voters in Brazil's first republic.

Additionally, Oliveira and Souza (2022) investigate whether dynastic mayors are more likely to have irregularities detected in CGU's national oversight program. They find that dynastic politicians have no differences regarding corruption compared to their non-dynastic counterparts. Yet, when authors analyze the irregularities separately, they find that dynastic mayors are more likely to execute over-invoicing. Considering the overall irregularities as proxies of corruption and mismanagement (BATISTA *et al*., 2021), Oliveira and Souza's results indicate that dynastic politicians are not worse managers than non-dynastic mayors.

Altogether, Borges' (2018) and Oliveira and Souza's (2022) evidence suggest that even in contexts that involve corruption, mismanagement and vote-buying, the contemporary political landscape in Brazil diverges from the picture from decades ago. Political dynasties are not necessarily more related to irregularities, and voters are more politically sophisticated when compared to the electorate in the 1940's, which had barely started to exert its right to vote in an electoral system that was not as institutionalized as the current one.

3.4 HOW CAN POLITICAL DYNASTIES HELP UNDERSTAND VARIATION IN COMPLIANCE WITH AIA IN BRAZIL?

The previous sections in this chapter brought two main discussions. First, what are the main arguments related to political dynasties, with an emphasis on what are the effects of political dynasties. Second, the debate focused on the role of family groups from its influence in the formation of the Brazilian state to the persistent presence of political dynasties in

contemporary politics. Now, my goal is to discuss what we can learn from these studies in terms of what to expect from dynastic politicians when it comes to compliance with access to information. In other words, how can political dynasties help to answer the question of why some local governments comply with transparency, while others do not.

The theoretical argument of this dissertation is focused on dynastic politicians, i.e., those whose at least one relative has been previously elected, in consecutive elections or not, and regardless of being in the same jurisdiction or not. In this chapter, I first discuss why to expect a distinct behavior from dynastic politicians in comparison to their non-dynastic counterparts, and why this matters for compliance with AIA. Then, I argue that the association between dynastic politicians and the local level of compliance with AIA depends on the level of consolidation of political dynasties. I expect that cities in which the ruling political dynasty is well established in the electoral realm and is inserted in the formal spheres of politics, having close connections with state and national level politics, for instance, will have a less intense negative association with compliance. Finally, I present the empirical implications of this argument, that is, the main hypothesis of this research.

3.4.1 The importance of examining political dynasties

Studies on political dynasties demonstrate their positive and negative effects in a myriad of areas. Among them, the quality of government can either be positively or negatively affected by dynastic politicians. The arguments regarding a positive association of political dynasties and policymaking account for the incentives that dynastic politicians have to deliver long-term policy outcomes. Because the members of a dynasty are worried not only with their electoral prospects, but also with their relatives' as well, they may improve policy making to build up their legacy and protect their reputation.

Concerning compliance with access to information, Schnell (2017) and Berliner (2014) explain that adopting freedom of information legislation can be advantageous for politicians as a way of signaling virtue, honesty. It can be read as commitment to fight corruption and so on. However, compliance, as seen in previous chapter, involves more than the regulation of access to information. It encompasses active and passive transparency. As schnell explains, only adopting FOI laws can be a window dressing, if politicians expect to enact the law but not comply with it.

Yet, access to information is a long-term policy. It demands ongoing effort to comply with the law. Its results, aside its regulation, that can be a clear signal to the electorate, are marginal. Mayors can try to capitalize with evaluations of transparency results, for instance, but it is something intermittent. Hence, reputational incentives for non-dynastic politicians regarding outcomes that are marginal and in the long term can be low. For dynastic politicians, incentives may be more salient, as they have greater inclination to consider not only their political horizon, but also their family members' as well.

On the other hand, negative effects are usually related to the argument of concentration of power and how it can harm the quality of government and development. Political dynasties are then understood in some studies as a proxy for this concentration. If a political dynasty is dominant in a certain place, it would have incentives to rent seeking, implementing extracting fundings, and building relationships with the electorate based on patronage and clientelism. In return, voters would evaluate not the performance of dynastic politicians, but the individual benefits they have received (clientelistic linkage) or personal features of dynastic leaders (charismatic linkage).

Reflecting these arguments on the compliance with access to information, if dynastic politicians are more inclined to rent seeking, clientelism and patronage, they will likely have less incentives to comply with the law. As dynastic politicians are able to accumulate power and influence, they can also use their political resources to avoid sanctions for no compliance. Having in mind that sanctions for no compliance with AIA are already low, it is reasonable that dynastic politicians would have even less incentives than non-dynastic politicians to comply.

I argue that the kind of relationship one can expect to find between political dynasties and the compliance with access to information depends on the political context in which the dynastic politicians are inserted. Three variables are important in this case: the level of political competition in the municipality, and the relationship between local dynasties and state level and national level politicians in the Executive.

3.4.2 Political dynasties and the compliance with AIA: a municipal-level perspective

The presence of political dynasties in Brazilian politics has been marked by the tension between the public and private spheres. Since the first republic, municipal leaders' political influence depended on formal spheres of power. In Brazil, there is evidence that citizens understand family groups that take part in politics as political groups (BOAS *et al*., 2018). I

argue that political dynasties represent, in some measure, an informal sphere of exercise of power. If dynasties manage to influence the political selection process, for example, they exert influence through a formal mechanism represented by the selection of candidates made by political parties. I argue that to understand the effect of political dynasties is crucial to consider how their influence is limited to the formal spheres of power, such as elections.

Furthermore, bearing in mind the negative effects of political dynasties on variables associated with the quality of government, I argue that, in general, dynastic politicians will be less inclined to comply with the AIA. They will be more prone to wrongdoing and, therefore, less willing to legally and practically comply with the prerogatives of access to information. With that in mind, I propose the following hypothesis (valid for *de jure* and de *facto* compliance):

H1: Dynastic mayors will be negatively associated with local compliance with AIA.

Nevertheless, I understand that this association between dynastic mayors and compliance with access to information is circumscribed in a formal instance of power. Therefore, it is important to remember that depending on these formal incentives, the negative association between dynastic mayors and compliance can be more or less intense.

I propose that this variation is linked to what I call the consolidation of dynasties at the local level. I am not referring to political dominance but to the extent to which political dynasties, as informal instances of power, are embedded in formal spheres of power and are politically well stablished. I then argue that less stable dynasties will tend to rent-seeking behavior, thus having a greater negative association with access to information. On the other hand, more consolidated dynasties will tend to act aiming at long-term returns. Therefore, the more consolidated the dynasty, the lower the negative association with the commitment to transparency.

I understand consolidation through two main aspects. The first refers to the level of electoral security of dynastic candidates. The greater the margin of victory for dynastic candidates, the greater their electoral security and the greater the dynasty's stability. On the other hand, dynasties with less electoral security may resort to illegal mechanisms to strengthen their candidacy and increase their chances of being elected, thus less inclined to promote transparency.

The second aspect refers to the alignment of dynasties with formal spheres of power. In this sense, if political dynasties can be understood as informal organizations for exercising power, political parties are the main formal institutions that allow politicians to exercise power legitimately. Parties define candidates, political platforms, aggregate voter preferences, and coordinate the actions of parliamentarians, among other aspects. Parties are the main channel of formal politics (although they are not limited to formal mechanisms for exercising power) (BESLEY, 2005; FIGUEIREDO; LIMONGI, 2001). Thus, dynastic politicians who have partisan alignment with state and national executive leaders may have more incentives to comply with the AIA. More transparent municipalities, with institutionalized and well-implemented mechanisms for providing information, may increase their ability to capture federal and state resources. With that in mind, I propose the following conditional hypotheses.

(H2) The higher the margin of victory, the lower the conditional marginal effect of dynastic mayors on local compliance with AIA.

(H3) If the mayor and the governor are from the same party, the conditional marginal effect of dynastic mayors over local compliance with AIA will be lower.

(H4) If the mayor and the President are from the same party, the conditional marginal effect of dynastic mayors over local compliance with AIA will be lower.

4 IDENTIFYING DYNASTIC MAYORS IN BRAZILIAN MUNICIPALITIES

In the previous chapter, I discussed the definition of political dynasties and presented the main argument of this dissertation. Defining dynastic politicians as the ones *whose at least one relative has been previously elected, in consecutive elections or not, and regardless of being in the same jurisdiction or not*, I contemplated the association of dynastic mayors with *de jure* and *de facto* compliance with access to information depending on the context in which politicians are inserted. I argued that consolidated dynasties allied with state and national politics tend to comply with AIA. Also, the decaying dynasties (conditional to the margin of victory in local elections) will dwindle the level of compliance.

Testing these hypotheses is only possible by adequately identifying dynastic mayors in Brazilian municipalities. In this chapter, I discuss the measures usually used in the literature and how the few studies on political dynasties in Brazil have been coding dynastic politicians so far. I reflect on the advantages and caveats of these approaches. I suggest a new strategy for this empirical challenge: using Google search results about elected mayors and trying to identify whether the text indicates dynastic relationships or not. I then describe the process I followed to build this measure, from data collection to prediction validation. I close this chapter with a brief note on the approach's main benefits and limitations.

4.1 EXISTING MEASURES OF POLITICAL DYNASTIES

Concerning the identification of political dynasties, the standard approach in the literature uses a dummy variable equivalent to 1 if the candidate is dynastic and zero if is not. What varies is what scholars define as dynastic. The first variation is that some studies focus on Congress, while others study the executive branch. When authors separate by the degree of kinship, they generally distinguish between people who are blood relatives, those who are in-law relatives, and those with marital ties (GEYS; SMITH, 2017). Studies with a very large N, such as Bragança *et al.* (2015), use an automatic and less restrictive classification, identifying family ties only by matching the candidates' surnames.

Regarding datasets, some countries have rich sources on dynastic kinship. In the UK, for instance, Van Coppenolle (2017) uses biographical data that report individual characteristics for all House of Commons members elected in general elections since 1832 up to and including 2005. In Ireland, Folke *et al*. (2016) code dynastic ties based on verified information from

yearly political almanacs, biographical dictionaries, and newspaper reports. For Sweden they use a complete record of every politician's ID code. They match each person to his or her 1) siblings, 2) parents, and 3) spouse, using highly accurate register data.

Another usual approach to code dynastic politicians is to use electoral data on elected officials and candidates and the tradition of the family name structure in a given country to match relatives by their surnames (BRAGANÇA *et al.*, 2015; LABONNE *et al.*, 2019; QUERUBIN, 2011, 2016; ROSSI, 2017). Bragança *et al.* (2015) were the first to explore this possibility in Brazil. Oliveira and Souza (2022) also use this approach. Albeit this can be an accessible solution to identify political dynasties in Brazil, its main caveat is the possibility of having a significant amount of false positives, that is, individuals who share a surname but are unrelated. Many surnames are common and may lead to this situation. Bragança *et al*. (2015) try to circumvent this obstacle by dropping the ten most common surnames in the sample.

However, it is also important to consider cultural aspects when discussing this measure. Let us use the Philippines as an example. Labonne *et al*. (2017) use name matching to identify dynasties and explain that false positives, i.e., individuals who share a surname but are not relatives, should not substantially affect their measure because of how family names are defined in this country. As they state, "(...) Governor Narciso Claveria y Zaldua created a catalog with a list of 61,000 different surnames. A different set of surnames (often starting with the same letter) was assigned to each town, and local officials had to assign a different surname to each family head." (LABONNE *et al.*, 2019, p. 8). In that way, common surnames are not a problem in the Philippines. On the other hand, even though Brazil and the Philippines share a similar name structure (i.e., first name + mother surname + father last name) (BRAGANÇA *et al*., 2015; LABONNE *et al*., 2019), in the former, it is much more common to share a surname despite kinship.

In Brazil, the main source of information about electoral candidates is provided by the Electoral Superior Court, henceforth TSE (*Tribunal Superior Eleitoral*). Municipal elections datasets present information about the elections results, percentage of votes, and personal data of the candidates, such as complete name, occupation, education, age, sex, and so on. However, there is not a dataset including the affiliation of the candidates, i.e., the name of the mother and/or the name of the father. *DivulgaCand*, a website from TSE, provides granular information on each candidacy. In some cases, there are available documents containing the names of the candidates' parents. However, there is still no certainty about how comprehensive this information is since many candidates do not have documents with this data in them.

Given these challenges, I propose exploring a new source of information about political dynasties: Google search results on elected officials. As Boas *et al*. (2018) explore in their research, local blogs are a rich source of information on local politics. News related to elections, such as the formation of the political groups (who is going to "adhere" to which political group), party conventions, election rallies, the release of election results, and so on, can present valuable information not only about kinship among politicians but also whether they are present or not in other municipalities, if there is a conflict inside a dynasty and other interesting topics. Also, news on tributes to politicians is quite rich about their biography and can also be an important source. These are only a few examples of how local news can help to find details on dynastic kinship among elected officials.

I propose combining automatized data collection, with Google's custom search JSON API and Beautiful Soup, with supervised classification to try and automatize the coding of politicians as dynastic or not based on Google search results. This approach can rely on actual reports of kinship between politicians. Also, because it collects texts that include information about politicians, it could be explored in the future as a source for qualitative research on political dynasties. As the news describes local politics, it can be used to gather detailed aspects such as how long the dynasty has been in power, if they are present in other municipalities or not, from what perspective local news describes these families, and how dynastic politicians communicate with their electorate (e.g., for news regarding election rallies).

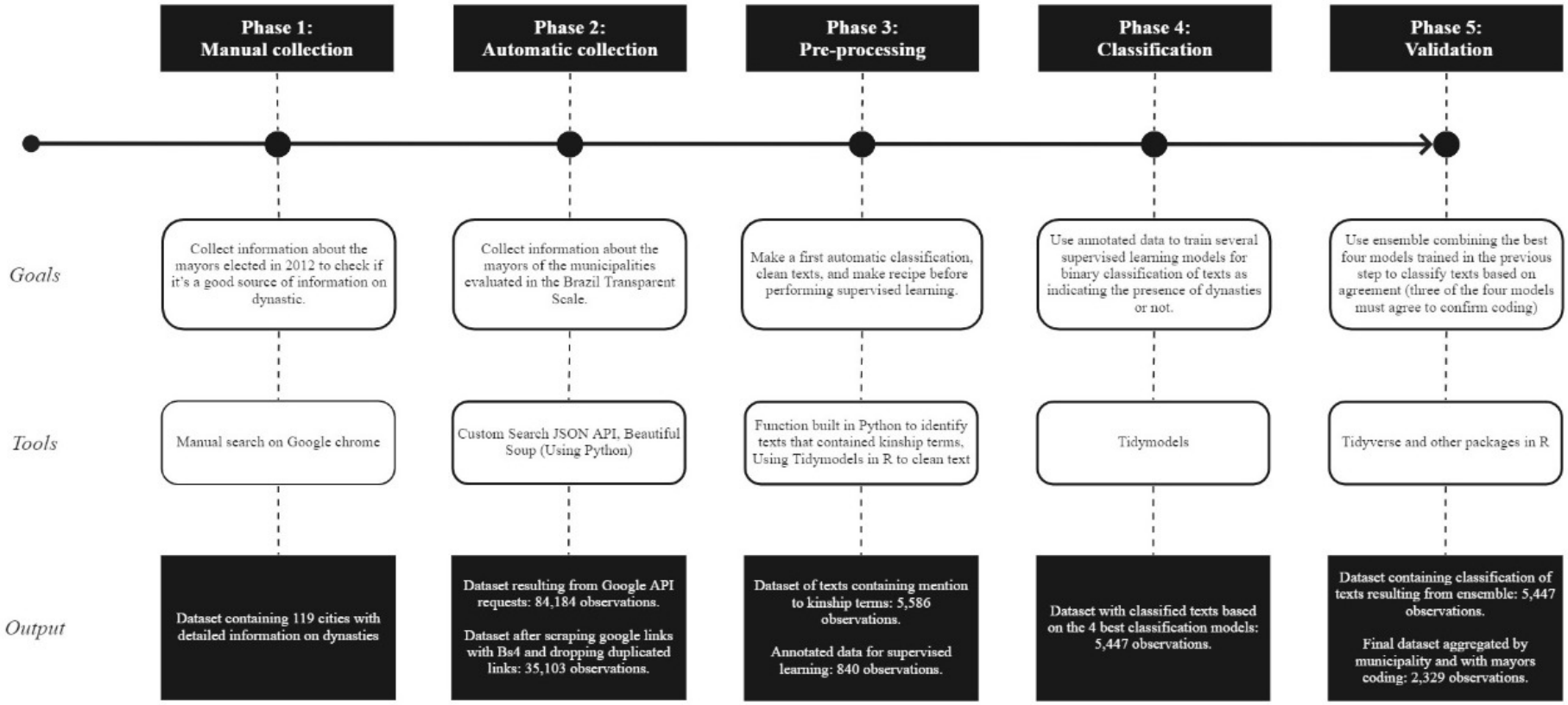
Yet, as text methods are incorrect at some level, in the sense that they will not fully capture the level of complexity that language entails (GRIMMER; STEWART, 2013; ROBERTS, 2016), using API, web scrapping, and supervised learning cannot be as accurate as manually searching and checking for information on politicians. In the next section, I discuss the step-by-step of this process and how I tried to overcome the limitations regarding this measure.

4.2 PROPOSING A NOVEL APPROACH TO IDENTIFYING DYNASTIC POLITICIANS

In this section, I describe the step-by-step identification of political dynasties based on Google search results. The process involved five phases: 1) manual search for information about mayors elected in 119 cities in 2012; 2) automatized collection of Google search results about mayors elected in 2012 in the cities that took part in GCU's Brazil Transparency Scale Evaluation; 3) pre-processing of the texts collected in the previous phase; 4) supervised learning

for binary classification; and 5) validation of the binary classification using ensemble technique and municipal-level calibration. This process is summarized in Figure 1.

Figure 1 - Identifying dynastic mayors with Google search results: Step-by-step.



Source: The author.

4.2.1 Exploring manual search results for mayors elected in 2012

The first phase in the process to build a measure of dynastic politicians based on Google search results was manually collecting information about the elected mayors in 2012 on Google. The goals in this search were twofold: 1) checking variability in the presence/absence of political dynasties in the municipalities; 2) examining the richness of information provided by Google search results. For that, I manually collected information about mayors elected between 2000 and 2016 in 119 cities randomly drawn from CGU's sample in the Brazil Transparency Scale (*Escala Brasil Transparente* - *EBT*, in Portuguese) evaluation. EBT will be used to measure the dependent variable in this research - compliance with access to information at the municipal level.

CGU's evaluations were performed between January 2015 and January 2017, which means the mayors of the evaluated cities were elected in the 2012 elections. The first EBT (performed between January 12th, 2015, and May 4th, 2015) evaluated cities with a population greater than 50,000 inhabitants, including state capitals, which resulted in 492 cities. In the second edition (from July 27th, 2015, to October 9th, 2015), they evaluated municipalities included in the first edition, plus 1,905 cities (1,587 cities in total). This number includes local governments that *asked to be evaluated*. Unfortunately, it was not possible, even after formally contacting CGU, to identify which city asked to be included in the evaluation. In the third and last edition (from June 27th, 2016, to January 16th, 2017), CGU evaluated the cities included in the second edition, plus 741 municipalities, resulting in 2,328 total. Since the electoral data is the same (2012 elections), I dropped repeated municipalities in the final dataset, which led to an N = 2,328.

After drawing the 5% representative sample at a 95% confidence interval, I performed a manual search on Google about the mayors of the remaining 119 cities. The intention was to collect information from local blogs that thoroughly describe local politics. To that end, I performed several searches on Google for each municipality. Firstly, searches had the following structure: "name of the city + state" + "*lista de prefeitos*" (which means a list of mayors, i.e., a list of all the mayors elected in the municipality, what was available in some cases). Subsequently, the search's structure was "name of the city + the initials of the state + the name of the elected mayor", including all elected mayors from 2000 to 2016 for each of the 119 cities.

I checked all the search results, discarding the ones unrelated to the candidates and reading any news that could bring information about the politicians. If, after these searches, I did not find any result indicating kinship between politicians, I performed a search including names associated with family ties in Portuguese. For instance: “mayor’s name + city’s name + state + *mãe* (mother)”. Another way to find information was to check if any ex-mayors had passed away. This type of news usually presents the background of the person who has passed away, which was very informative for this research. If, after all these searches, I still did not find any indicator of political dynasties, I coded the mayors as non-dynastic (0). The city received a code 0, meaning a political dynasty is absent.

Throughout the search, it was possible to find many details about the families. Some examples are a) identifying who started the dynasty (i.e., the first member of the family to be elected), b) checking if the family had some kind of disagreement that resulted in two distinct political groups, and c) verifying if the family had members that were elected mayors/city councils in other cities or that were elected as state/federal deputies and/or senators and d) learning about political campaign strategies. That was not always the case, but I registered as much information as possible.

A good example of the cases in which ample information is available about local politics is the case of Abreu e Lima, a city in the state of Pernambuco, in the Northeast region of Brazil. Here is an excerpt of a local blog’s post about the ex-mayor Jerônimo Gadelha, from which it is possible to learn about politicians’ kinships, confirm the presence of the Gadelha dynasty in the city, and identify conflict inside this political group:

> Jerônimo Gadelha was once one of the most important mayors in the metropolitan region [of Recife, Pernambuco], courted by governors, deputies, and other politicians. His fall into politics, according to residents of Abreu e Lima, began with the alleged "betrayal" of his nephew, former mayor Flávio Gadelha (PMDB). According to local residents, [Flavio] who was elected mayor in 2004, broke up with his uncle, [who were also] his political godfather, just hours after the results of the polls came out (BLOG DO JAMILDO, 2017, free translation).

In general, each city had 15 to 25 relevant news. For one city, there was not enough data available on Google to either confirm or discard the presence of political dynasties. This case was coded as “NA”. Finally, in some cases, there were two or more former mayors with a surname in common. However, neither Google results nor electoral data from the Electoral Superior Court (TSE) were sufficient to confirm the kinship. These cases were coded as “NC”, i.e., “Not Confirmed”.

In summary, cities were coded as 1 if they had at least one mayor or ex-mayors (from 2000 to 2016) coded as dynastic (presence of political dynasty), and 0 if the city did not have any mayor or ex-mayor confirmed as dynastic (absence of political dynasty), NA for cases with insufficient information on Google and on TSE, and NC for cases in which kinship between politicians was not confirmed. After this search, from the 119 municipalities, 58 cities were coded as 1 (presence of dynasties), 55 as 0 (absence of dynasties), 1 as NA, and 5 as NC. In 2012, 93 cities were coded as 0, 20 as 1, 4 as NC, and 2 as NA.

Even though I tried to be as transparent and as rigorous as possible, recognizing the limits of this coding is unavoidable. Its main caveat is that only one person performed it. I doubled checked the data, but the ideal situation would be to have a second person checking the coding. Yet is crucial to bear in mind that the main goal of this coding is explorative. For this aim, the collected data is extensive and allows the conclusion that the presence of political dynasties in Brazilian municipalities is currently diverse. In other words, there is indeed variation - not only in the presence/absence of dynasties but also in the structures of families that participate in politics. Also, manual search indicates the richness of the Google search results as a source of information about dynastic politicians. The possibility of emulating the manual coding in an automatized way will be discussed in the next section.

4.2.2 Automated collection of Google search results

As discussed, Google results can return local blog posts that detail the local political context. In this sense, adopting Google search results as a source for political dynasties coding can be a successful path to identifying dynastic politicians. However, manually checking dozens of news for more than 2,000 mayors for this dissertation was unfeasible. One solution would be trying to automatize this process.

With that in mind, I propose using Google search results about elected mayors as a novel source for political dynasties coding. To do that, I used Custom Search JSON API, and Beautiful Soup. The former is an API from Google that provides titles, snippets, and links to Google search results (GOOGLE, 2022). The latter is a Python library for web scraping (BEAUTIFUL SOUP, 2022).

The Custom Search JSON API allows requests to be made with an instance of a Programmable Search Engine. I created a search tool that did not limit the type of sites that would be returned in the search. I also specified that the results should be in the Portuguese

language and for Brazil[1]. Google allows a user to create a free Google cloud account with a limit of 300 dollars. In addition, Google also allows 100 requests to be made per API key (GOOGLE, 2022). Since this research had no funding[2], I opted for the free versions, limiting the number of collected results.

After creating the account and the Google search engine (GSE), it is possible to use the API to collect Google search results. As Google offers API usage guidelines for Python, I decided to use this language to perform the analysis. The main challenge in this process was to automatize the API request input. In this sense, I built a function that creates the search tags of the Google search function based on a TSE database of mayors elected in 2012 (and, therefore, municipal managers in 2015 and 2016, the period of evaluation of the Brazil Transparent Scale)[3]. The search tags followed the manual search structure: name of the city + state + name of the elected mayor. Requests were made both for the mayor's full name and ballot name. For the mayors of the cities included in the first edition of the Brazil Transparent Scale, I also tested other search tags, including the type of kinship in the tag, for instance, "*irmão* + city + state + elected mayor's name" (for brother). The results of these alternative search tags did not seem very different from the main ones (complete name and ballot name).

For each request, the API returns 10 results (just like when you manually search for something in the Google search bar). It is also possible to obtain the results of subsequent pages. Based on Google's algorithm and the experience of manually collecting this data, I knew the search results become less related to the search request as the pages go through. In this sense, I only collected the results of the first two pages of the Google search. It is essential to remember that the more pages, the more requests for the same candidate and the less the possibility of using free keys for data collection. Limiting the results to two pages allowed collecting results related to what was searched and performing searches for all mayors in the CGU sample.

As a result, I collected 84,184 search results containing the title, snippet, and link of each web page. The snippet is an excerpt of the text on the webpage, possibly the preview of the content that appears below the title and link on the Google results page. Nonetheless, using

[1] The arguments were the following: country='countryBR'; language='lang_pt'; termsToExclude = ['médicos jus brasil google maps'] (somewhat frequent terms in the tests I did before running the search for the complete dataset); numberOfResults = 10.

[2] The absence of funding is exclusively related to the collection of data. As a PhD student, I received a CAPES scholarship, for which I am very grateful for. Without this financial support, I, like many other graduate students, would not be able to have an exclusive dedication to research.

[3] I would like to thank Cláudio Alves Monteiro, who significantly helped me to learn Python programming language and revised my coding to collect automated data with Google's API and Beautiful Soup.

only this data, the content for coding the dynasties would be pretty limited since the snippet contains few lines of text. Also, the chances that the snippet contained a reference to kinship among politicians would be smaller than using the complete text from the page.

There were two possible solutions in this case. The first would be to try to collect Google search results via Beautiful Soup (bs4). This Python library allows you to extract data from HTML and XML files. Through Beautiful Soup, it is possible to collect titles, snippets, and links from Google search results, but also the entire content of the page. Compared to a Google search done with only one candidate, a single input to the request, the Beautiful Soup results are cleaner and better structured than those from the Custom Search JSON API. Also, its results are similar to what I found while performing the search manually. This is expected because the browser's history of searches influences results in bs4. Howbeit is not clear how Google's API algorithm performs its research, it does not seem to be influenced in that way.

Thus, I tried to collect Google search results in an automated way for the entire database of mayors elected in Brazil in 2012 with Beautiful Soup. The results contained several errors, which possibly stem from barriers imposed by Google aiming to prevent web scraping. It is important to highlight that, although no Google guidelines were found prohibiting web scrapping, the high volume of requests made in a short period can overload servers, and if the company imposes protection barriers, it is understood that persisting in the attempt to scrape its data would be unethical.

The solution, in this case, was to use the links provided by Google's Custom Search JSON API and, from them, scrape the content of the web pages. With this strategy was possible to obtain the full text from the Google links, which were later tokenized. As bs4 scraped each link separately, instead of scraping Google search result pages, it did not have the risk of overloading the servers. To reduce the size of the final dataset and facilitate its handling, the observations in which there was an error in the return of the scrape were deleted. This resulted in 46,162 observations. After dropping repeated page links, the dataset had 35,103 observations for all the mayors from the cities evaluated in the Brazil Transparent Scale.

4.2.3 Pre-processing the text extracted from Google search results

Pre-processing is a fundamental step in performing supervised learning classification (HVITFELDT; SILGE, 2021). Regarding this point, the quality of the text extracted from Google's search results is one of the main challenges in measuring political dynasties with this

data. In the first place, the text consists of the entire web page content. This includes not only the main text from the web page but also tags, related links, and so on. Below is an example of a web page delivered by Google's API, the *Blog do Finfa*.

Figure 2 - Example of a web page from Google results scraped with Beautiful Soup.

Source: BLOG DO FINFA (2014). Available at: <https://blogdofinfa.com.br/category/brejinho>.

This web page is a tag link for all the blog news related to the city of Brejinho. That means that this link contains several blog posts. There are other links with this same configuration in the dataset. However, this type of result is not the majority of the texts. This figure also shows other elements that will be captured with Beautiful Soup but are irrelevant for classification. For instance, the menu of the website at the head of the page (i.e., "Home",

"Contact", etc.). Another issue is that because I could not collect only the page's main text, it is not possible to separate the text by paragraph. This could be an important aspect of the text structure for supervised classification. Figure 3, as follows, was extracted from Blog do Finfa's web page. It shows the case in which the references for a dynasty are in different paragraphs.

Figure 3 - Excerpt of a web page containing dynastic kinship indication.


5 de setembro de 2013 / Comente sobre

Prefeito José Vanderley e Seu João Pedro

Faleceu na noite desta quarta-feira (04) João Manoel da Silva, Seu João Pedro, ex-prefeito por três vezes da cidade de Brejinho.

Considerado um dos maiores políticos do Pajeú, seu João Pedro, como era conhecido popularmente, exerceu o cargo de prefeito por três mandatos e um mandato como vice-prefeito ao lado de seu genro o atual Prefeito de Brejinho José Vanderlei (PSB).

João Pedro, governou Brejinho, em uma época que não existia Bolsa Família e quando a seca castigava e, muitas vezes não exercendo o cargo de prefeito, João Pedro não hesitava e as famílias carentes do seu município.

Homem público, pai de família, líder político, de caráter forte mais, ao mesmo tempo enfrentava os principais problemas de Brejinho; João Pedro deixa mais que saudades a parentes e amigos, deixa a história de um homem que ajudou a desenvolver um pedaço do Pajeú, homem simples de uma conversa franca, c,de muitos amigos e de poucos inimigos, que ao fim da vida fez questão de procurar e apaziguar as desavenças, Seu João, chegou a quase 90 anos, com a saúde já debilitada, faleceu na noite de ontem (04) no Hospital Regional de Patos, de falência Múltipla dos órgãos.

Source: BLOG DO FINFA (2013).

In a free translation, the paragraph in highlighted with a blue line means (text in bold refers to sentences highlighted in red in Figure 3):

> João Manoel da Silva, Seu **João Pedro, three-time former mayor of the city of Brejinho**, died on the night of this Wednesday (04).

> Considered one of the greatest politicians in Pajeú, João Pedro, as he was popularly known, **held the post of mayor for three terms and one term as vice mayor alongside his son-in-law, the current Mayor of Brejinho José Vanderlei (PSB).** (BLOG DO FINFA, 2013, free translation, added emphasis).

Beautiful Soup contains a function called soup.findAll. It allows you to extract information from several Google results simultaneously (RICHARDSON, 2007). However, this is not available to extract the main text from the web page but rather to return information such as the title, the snippet, and the link from the pages. An alternative would be scraping the page with the class related to the website's main text. The matter is that this class changes among web pages, so it is not possible to use a pattern to scrape all pages with a unique function. Hence, the best solution available was scraping the page's entire content and cleaning the text afterward.

Three steps were involved in this pre-processing phase. First, I made a function in Python that simply identifies whether the text contains the terms related to kinship and political office or not. The goal was to discard Google results that were completely unrelated to kinship. The final function to code the texts included the previous 7 functions made to identify specific kinship terms and the function built to detect mentions of political office[4].

The terms included in the function contained a space between the quotation marks, indicating the expression or word needed to be separated in a sentence. For example: " tio do ". This aimed to reduce results including kinship terms yet unrelated to the topic - e.g., when the expression " tia do " (aunt of) was part of unrelated sentences, such as "(...) partia do princípio que (...)".

To check the precision of this coding, I performed a manual check on the texts coded as "1", that is, text containing at least one mention of kinship and political position. I manually checked all the links coded as 1 for the municipalities evaluated in the first Brazil Transparent Scale edition. This resulted in 867 texts after eliminating links that were either duplicated or corrupted. A manual checking indicated a 9% compatibility between the automatic and the manual coding. That is, of the 867 texts classified by the function as containing dynasty indication, only 78 of them have this indication.

A second check of the automatic coding resulting from applying the rule of kinship and political office was carried out at the municipal level. I aggregated information from the 867

[4] The terms included in the functions to check for kinship and the function used to check for political office are available in the appendix.

texts and coded as dynastic those mayors whose texts contained at least one mention of kinship and political office (coded as “1”). I assumed that the mayors whose texts did not return any mention of kinship and political office were not dynastic (coded as “0”).

I compared the automated coding results with the manual coding of the 119 municipalities randomly selected from the CGU sample in the Brazil Transparent Scale (EBT). The compatibility between the two classifications was 28.2%. I analyzed the cities where the mayors were correctly classified as dynastic and identified that they were places in which the dynastic presence was quite strong. This result is intuitive since the more influential a political dynasty is, the more the news will be published about them.

This information sheds light on one limitation of using Google search results. Dynasties with more political members or who have been in power for a longer period are more likely to have more news on Google. An implication is a difficulty detecting less well-established dynasties because there may be less information available about them. In addition, there is no specific feature in the structure of the correctly coded texts of the municipalities that differentiates them from the other texts collected.

One way to circumvent this caveat is to gather more information about each mayor. For instance, collecting the first two pages of Google search results, 20 links per mayor. Also, with the manual coding, gathering information about minor dynasties was possible. That richness of information allowed me to identify both incipient and consolidated dynasties. In this way, this 28% compatibility indicates how the kinship and political office terms can still wrongly code as dynastic texts that are unrelated to dynastic ties. That is not a problem since I only applied the Python function to filter Google search results and arrive at a cleaner dataset allowing a better-supervised classification of the texts.

After dropping the texts that did not contain any mention of kinship terms and political office, I performed the second pre-processing step: cleaning the text with the usual proceeding researchers follow when dealing with this type of data: removing accents, special characters, extra space, etc. I also removed words that appeared to be in a foreign language, some of them containing only consonants. The third and last step, also performed in R, used Tidymodels’ Recipe to transform the text through three steps: step_tokenize(), step_tokenfilter(), and step_tfidf(). The first one tokenizes the text, limiting the number of tokens to 1,000. The second will turn the token into a variable that allows filtering the token by its frequency. Finally, the last step transforms the tokens into several variables containing the term frequency-inverse document frequency. Then, the tokens are inserted as predictors in the models that use this step

in their recipes. In that way, the most frequent the word, and, therefore, the less it contributes to supervised learning, the lower the importance of the token in the model (KUHN; WICKHAM, 2020).

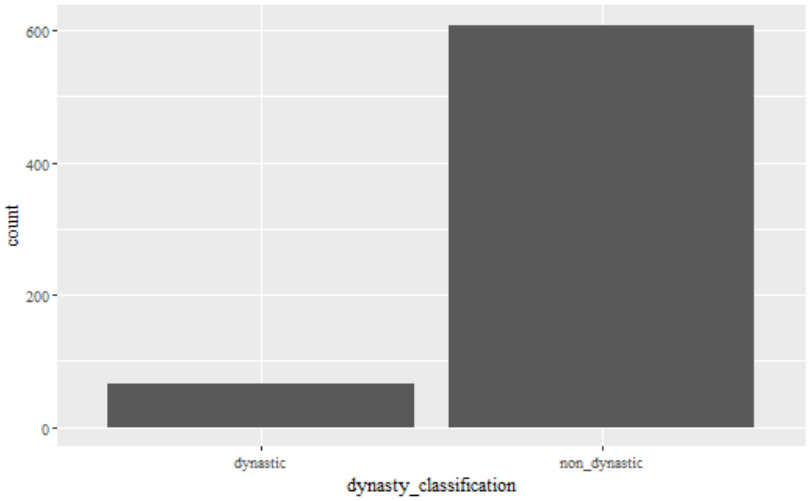
4.2.4 Coding dynastic politicians with Supervised Learning Classification

In this subsection, I discuss how the data collected via Google API and Beautiful Soup and filtered with a Python function was used to perform supervised learning classification. Supervised models can solve one of these three problems: 1) generate accurate predictions with predictive models; 2) test hypotheses with inferential models; and 3) describe properties of observed data with descriptive models (HVITFELDT; SILGE, 2021). In this thesis, I will use predictive models to solve a binary classification problem (*ibidem*). The goal is to train models capable of predicting whether the text is part of the class of texts that indicates dynasty or not.

Supervised learning uses the outcome variable to predict classes (KUHN; JOHNSON, 2013). In this sense, I use a dataset of 840 manually classified observations to train the supervised models I tested for this measure. In the previous subsection, I explained how I checked the Python function to filter texts for supervised learning. I selected all the links for the municipalities evaluated in the first edition of the Brazil Transparent Scale, which led me to 867 texts. To train the models, I excluded links that did not contain texts, such as those related to the *Cadastro Nacional de Estabelecimentos de Saúde* (National Register of Health Establishments) forms or links for downloading documents.

This generated a final dataset of 840 links, of which 78 were manually classified as dynastic, i.e., the text indicates the mayor is dynastic. The remaining texts were classified as non-dynastic. Non-dynastic cases mean that the text did not mention dynastic kinship among the mayor and other politicians. This includes cases in which the absence of dynastic kinship was confirmed, such as texts containing a small biography of the mayor, and texts that do not approach anything related to dynastic relations. Thus, the data used to train and evaluate the models contains class imbalance. The positive class, which confirms the presence of dynastic ties, is the minority class (9% of the sample). Figure 4, as follows, shows the difference between the two classes.

Figure 4 - Class imbalance in the manually classified dataset.



Source: The author.

Class imbalances affect models predictions by decreasing its sensitivity (i.e., the true positives rate), while its specificity (true negative rate) is high, which generates high accuracy measures that do not proper show the quality of the model (KUHN, JOHNSON, 2013). Among the strategies to deal with class imbalance, tuning the predictive model in a way that maximizes the accuracy of the minority class (i.e., texts that indicate dynastic ties) is a possible solution (*ibidem*).

For this analysis, I tested several classification algorithms, namely: Naive Bayes Model, Lasso Regularization Model (in three variations: with fixed penalty (at 0.01), with penalty defined with regular grid, and with sparse encoding), Bagged Tree, Quadratic Discriminant Analysis, Random Forest, Random Forest with Case Weight, Extreme Gradient Boosting (with and without classification cost in tuning), and Support Vector Classifier. The models that presented the best performance were Random Forest, Extreme Boosting (in the two variations), and Support Vector Classifier. The Quadratic Discriminant Analysis and Random Forest with Case Weights presented problems in execution and could not be compared to the other models.

In this subsection, I discuss the performance of the four main models: Random Forest, Extreme Gradient Boosting, Extreme Gradient Boosting including Classification Cost, and Support Vector Machine. Random Forest (RF) can be applied for binary classification problems by using a classification tree. This algorithm is a modification of bootstrap aggregation, also known as bagging. Each tree in the forest will give a vote in the classification for a new sample. The predicted probability vector is formed by the proportion of votes in the ensemble (KUHN; JOHNSON, 2013).

For this model, I used a tune grid for two parameters: mtry and ntree. The first represents how many variables are going to be randomly sampled as candidates for each split of the data. Ntree regards the number of trees to grow when fitting the model (BROWNLEE, 2020). The values were mtry = 3, 4, 5, and ntree = 500. However, Kuhn and Johnson explain that Random Forest models are somewhat insensitive regarding mtry values. Tidymodels allows the use of RF with rand_forest specification, mtry tuning parameters, and ranger as the engine. Ranger creates several decision trees which are independent of each other. As it is an ensemble approach, the final prediction combines the predictions previously created for each tree (KUHN; JOHNSON, 2013; KUHN; WICKHAM, 2020).

Boosting algorithms were first presented in the 1990s with the approach of combining weak classifiers to create an ensemble classifier. To do so, they used the best-generalized classification error rate (KUHN, JOHNSON, 2013). The first boosting algorithm was the AdaBoost. Currently, the stochastic gradient boosting machine is the most used (*ibidem*). In this approach, the algorithms use weak learners to find an additive model that can help to minimize the loss function. This process happens through the calculation of the residual of the best prediction of the response (called gradient) and these values are used to fit the model and minimize the loss function. This additive process repeats subsequentially until it reaches a specified number of iterations (*ibidem*).

In this sense, the values of each tree depend on the previous values. I used *xgboost* from the Tidymodels framework to train the model (KUHN; WICKHAM, 2020). For tuning the *xgboost* model I used the early stopping feature. The models used a validation set of 0.2 to limit the proportion of analysis sets the models can use to decide when it should stop iterating (SILGE, 2021). The number of trees is held constant at 500 to avoid a significantly high value (*ibidem*). Additionally, I trained a variation of this Boosting Model specification. In this modified version I tuned the model with the classification cost, from the Yardstick package that is part of the Tidymodels framework (KUHN; WICKHAM, 2020). This is aligned with what Kuhn and Johnson (2013) call cost sensitivity training.

Finally, for the Support Vector Machine Classifier, I use a polynomial SVM model. I set the linear model specification to have a degree of svm_poly() = 1. I also set the engine as kernlab. In this way, the model will try to maximize the thickness of the margins between the classes. Additionally, I tune the model with cost = 10 (HVITFELDT, 2021).

However, changing model tuning may be insufficient to deal with class imbalance, because it does not generate enough increase in sensitivity (i.e., the true positive rates) (KUHN,

JOHNSON, 2013). Thus, Kuhn and Johnson (2013) also suggest alternating cutoffs. The ROC (Receiver Operating Characteristic) curve can be used as a cutoff to determine the best fit since it calculates both sensitivity and specificity for a continuum of cutoffs. Another possibility is using the Youden Index, also called the j-index, which indicates the capacity of a diagnostic test to determine if a patient has or not a given disease. For predictive models, it indicates whether the model is capable of correctly predicting the positive class (YOUDEN, 1950).

Both the ROC curve and the j-index range from 0 to 1 and indicate that the closer to 1, the better the accuracy of the model. The Youden Index is calculated as sens() + spec() - 1. J-index is also considered a measure that summarizes the ROC curve. Thus, I used the j-index as the cutoff point to choose the best predictive model (RUOPP *et al*., 2008), except for the Extreme Gradient Boost model with classification cost, in which I used the latter as a cutoff.

Sampling methods can also help to deal with class imbalance (KUHN; JOHNSON, 2013). This approach aims to reduce the impact of class imbalance in the model training by balancing the classes in the training set. Sampling methods can be performed *a priori* or *post hoc*. The latter involves *down-sampling* and *up-sampling* (*ibidem*). Following the option available in the Tidymodels framework, I tried both approaches, down-sampling, and up-sampling. I also used combinations of the two strategies, such as step_SMOTE, and step_rose. Down-sampling offered the best results, so I applied it to the models' recipes.

Besides text pre-processing and down-sampling that were included in the recipes of the models, I also used step_nzv() to remove parameters that do not contribute to prediction, step_normalize() to make numeric parameters normally distributed, and step_dummy() to transform categorical variables in traditional dummies. The parameters included in the model were mainly derived from the text. Table 2 presents the description of each variable included in the analysis.

Table 2 - Variables included in the classification models (Continues).

Name	Type	Description
page_text_new_clean	character	Cleaned text scraped from Google search results links, collected via Custom Search JSON API and Beautiful Soup
sentence_words_categorical	character	A categorical variable based on the distance (number of words) between terms related to kinship (such as "filha de" (i.e., "daughter of") and terms related to political office (e.g., "senador", which means senator) found in texts that indicated dynastic kinship

Table 2 - Variables included in the classification models (Ends here).

Name	Type	Description
mayor_name_is_in_text	character	A dummy variable that identifies whether the mayor name or their ballot name is in the text (1) or not (0)
kinship_sentences_count	double	A continuous variable that represents the number of sentences in a text containing kinship-related terms
type_of_search_tag	character	A categorical variable that regards the search tags used in Google custom search requests
categorical_kinship_function_python	character	A categorical variable that indicates the type of kinship found in texts (if children, siblings, grandparents, etc.)
dynasty_classification	character	Is the outcome variable. It is the classification of the texts as indicating dynastic kinship or not

Source: The author.

After describing the main features of the models, it is possible to evaluate the metrics resulting from training and evaluation. The following table, Table 3, presents the main metrics regarding the best fit in the training phase. As already discussed, the cutoff to find the best fit of the models for Random Forest, Extreme Gradient Boosting, and Support Vector Machine was the J-Index. For the Extreme Gradient Boosting with Classification Cost model, I used the classification cost itself as the cutoff. Table 4 displays the following metrics: ROC curve, J-Index, Sensitivity, Specificity, and Classification Cost.

Table 3 - Metrics of the Best Training Results

Model	ROC Accuracy	J-Index	Sensitivity	Specificity	Classification Cost
Random Forest	0.755	0.401	0.784	0.617	-
Extreme Gradient Boosting	0.739	0.371	0.727	0.645	-
Extreme Gradient Boosting + Classification Cost	0.772	0.372	0.721	0.651	0.428
Support Vector Machine	0.715	0.254	0.566	0.687	-

Source: The author.

I explore the J-Index and the Sensitivity and Specificity metrics separately because the main challenge in the classification problem of this research is the prediction of the positive class, that is, the indication of dynastic kinship in a text. Due to class imbalance and possible non-linear distribution, this task is not easy.

ROC accuracies seem good and suggest that predicting text classes with the models is still better than defining classes randomly. However, the J-index indicates that the balance between positive and negative classes is not ideal. For a test to be confidently able to define the positive cases Youden's Index should be greater than 0.5. Also, sensitivity and specificity values demonstrate that the measures to deal with class imbalance were able to increase the models' sensitivity. Yet, specificity, for 3 of the 4 models, is lower, confirming that both rates are not balanced.

Table 4, below, presents the evaluation metrics after fitting the models with the best J-index and cost classification metrics to the split dataset. Accuracy is not the best metric to use as a reference for data with class imbalance, but ROC accuracy could also be influenced by higher values of sensitivity.

Table 4 - Evaluation Metrics of the Fitted Models.

Model	**Accuracy**	**ROC Accuracy**
Random Forest	0.631	0.863
Extreme Gradient Boosting	0.649	0.831
Extreme Gradient Boosting + Classification Cost	0.738	0.763
Support Vector Machine	0.714	0.840

Source: The author.

Because of that, it is important to use other metrics to evaluate the models. One important and simple evaluation metric is the Confusion Matrix (HVITFELDT; SILGE, 2021). Table 5 summarizes its interpretation. The numbers in each quadrant sum up the cases predicted by the model. The darker the quadrants in the diagonal line, the better.

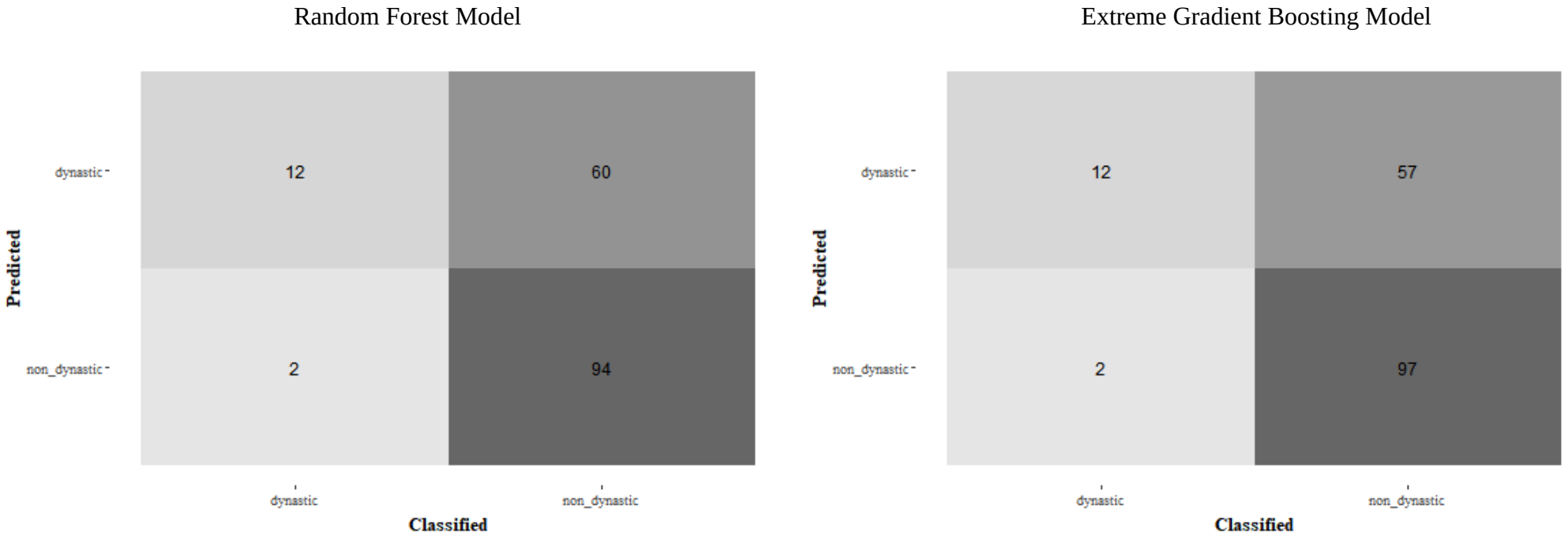
Table 5 - Interpretation of the Confusion Matrix.

True Positives	False Positives
False Negatives	True Negatives

Source: The author.

Figure 5, as follows, presents the confusion matrix for the four models. As expected, based on the J-index, the number of false positives is greater than the number of true positives, reinforcing that the main limitation of these models is their capacity to correctly predict the minority class. On the other hand, although training metrics indicated lower specificity, the models generated more true negatives than false negatives.

Figure 5 - Confusion Matrices for Random Forest and Boosting Models.



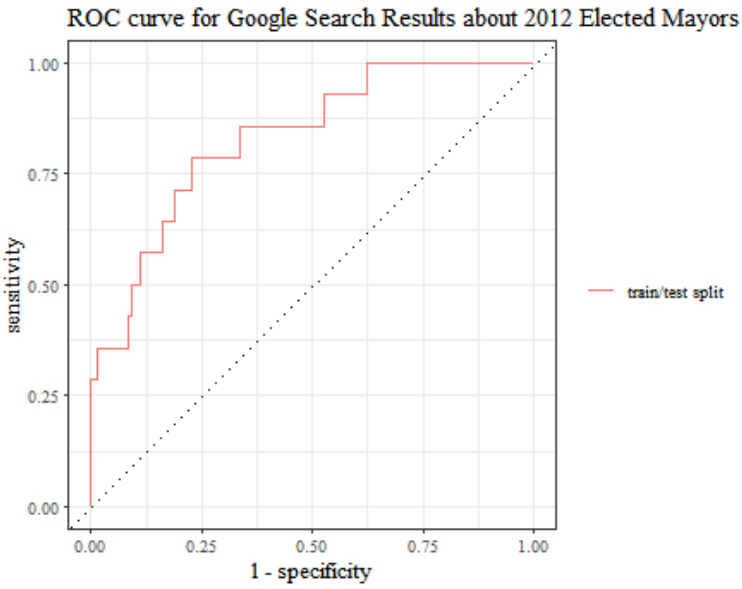
Source: The author.

Figure 6 - Confusion Matrices by for Boosting with Classification Cost and SVM Models.

Source: The author.

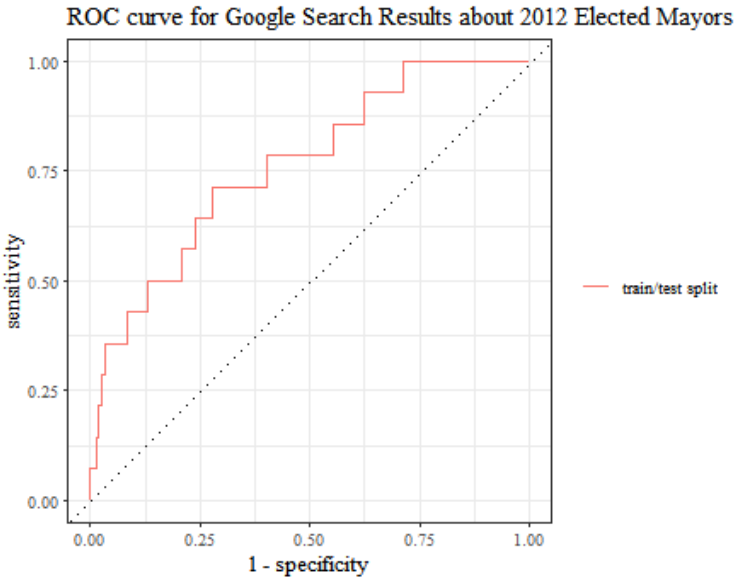
Another important measure is the ROC curve, also known as area under the curve. The wider the area, the better the model. If the curve was almost flat, this would mean the models are not able of confidently predicting the classes. The following figures demonstrate that the models are better than randomly assigning the classes of the texts. Hence, even though the models have notable limitations that need to be recognized, they are still capable of generating confident information about the texts that do not contain dynastic kinship indications.

Figure 7 - ROC Curve for the Extreme Gradient Boosting Model.



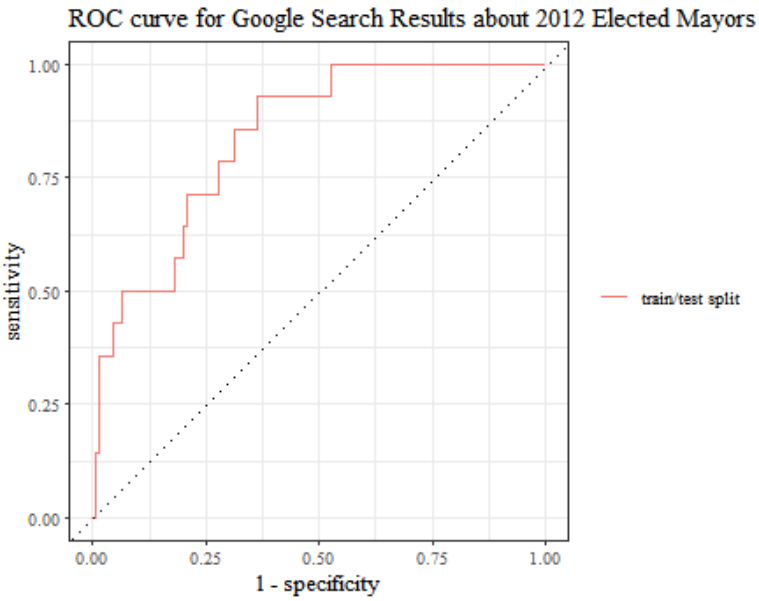
Source: The author.

Figure 8 - ROC Curve for the Boosting + Classification Cost Model.



Source: The author.

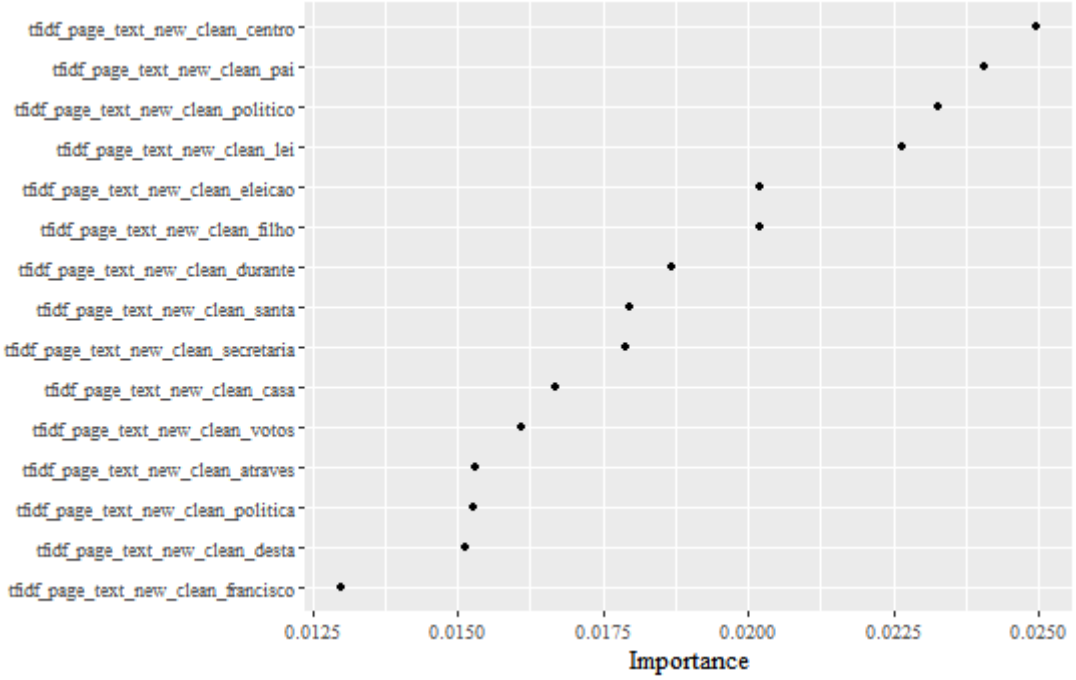
Figure 9 - ROC Curve for the SVM Model.



Source: The author.

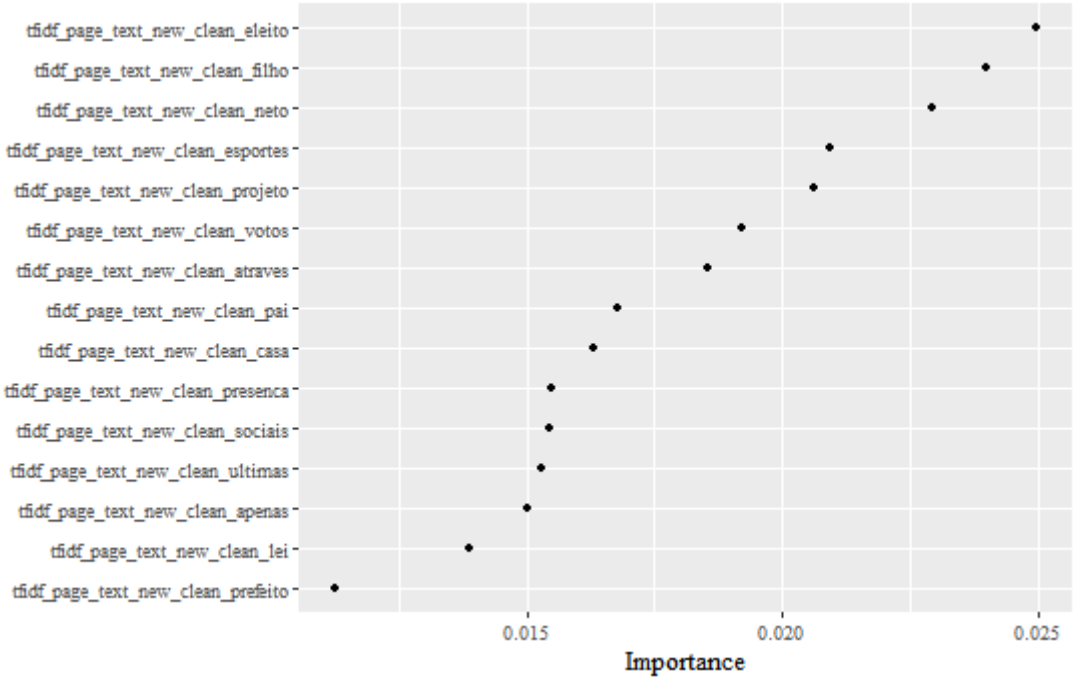
Finally, it is possible to check the most important variables for predicting a text as indicating dynastic kinship for the two specifications of the Extreme Gradient Boosting algorithm. Since I used term frequency-inverse document frequency, each token was transformed in a parameter for the model. In Figures 10 and 11 is clear that the most important variables for models' predictions are the text itself. Between the most important terms are tokens such as pai (father), político (politician), eleição (election), and filho (son). On the other hand, there are also words that do not seem directly linked to dynasties, e.g.: centro (center), santa (saint), and francisco.

Figure 10 - Variable Importance for Boosting Model.



Source: The author.

Figure 11 - Variable Importance for Boosting Model + Classification Cost.



Source: The author.

Taken together, these results suggest important limitations in using these models to predict text class and extract dynastic kinship from this data. The main caveat of these models is their capacity to identify the positive class. Metrics such as the j-index and the confusion matrix demonstrate that predictions still lead to a substantive number of false positives. On the other hand, measures such as the ROC curve, and the values of sensitivity and specificity signal that strategies such as down-sampling and model tuning cost-sensitive learning may have helped to deal with class imbalance. Yet, these corrections were insufficient to generate strong predictors for the presence of dynastic kinship mentions in the texts. In the next subsection, I discuss how to deal with these limitations.

4.2.5 Validating the identification of dynastic mayors with Google search results data

Individually, the models discussed in the last subsection cannot satisfactorily predict the positive class, i.e., whether there is a mention of dynastic kinship in the texts. The best models were the ones based on ensemble learning, such as Random Forest and Boosting. I intend to follow this ensemble logic and combine the four models previously discussed to arrive at a more complete and rigorous classification of the texts. I included the SVM classifier even though its j-index is the lowest one, compared with the other models. SVM training metrics showed higher levels of specificity than if sensitivity, which is different from the other models. This could help to balance classification.

To do so, I checked the four models' predictions and accepted their classifications if at least three of the four models agree in their coding. Hence, if at least three models agreed that a text does not indicate dynasty, it was coded as 0. Symmetrically, if three of the four models agreed that the text indicates kinship between politicians, it was coded as 1. For the other cases, that is, where half of the models coded texts as 0 and the other half as 1, I checked and classified the texts manually.

According to this approach, from the 4,607 automatic classified texts, 1,607 contained indication of political dynasty, being coded as dynastic. From these, 942 texts received code 1 in a consensus, while 665 were considered dynastic because three of the four models agreed. On the other side, 2,607 texts were coded as 0. All models agreed in 2,129 of the cases, while 478 texts received 0 code because the majority of the algorithms agreed. For 393 texts there was no agreement about the classification. I classified these texts manually. 21 of them

indicated the mayor had relatives in politics, while for 372 of the cases there is either no mention to kinship or dynasty, or the text confirmed that the mayor was not part of a dynasty.

While checking the cases with no agreement between the algorithms, some types of texts and news became more salient. These cases reveal challenges regarding dynasty identification through texts that could be circumvented in future version of this research, since they entail more profound changes that are not possible to execute as of now. In this sense, among the types of texts that do not indicate dynastic kinship there is considerable variation. In some cases, the text confirms that the mayor is non-dynastic. For instance, when the text presents a mini biography of the politicians and do not mention any relatives in politics or even states that the mayor represents renovation or that it has no traditional ties in politics.

Other news, the majority of them, are not related to dynasty. They approach different themes related to the mayor's management and so on. Some cases concern nepotism. There are still cases in which the web page's text is completely unrelated to the mayor. There are also cases in which the text mentions dynastic kinship, but it is unrelated to the mayor. This type of text is a complex case for the algorithm since its structure is similar to a content that indicates dynastic ties for the mayor. Finally, there are also situations in which the text informs that the mayor elected in 2016 is related to the politician elected in 2012. However, if the 2012 elected mayor was the first of her family to be elected, then she cannot be classified as dynastic.

Possible solutions for future developments of this measure are the following. First, increase annotated dataset to train the models and use multiclass classifications. This may help to reduce class imbalance by dividing the types of texts that do not indicate kinship. In terms of pre-processing, two changes could improve the analysis. First, limiting Google search results to a certain date, for instance, until 2016. Yet, I did not find any parameter for Google's Custom Search API that enables filtering results by date. Second, collect the main texts of the page links collected via Google's API. This would make the text less noisy and could improve the classification.

While these changes are not feasible for this present research, the ensemble approach may help to use this classification to identify dynastic mayors. The concatenation of the 840 manually classified texts and the 4,607 coded with ensemble learning generates a dataset of 5,447 cases. I used this dataset to group texts coding by the municipality. For cities in which no Google results indicated the presence of political dynasties, I coded the mayor as non-dynastic. For the cities that had at least one Google result classified as indicating dynasty, I coded that mayor as dynastic.

As discussed before, one of the drawbacks of using Google search results, such as local blogs posts, to identify dynasties is that the stronger the presence of a dynasty in a given municipality, the greater the chances of having more news and information about it. On the other side, smaller dynasties might not be detected. I try to register this bias by creating a new variable called "Dynasty Consolidation" that sums up all the texts predicted as containing dynastic mention by ensemble learning. This final dataset has 2,329 observations, which is compatible with the number of cities evaluated in the Brazil Transparent Scale.

Finally, to check if the measure based on ensemble learning is fairly accurate for the Brazilian municipalities, I compared the automatized coding with the manual coding of the 119 municipalities discussed at the beginning of this chapter. The level of compatibility between the two measures is 53%. In 64 municipalities they were equally coded, in 14 they were wrongly coded as non-dynastic (false negatives), and 39 mayors were incorrectly coded as dynastic (false positives). This result is aligned with the metrics of the algorithms used in the ensemble classification. It is better than randomly assigning the classes, but it still presents a high number of false positives.

One way to overcome this limitation is using the sum of Google results for each municipality. Since this classification is biased to classify the texts with the positive class even when there is no indication of dynastic kinship (false positives), defining a more rigorous cutoff may reduce the number of false positives. With that in mind, I use the dynasty consolidation variable to calibrate the coding by the number of texts mentioning dynastic ties. I define a new cutoff in which only mayors that had three or more texts classified as containing a dynastic indication will be coded as dynastic politicians.

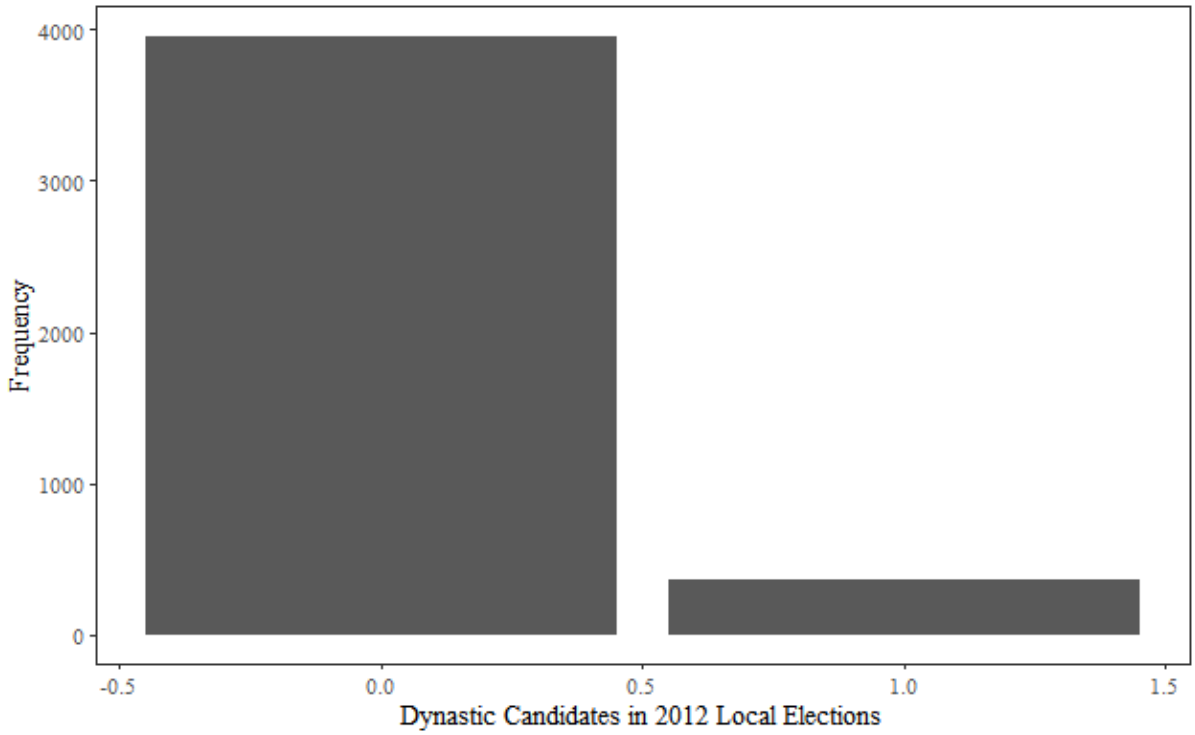
Verifying this calibrated measure with the manually coded sample, the level of compatibility between the two measures increased to 73%. For this calibrated coding, 85 mayors were equally coded. There were 22 false negatives and 10 false positives. The increase in false negatives is probably related to the mentioned bias that may affect small or weaker dynasties for which news are less frequent. Yet, for the purposes of this research, is more important to decrease false positives than false negatives. In this sense, I use this calibrated measure for the entire dataset.

4.3 EXPLORING DATA ON DYNASTIC MAYORS

With a classification of dynastic politicians on hand, I present in this section some data regarding the presence of dynastic mayors elected in 2012 in Brazil. I also compare dynastic

and non-dynastic candidates based on individual characteristics informed by the Super Electoral Court. Figure 12 shows the proportion of dynastic and non-dynastic mayors for the EBT's sample. Only around 9% of the mayors are dynastic.

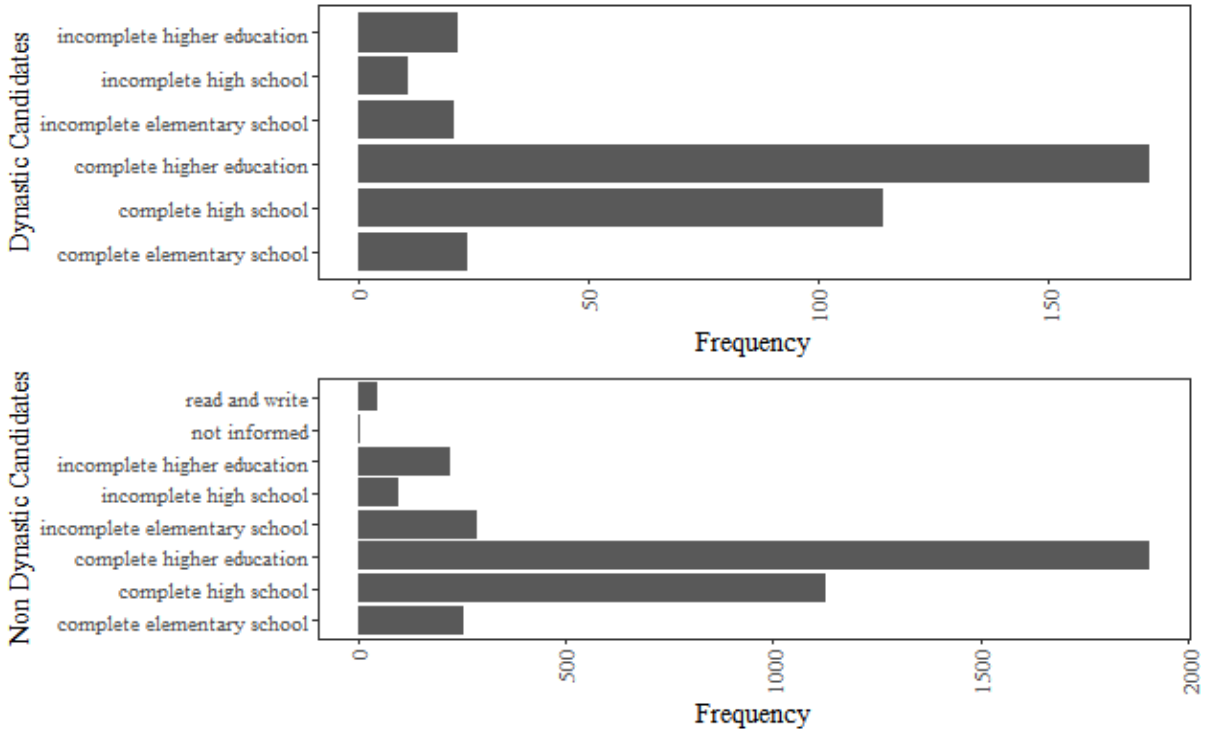
Figure 12 - Frequency of dynastic and non-dynastic candidates elected in 2012.



Source: Source: Superior Electoral Court.

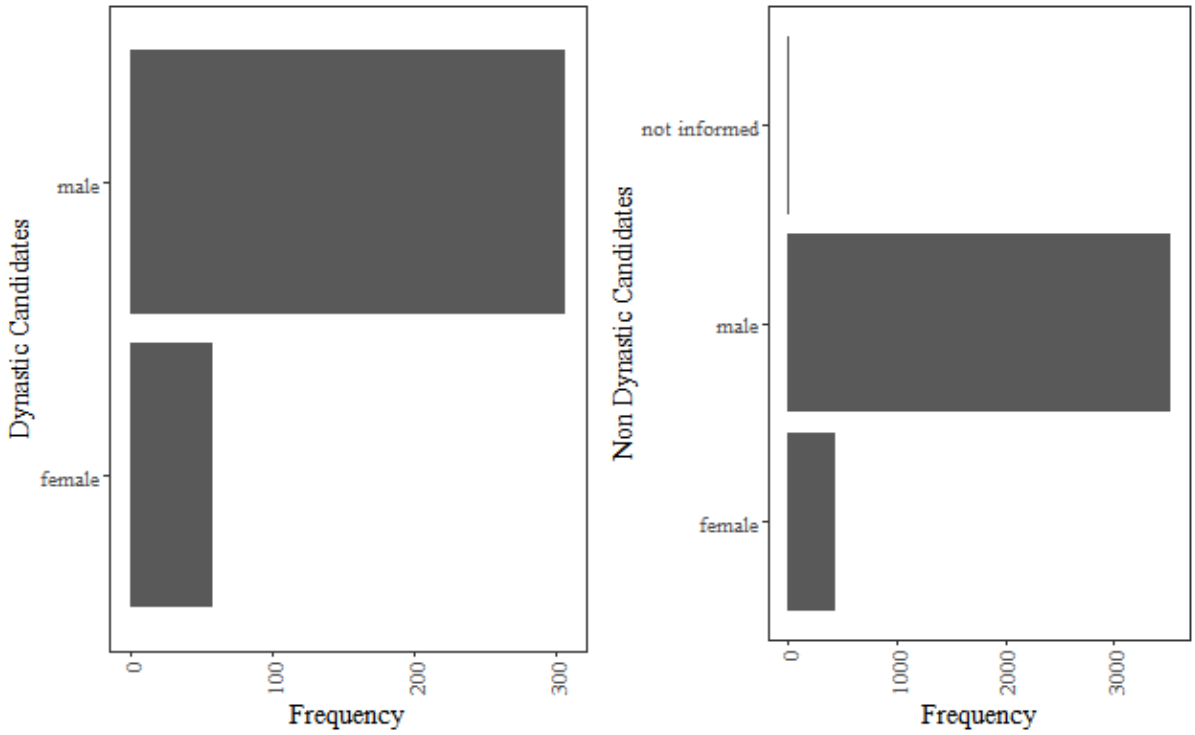
Another interesting aspect is checking whether dynastic politicians vary among observable characteristics, such as education, gender, and occupation. To so, I splitted the sample for dynastic and non-dynastic candidates and compared their distributions regarding these features. As following figures show, dynastic politicians are not very different from the non-dynastic ones. This is interesting because some studies argue that dynastic influence on the political selection process would lead to worse candidates. However, the data on education, for instance, shows that there is little difference between the two groups, though I do not test if the differences are significant. For education, for example, the dynastic group is slightly better than the non-dynastic set.

Figure 13 - Level of education: dynastic and non-dynastic candidates elected in 2012.



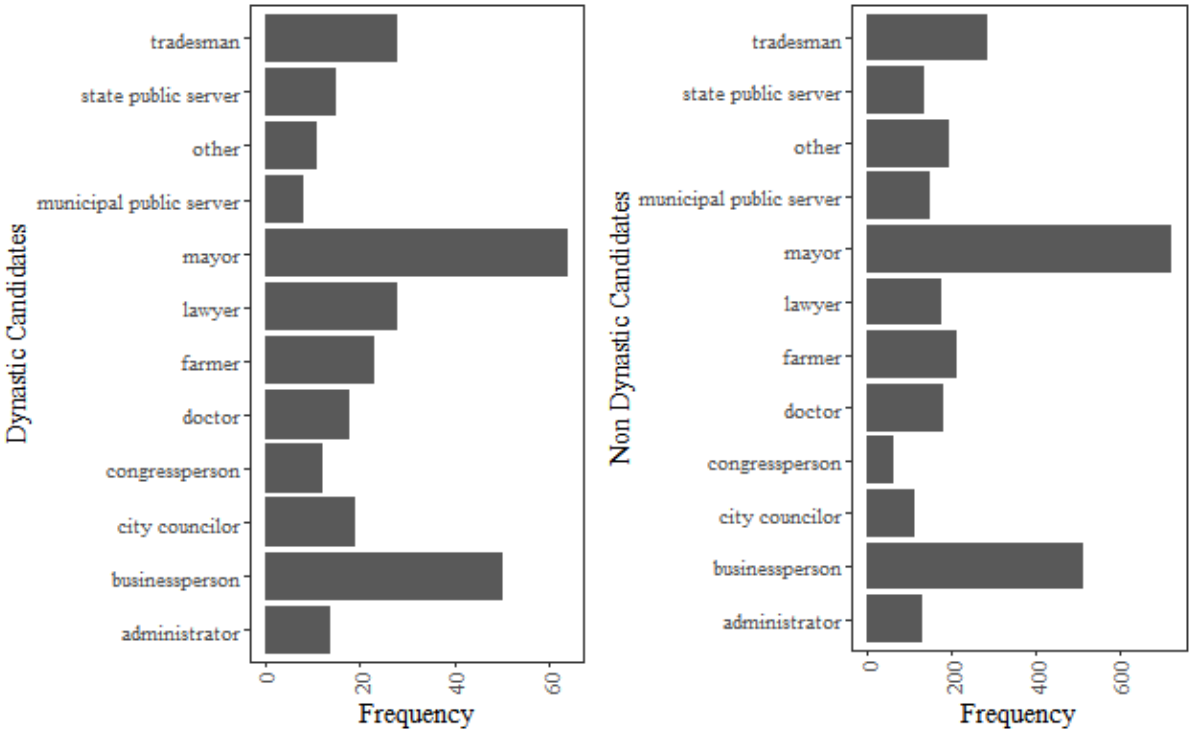
Source: Superior Electoral Court.

Figure 14 - Gender distribution for dynastic and non-dynastic candidates elected in 2012.



Source: Superior Electoral Court.

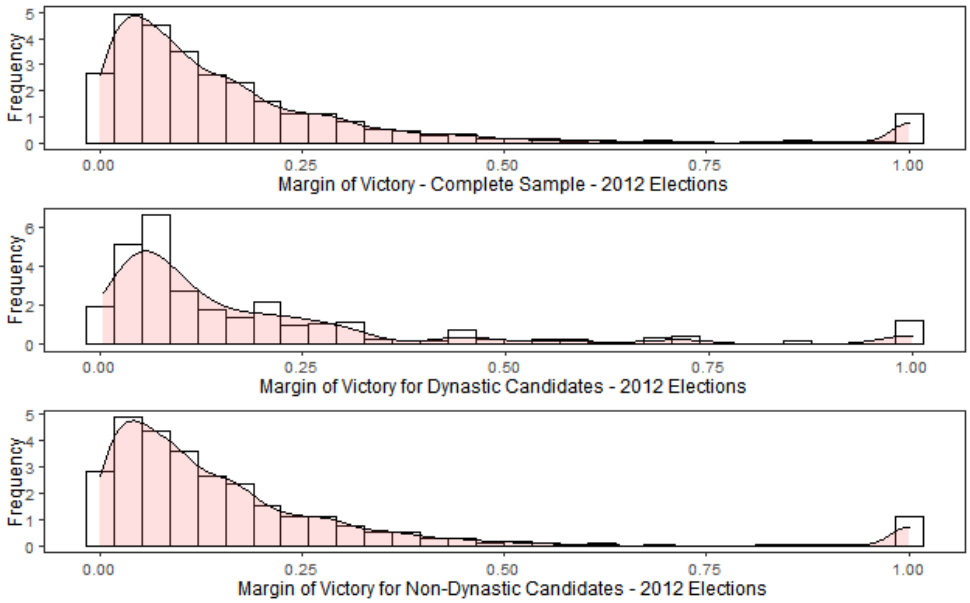
Figure 15 - 10 most frequent occupations: dynastic and non-dynastic candidates elected in 2012.



Source: Superior Electoral Court.

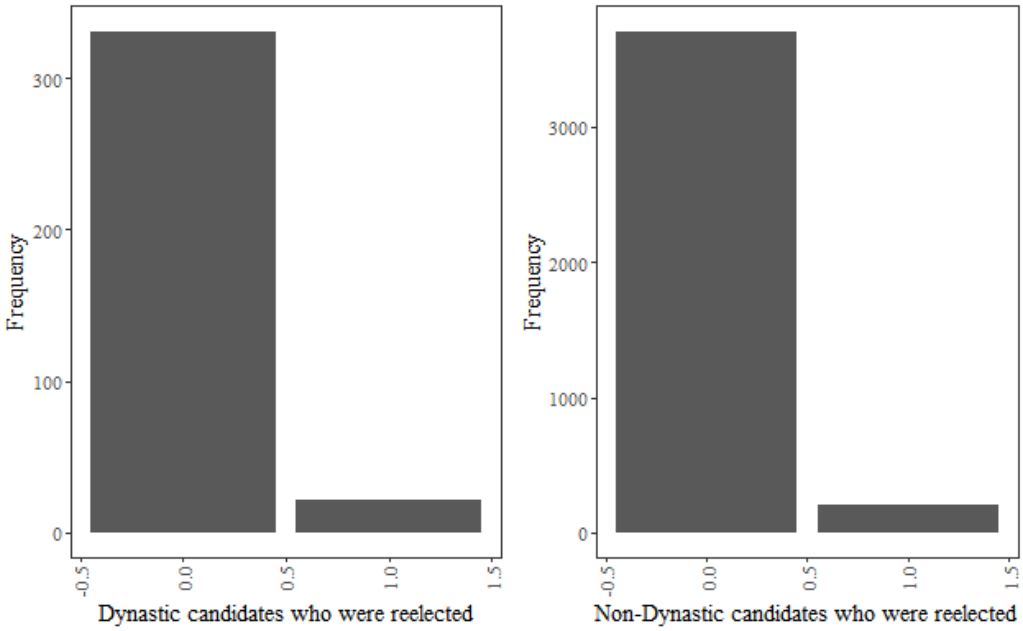
As for the mayors' electoral environment, I also checked whether dynastic mayors were elected in more competitive elections and if they had different rates of electoral success regarding reelection. The results are once again very similar.

Figure 16 - Margin of Victory for Dynastic and Non-Dynastic Elected Mayors in 2012.



Source: Superior Electoral Court.

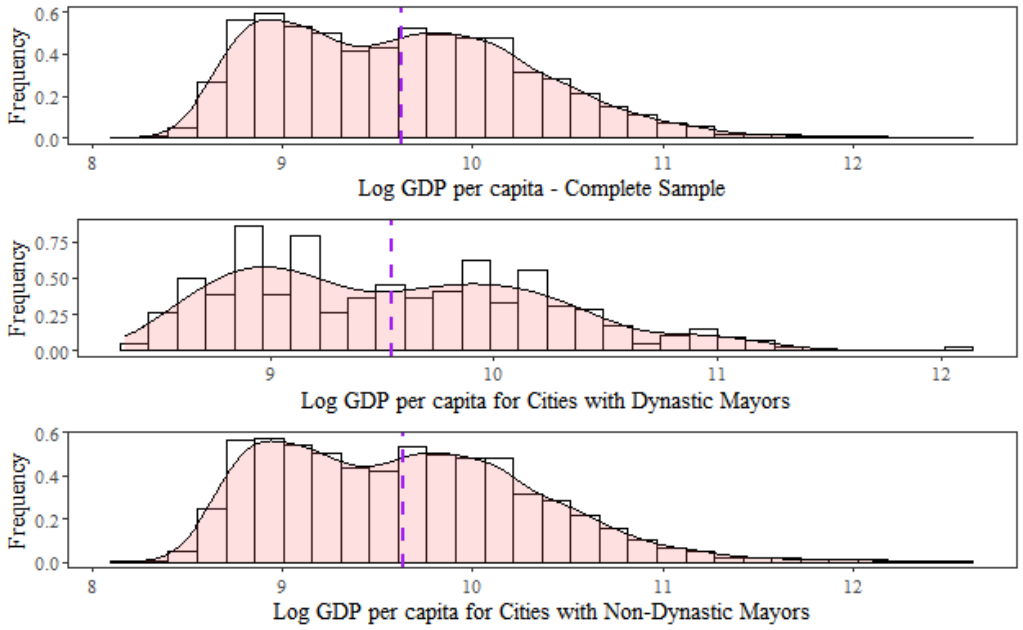
Figure 17 - Reelection for Dynastic and Non-Dynastic Elected Mayors in 2012.



Source: Superior Electoral Court.

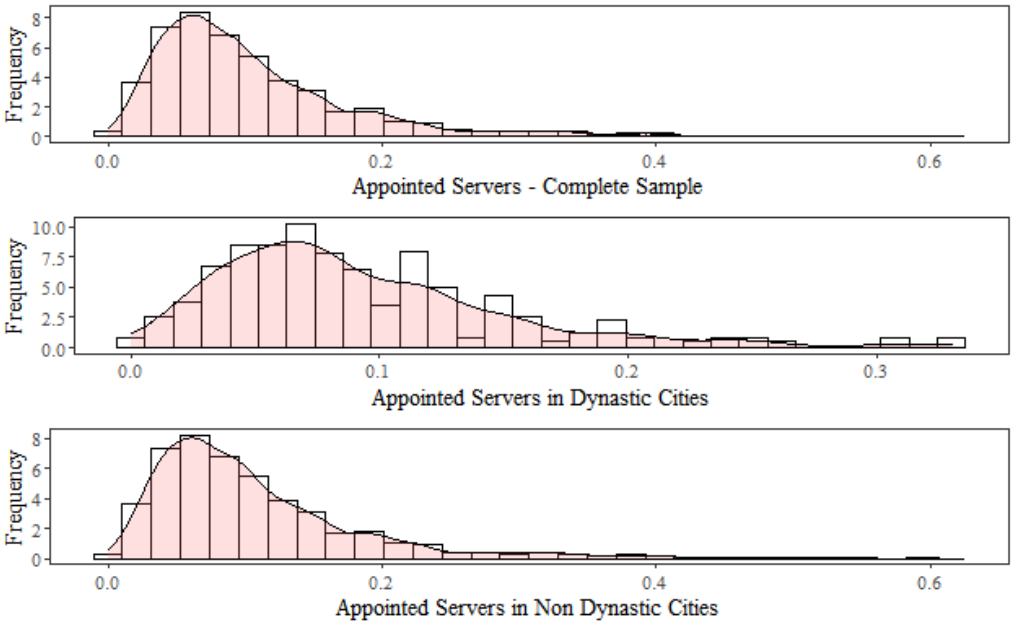
Concerning social aspects, such as the GDP per capita and the proportion of appointed servers (for 2015 and 2016), dynastic mayor's municipalities are quite similar to those ruled by non-dynastic politicians. However, dynastic cities (i.e., cities ruled by dynastic mayors) have more appointed servers in public administration. The dashed purple line indicates the mean.

Figure 18 - GDP per Capita for Dynastic and Non-Dynastic Elected Mayors in 2012.



Source: Brazilian Institute of Geography and Statistics.

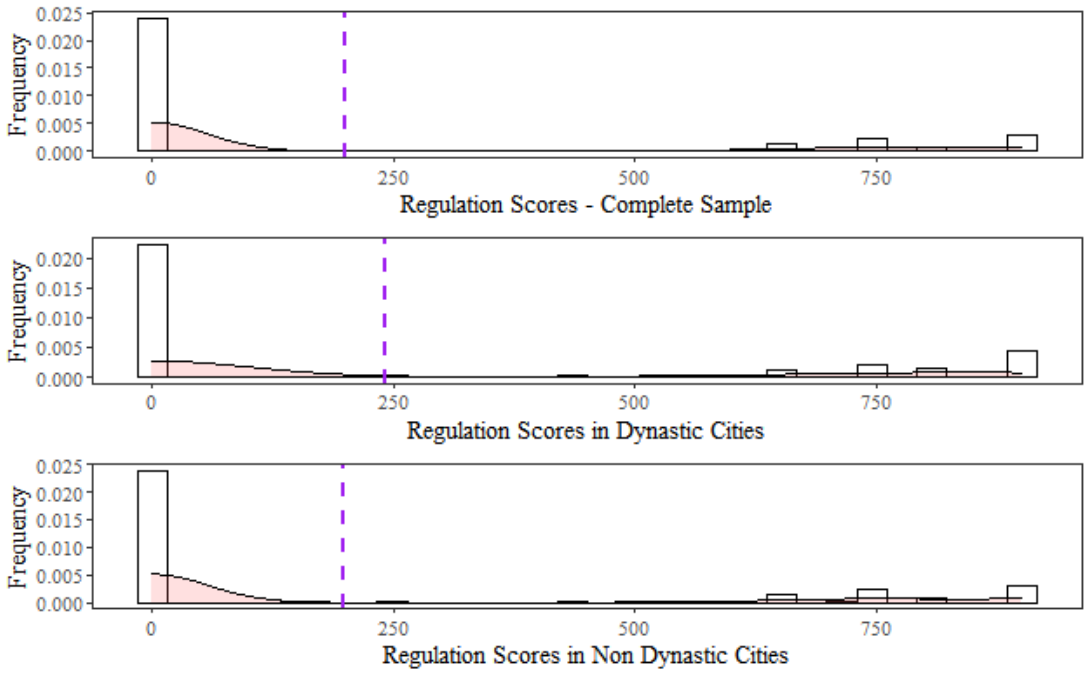
Figure 19 - Proportion of Appointed Servers for Dynastic and Non-Dynastic Elected Mayors in 2012.



Source: Brazilian Institute of Geography and Statistics.

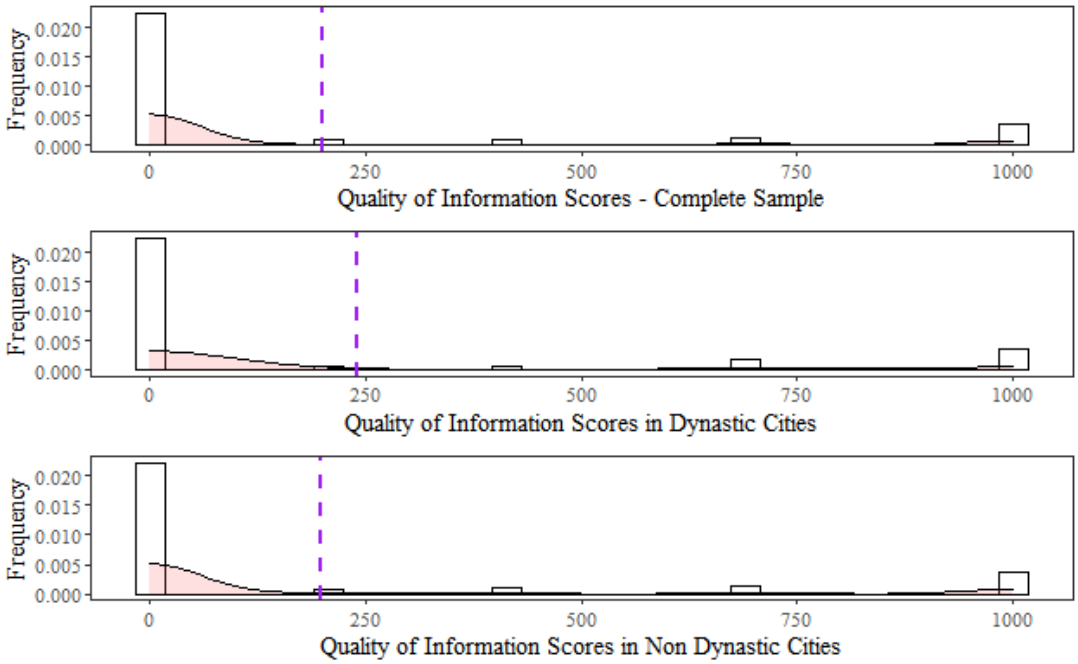
Finally, for EBT scores, the results are, again, quite similar. However, dynastic cities have slightly better means (defined by the purple dashed line in the figures) for all the three proxies of compliance, when compared to the complete sample and the non-dynastic sample.

Figure 20 - Figure 20: EBT's Regulation Score – Dynastic and Non-Dynastic Mayors.



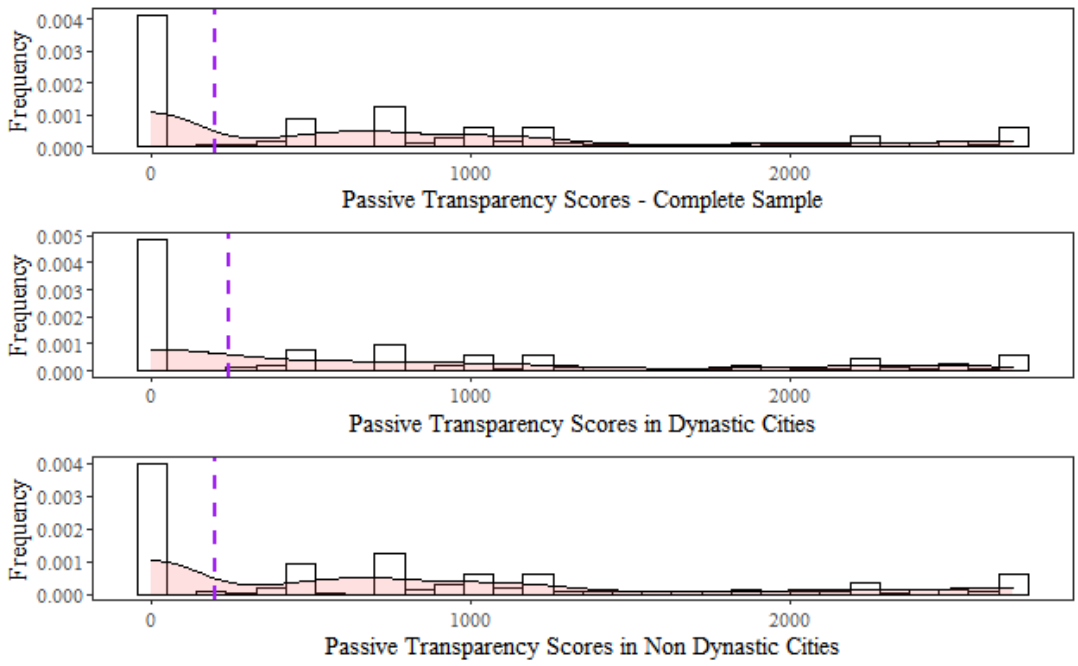
Source: Brazil Transparent Scale.

Figure 21 - EBT's Quality of Information Score – Dynastic and Non-Dynastic Mayors.



Source: Brazil Transparent Scale.

Figure 22 - EBT's Passive Transparency Score – Dynastic and Non-Dynastic Mayors.



Source: Brazil Transparent Scale.

5 COMPLIANCE WITH ACCESS TO INFORMATION IN BRAZILIAN MUNICIPALITIES (2015-2016)

Why do some local governments comply with transparency, while others do not? This is the central question of this dissertation. Although it focuses on "why" compliance with access to information vary among municipalities, my aim is to understand the association between political and social variables and compliance, especially when it comes to political dynasties. In this chapter, I present the methodology and the empirical strategies used to try to answer the research question, as well as the analysis results.

Therefore, this is an observational study that uses data from the *Escala Brasil Transparente (EBT)*, here roughly translated as Brazil Transparent Scale, as *proxy* of *de jure* and *de facto* compliance with the Access to Information Act. The EBT is an evaluation of several aspects regarding how the municipalities have complied with the national law. Before changing their methodology, the *Controladoria Geral da União*, the federal instance responsible for the evaluation, performed three evaluations in 2015 and 2016. They drew representative random samples of Brazilian municipalities. They expanded the samples between the editions, but they kept the municipalities evaluated in previous rounds of EBT. Hence, the dataset generated by the evaluations includes 4,407 ranging more than 2,000 in total (CONTROLADORIA GERAL DA UNIÃO, 2020).

Michener and Nichter (2022) use six measures derived from the EBT as proxies of both *de jure* and *de facto* compliance. I focus my analysis on three proxies extracted from EBT: 1) the regulation score, as a proxy for *de jure* compliance, 2) the score for municipalities' responses to four questions made by CGU, which, as Michener and Nichter (2022) state, can be used to measure of the quality of information given by the local civil servers, and 3) the passive transparency score. The two last measures are proxies of *de facto* compliance.

As for the central argument of this thesis, beyond the measurement of political dynasties, already discussed in the previous chapter, I used the Superior Electoral Court data to measure political variables. Since the evaluations were performed in 2015 and 2016, I used data from the 2012 municipal elections and the 2014 state elections.

From this electoral data, I derive the following variables: margin of victory (i.e., the difference between the proportion or the percentage of votes between the first and second most voted candidates), party identification with the President (a dummy that indicates if mayor and President are from the same party, in this case, PT), party identification with the governor (a dummy to signal whether the mayor and the governor, elected in 2014, are from the same party)

reelected mayor (a dummy for mayors who were reelected in 2012), a dummy that check if the mayor is affiliated to the PT party (as a proxy for left-wing mayors), and a dummy the check mayor affiliation with the PSDB (a proxy for right-wing mayors)[5].

Additionally, other control variables were included in the models to account for the alternative explanations discussed in chapter 2. The data for the following variables come from the Instituto Brasileiro de Geografia e Estatística (IBGE) and range from 2015 to 2016. The variables are the log of the municipality's population, the log of the GDP per capita, and the proportion of civil servers from local public management who were appointed, as a measure of state capacity. The following table present the main descriptive statistics of the continuous variables that were included in the analysis.

Table 6 - Main Descriptive Statistics.

Variable	**Mean**	**Median**	**SD**	**Min**	**Max**
EBT Final Score	2.72	1.94	2.98	0	10
EBT Regulation Score	200.32	0	344.59	0	900
EBT Quality of Information Score	173.27	0	347.73	0	1000
EBT Passive Transparency Score	780.83	700	855.57	0	2700
Margin of Victory (percentage)	19.88	10.65	31.80	0	100
Log GDP per capita	9.63	9.59	0.68	8.24	12.63
Lop Population	9.49	9.38	1.17	6.71	15.68
Proportion Appointed	0.11	0.09	0.08	0.00	0.61

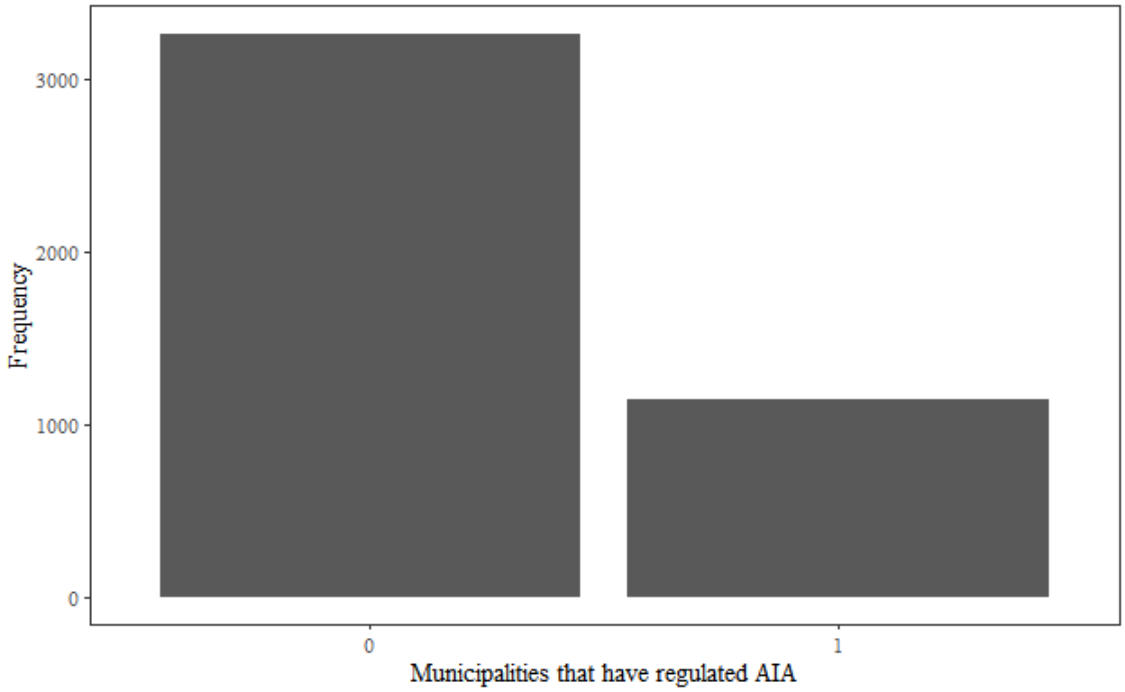
Source: The Author.

Figure 23 demonstrate that, as of 2016, many municipalities had not yet regulated access to information. In total, 1,150 cities have defined their own legal device for AIA, while 1,670 have not regulated it. For 1587 cities, Brazil Transparency Scale displays that they could not find information about regulation available. In this sense, the bar on the left, coded as 0, represents lack of regulation and lack of information about local regulation. The bar on the right, coded as 1, represents the cities that have regulated the law. Figure 24 displays how regulation happened over the years. 2011, the year in which AIA was enacted at the national

[5] This refers to PSDB in 2012, the year of the electoral data used in this analysis.

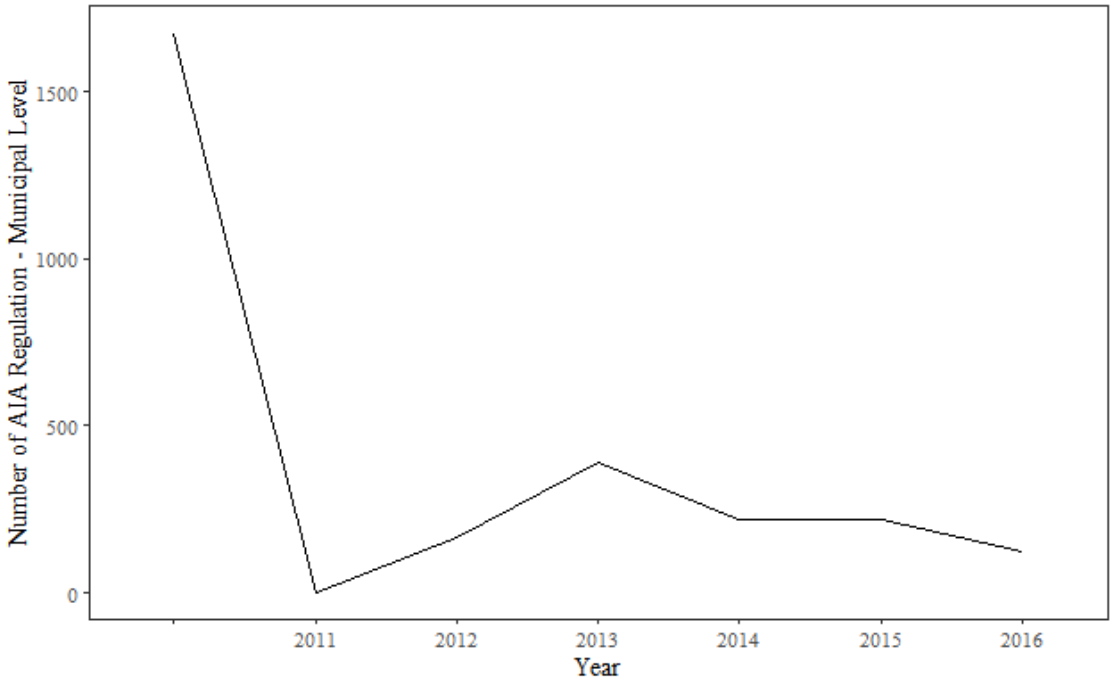
level, contains a greater number of regulations. In 2013 there was a timid increase in regulation and then the numbers started to dwindle again from 2014 on.

Figure 23 - Regulation of AIA at the municipal level (2011-2016).



Source: Brazil Transparent Scale.

Figure 24 - Regulation of AIA at the municipal level (2011-2016).



Source: Brazil Transparent Scale.

A central aspect to define what is the best model to fit the three proxies of the dependent variable, compliance with AIA, is to observe the distribution of these measures. Figure 25, below, presents the frequency of the three proxies of compliance. The dashed blue line indicates the mean. The distribution is highly concentrated in the zero value, for all proxies.

Figure 25 - Histograms of the three proxies of compliance with AIA.

Source: *Escala Brasil Transparente (EBT)*, Brazil Transparent Scale (2017).

According to Lambert's (1992) model, an excessive number of zeros can either be due to chance (false zeros) or it can be what is called a structured zero (true zeros). The first case regards to a group that is susceptible to event of interest, while in the second, observations have no chance of experiencing the event (YANG *et al.*, 2017).

Long (1997) explains this scenario with an example of scientific publication. Comparing the publication chances of different scientists, a regular model will assume that any researcher may publish a paper. However, some individuals may not publish for several reasons, i.e., lack of resources, subject of study do not usually get published and so on. On the other hand, another scientist may have submitted various manuscripts but did not receive an approval. This researcher was then susceptible to experience the event. This two types of zeros indicate that there can be a separated process that generated the excessive amount of zero (LONG, 1997).

For what concerns compliance with AIA, there is no way of knowing if the excessive zeros are structured or generated by chance. For some municipalities, the lack of resources may be so severe that complying with law is simply out of the table. While other municipalities are susceptible to comply but did not do so. Zero-inflated regression allows us to test what parameters influence the odds of experiencing the event for those who are susceptible to it.

In this context, the zero-inflated negative binomial model change its conditional mean and the conditional variance. Hence, it allows the independent modeling of the excessive zeros (LONG, 1997). In this way, "(…) the probability of being in a susceptible group can be estimated by information from covariates with a logistic link." (YANG *et al.*, 2017).

Zero-inflated can be combined with more than one regression model. To define the best approach, is important to observe other aspects of the dependent variable. The first aspect is that, even though EBT is a scale, it does not censor the data because there is no observation that could have greater values than 3600 points, which is the maximum value by definition. This is different from a situation in which you can score as many point as you can, but your final score will be limited to a certain value (i.e., right censoring).

Because EBT's scores sum up several disaggregated scores to from a final value for each evaluation dimension – regulation, quality, and passive transparency – it can be considered a count data. In this case, there are two main possibilities: Poisson and Negative Binomial regression, which can both be combined with a zero-inflated approach (LONG, 1997). Poisson assumes that the conditional mean and the conditional variance are equal. When data is over dispersed, this assumption is violated. In this case, the Negative Binomial regression is preferable, as it allows conditional mean and variance to be distinct (*ibidem*). Zero-inflated distributions are likely over dispersed, and for that reason I chose to use Zero-Inflated Negative Binomial (ZINB) regression to test the main arguments I proposed in this research.

To do so, I use zeroinfl function for the pscl package in R. Zeroinfl is a zero-inflated regression that fits count data using maximum likelihood. This regression in suited to dependent variables that have a large concentration of observations which value equals zero (JACKMAN, 2020). Zero modified count models account for the underprediction of 0 in the Poisson Regression Model. In both models, 0 values are limited to positive counts.

To test the conditional hypothesis presented in chapter 3, I include in the regression interactive terms between dynastic mayor and party identification, and dynastic mayor and margin of victory. To account for possible meaningful differences between the automatized and the manual codification of dynastic politicians, I compare the results for the automatized coding

of dynastic mayors, including all municipalities evaluated in the three editions of the EBT, with the analysis results of manual coding, which was also focused on mayors elected in 2012, but only for municipalities of the first edition of the EBT. Regression tables for manual coding are available in appendix B.

To present and discuss the analysis results, I proceed as follows. First, I present the analysis results for the three specifications of compliance, using as a proxy of dynasties the automated classification. The ZINB regression displays two outputs: a count model and a zero model, and the models do not have to match their results. The former models the count data and regards the negative binomial regression results. The latter models zero-inflated for a logit distribution, adding zero counts to the model (RPUBS). To interpret coefficients as odds ratio, I convert estimates with exponentiation.

The count models allow us to observe what variables are significant to reduce or increase the predicted counts for EBT's scores. The zero model is focused on the municipalities that are not complying but could be doing so (i.e., the susceptible group of municipalities that represent the false zeros. While count models expand our understanding about what may increase or reduce the predicted count of EBT's scores for the cities that already complying with the law in some way, the zero model clarifies what factors may increase the chances that a city start to comply.

For the interaction terms, I present the conditional marginal effects of dynastic mayors for the three interactions included in the analysis. Marginal effects refer to the changes that happen in the predicted probabilities on the dependent variable. For a binary parameter, the marginal effects measure discrete change. For continuous variables, it measures the instantaneous rate of change (WILLIAMS, 2021). To be able to observe the changes in the counts when the independent variable is conditioned to a given term it is necessary to observe the conditional marginal effects, since the coefficient reported in a regression output refers to the effect of the independent variable over the dependent variable when the conditional term is equal 0 (BRAMBOR *et al.,* 2006). Also, the visualization of the conditional marginal effects helps to understand the direction of the interaction, which one cannot fully understand only with the regression results (NORTON *et al.,* 2004).

5.1 ZERO INFLATED NEGATIVE BINOMIAL REGRESSION RESULTS USING AN AUTOMATIZED MEASURE OF POLITICAL DYNASTIES

5.1.1 Results for EBT's Regulation Scores – *De jure* Compliance

Table 7 presents the results for four models for compliance with access to information measured via the EBT regulation score, for count and for zero models. The first model includes the main political variables in the analysis. The second model includes the main explanatory variables related to the argument of the dissertation and the interaction terms among them. I include an interaction between dynastic mayors and margin of victory, an interaction between dynastic mayors and party identification with the governor, and finally an interaction between dynastic mayor and party identification with the President. The third model inserts the control variables. Finally, the fourth model includes all the parameters that presented significance in the previous models (for both count and zero model).

Table 7 - Result for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Regulation Score (Continues).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor	0.020	-0.005		-0.009	Dynastic Mayor	-0.22793*	0.03053		0.01773
	(0.023)	(0.034)		(0.029)		(0.12255)	(0.17863)		(0.15255)
Margin of Victory	0.021	0.018		(0.026)	Margin of Victory	-0.29655*	-0.34320**		-0.22397
	(0.030)	(0.031)		(0.030)		(0.10467)	(0.16398)		(0.16421)
Party Identification with President	-0.014	-0.031		-0.030	Party Identification with President	0.07673	0.09994		0.16806
	(0.021)	(0.022)		(0.022)		(0.11092)	(0.11624)		(0.12006)
Party Identification with Governor	-0.024	-0.025		-0.026	Party Identification with Governor	-0.21972**	-0.09782		-0.11743
	(0.017)	(0.018)		(0.018)		(0.09112)	(0.09667)		(0.10379)
Reelected Mayor	-0.038			-0.036	Reelected Mayor	-0.27342*			-0.22274
	(0.028)			(0.028)		(0.15094)			(0.15758)
Dynastic Mayor:Margin of Victory		-0.012			Dynastic Mayor:Margin of Victory		0.14319		
		(0.103)					(0.53013)		
Dynastic Mayor:Party Id with President		0.140**		0.123*	Dynastic Mayor:Party Id with President		-0.43919		-0.40202
		(0.067)		(0.067)			(0.41380)		(0.42560)

Source: The author.

Table 7 - Result for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Regulation Score (Ends here).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor:Party Id with Governor		0.017		0.018	Dynastic Mayor:Party Id with Governor		-1.25757***		-1.09212***
		(0.050)		(0.050)			(0.31736)		(0.32993)
Log of GDP per capita			0.005	0.005	Log of GDP per capita			-0.70510***	-0.70334***
			(0.010)	(0.010)				(0.05249)	(0.05329)
Log of Population			0.009*	0.008*	Log of Population			-0.19894***	-0.19933***
			(0.005)	(0.005)				(0.02957)	(0.02961)
Proportion of Appointed Civil Servers			0.015		Proportion of Appointed Civil Servers			-0.39281	
			(0.096)					(0.47379)	
Dummy PT			-0.020		Dummy PT			0.12048	
			(0.021)					(0.11273)	
Dummy PSDB			-0.020	-0.018	Dummy PSDB			0.20686*	0.26846**
			(0.022)	(0.022)				(0.11895)	0.12656
Constant	6.644***	6.643***	6.511***	6.520***	Constant	1.18926***	1.15948***	9.81287***	9.83916***
	(0.010)	(0.010)	(0.103)	(0.098)		(0.05105)	(0.05169)	(0.57882)	(0.57318)
Observations	4,257	4,268	4,338	4,257	Observations				
Log Likelihood	-9,522.750	-9,523.828	-9,567.551	-9,385.018	Log Likelihood	-9521 13 DF	-9522 17 Df	-9568 13 Df	-9384 23 Df

Note: *p<0.1; **p<0.05; ***p<0.01

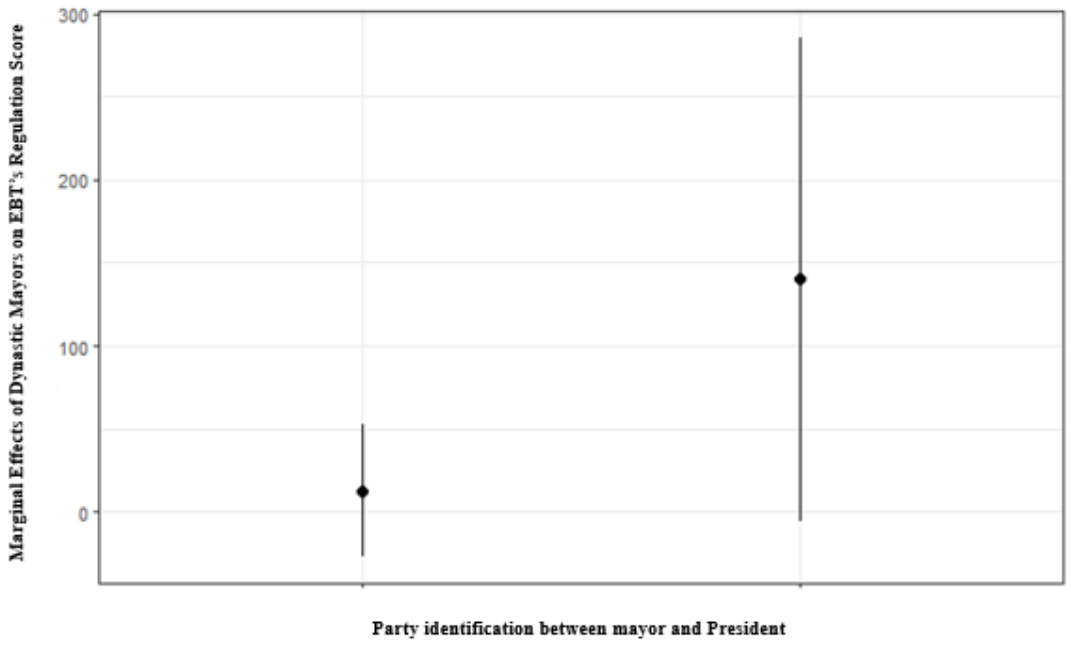
Source: The author.

For the count model, results demonstrate that none of the political variables presented a significant effect. Also, the direction of the association is not consistent across the models since dynastic mayors change the direction from positive to negative. Margin of victory is consistently positively associated with compliance. On the other hand, party identification coefficients, both for President and for Governor, point to a negative association between party alignment and *de jure* compliance. For reelection, it seems that mayors in the second term would increase predicted regulation scores, but the term is not significant.

Among the interaction terms that test the three conditional hypotheses of this research, only party identification between mayor and President is significant for the count model. In model 2 the baseline for EBT regulation score is 6.642918. An increase of one unit in dynastic mayors, when aligned with the President's party, raise the odds of better in regulation of in 1.15 points. According to model 4, where the intercept is 6.7882, the increase is of 1.13 points.

Figure 26 shows the conditional effects of dynastic mayors. Indeed, when the mayor and the President are from the same party, the marginal effect of dynasties on *de jure* compliance is higher. Nonetheless, the interval of confidence for this result is quite large and it almost to touches the 0. Even though the interactive term was significant, this may indicate the conditionally is not strong and consistent.

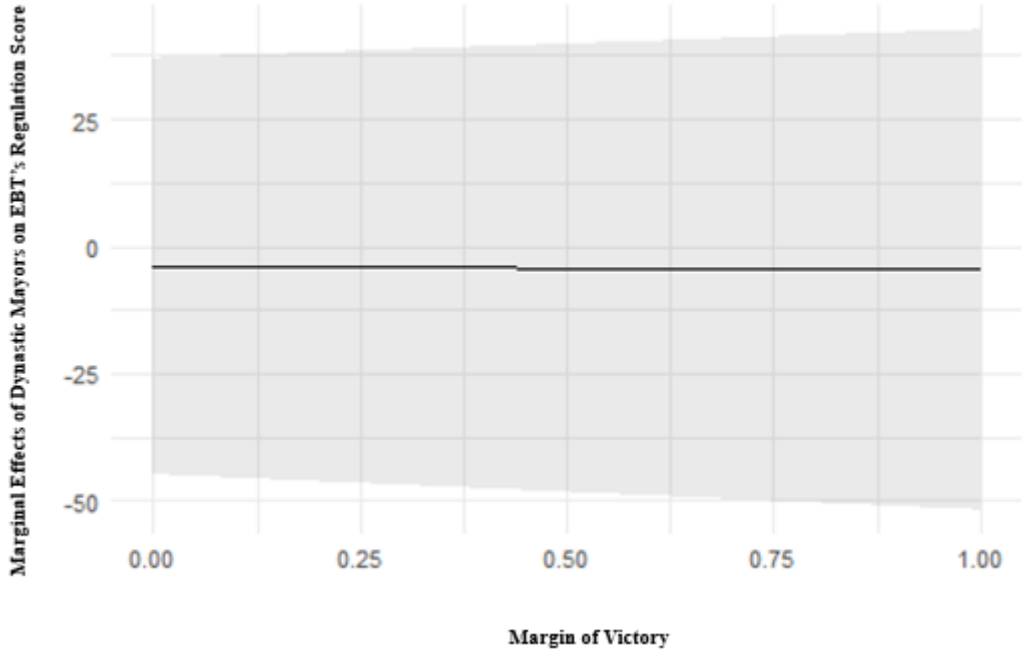
Figure 26 - Marginal effects of dynastic mayors conditional to Party Identification with the President – EBT Regulation.



Source: The author.

The interaction between dynastic politicians and margin of victory is negative. That means that for a one unit change in dynastic mayors, the predicted count of points for the EBT's regulation score is smaller, as margin of victory increases. I expected the opposite: H2 states that as margins of victory increase, the conditional effect of dynastic mayors would be positive. That is, it would raise the predicted counts of points for EBT scores. Figure 27 gives a visualization for the conditional effect, where is possible to see a slight negative line in the predicted counts, as margin of victory increases. Confidence intervals are wide, what is another indicative that this interaction is not significant.

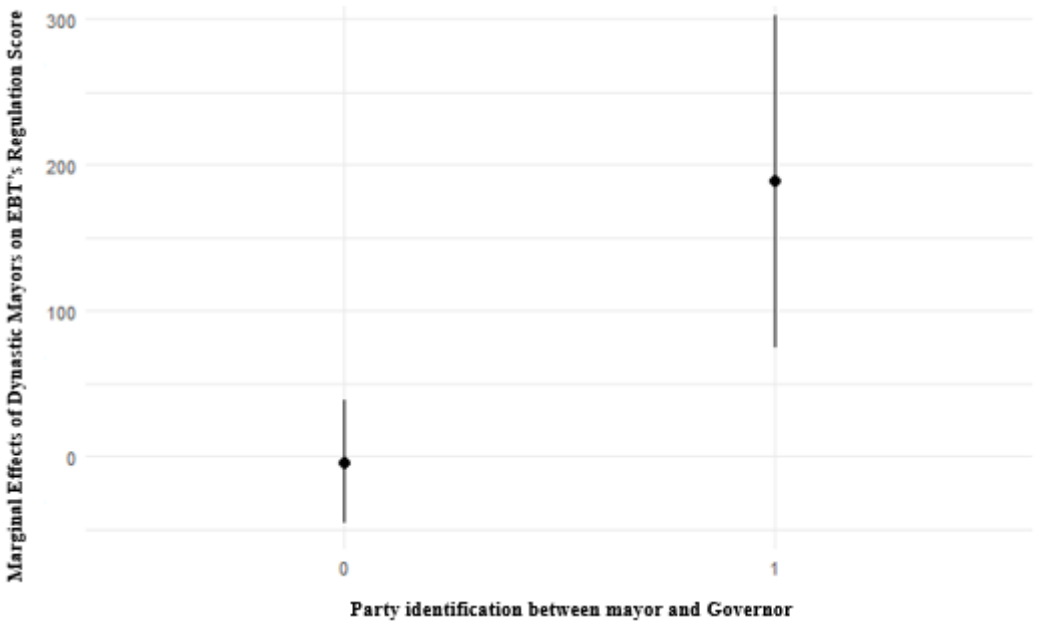
Figure 27 - Marginal effects of dynastic mayors conditional to Margin of Victory – EBT Regulation.



Source: The author.

Figure 28, below, displays the conditional marginal effects for third and last interaction, dynastic mayors who are aligned with the governor's party. According to the count model results, the association is positive, but not significant. Figure 28 demonstrates that only when party identification is not present, the conditional effect is not significant. However, when there is indeed alignment between the parties, the change in the predicted count for the Regulation index for one unit change in dynastic mayors increases (when compared to the absence of alignment).

Figure 28 - Marginal effects of dynastic mayors conditional to Party Identification with the Governor– EBT Regulation.



Source: The author.

For the control variables in the count model, only log of population is statically significant at 0.1. The exponentiated coefficient show that one unit change in the lop of the population generates a change of 1 point in the odds ratio. Log of GDP per capita, the proportion of appointed civil servers and the two dummies related to the ideology of the mayor (whether she is affiliated to PT or PSDB, as proxies for left-wing and right-wing mayors, respectively) are all positively associated with regulation, but are not significant.

For the zero model, as shown in table 7, among the political variables four are significant: dynastic mayors, margin of victory, party identification, and reelected mayors. Dynastic mayors are significant at 0.1 only in the first model. The coefficient is negative, denoting that a one unit change in dynastic mayors will reduced the predicted count for EBT's regulation score. The baseline regulation score for the cities that did not comply with AIA is 3.2846565. One unit change in dynastic mayor reduces the regulation by 0.79 for the cities that have chances to comply with the law. For the margin of victory, a change in one unit of the variable reduces the score to 0.74, according to the first model, and to 0.70, both among the municipalities that are susceptible to compliance. the second model. For party identification, the result of one unit change is a reduction of 0.80 among the cities with chances to comply with the law. Finally, for one unit change in mayors in the second term, one unit change will lead to a 0.76 change is expected, among the cities that have chances to comply with AIA.

As for the interaction effects, the interaction between dynastic mayors and margin of victory, as well as between dynastic mayors and party identification with the president are both non-significant. The interaction between dynastic mayors and party identification with the governor is significant and negative. According to exponentiated coefficients, a change of one unit in dynastic mayors, when there is no party identification (since the conditional term is hold at 0), leads to a 1.25, in the second model, and 1.09, in the fourth model, for the cities that are susceptible to *de jure* compliance.

As discussed, the regression coefficients do not consider the variation in the conditional term, but only the effect of dynastic mayors on regulation when they are hold constant. Unfortunately, I could not only visualize conditional marginal effects for the count models, which seems to be the default in R. I then report information about the interaction terms with the same procedure used for the other variables: significant, direction and, when significant, the exponentiated coefficient of the variable. Consequently, the interpretation of zero models results for interaction must acknowledge this limitation.

Finally, among the control variables, log of the population, log of GDP per capita, and the dummy for PSDB are all significant in two models (model 3 and model 4). Population and GDP per capita affect compliance negatively, while a mayor being affiliated to the PSDB is positively associated with regulation. That is, for the cities who have chances to comply with AIA, the size of the population and GDP per capita may hinder compliance instead of increasing the chances of compliance. For population, a one-unit change leads to a reduction of 0.19, for the third model, and 0.81 for the fourth one. For the log of GDP per capita these numbers are -0.70 and -0.49. For Dummy PSDB, the values 0.20 are 1.30 of increase.

In summary, for EBT's regulation score, which a proxy for *de jure* compliance, the results indicate that dynastic mayors who are either aligned with the party of the President or of the Governor, are associated with higher predicted counts for the regulation score. The expected change is around 1,1 points, which, considering a baseline of approximately 6.5, is a relevant magnitude. The size of the population also matters and can increase the predicted count by 1 point. This is valid for the cities that are already complying with AIA at some level.

For the cities that are not complying with that law but have chances to do so (false zeros), outcome indicates suggests that political variables have more influence in this scenario. Dynastic mayors, margin of victory, party identification with the governor, and mayors in the second term are all significantly and negatively associated with compliance, for the cities that are susceptible to it. Population and GDP per capita also dwindle the predicted counts for EBT's

score. On the other hand, mayors who are affiliated to with PSDB, or right-wing mayors, are positively associated with regulation, for the cities that have not yet complied, but have the chances to so.

5.1.2 Results for EBT's Quality of Information Score – *De facto* Compliance

Table 8 presents the results for the Quality of Information Score.

Table 8 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Quality of Information Score (Continues).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor	0.040	0.042		0.034	Dynastic Mayor	0.06226	0.36346*		0.23827
	(0.057)	(0.085)		(0.063)		(0.13482)	(0.19672)		(0.15340)
Margin of Victory	-0.037	-0.027			Margin of Victory	-0.09485	-0.10615		
	(0.071)	0.074)				(0.16630)	(0.17401)		
Party Identification with President	0.007	-0.002		0.003	Party Identification with President	0.16295	0.25120**		0.29065**
	0.049)	(0.052)		(0.052)		(0.11736)	(0.12403)		(0.12623)
Party Identification with Governor	0.042	0.040		0.041	Party Identification with Governor	-0.21956**	-0.18267*		-0.12933
	(0.039)	(0.041)		(0.039)		(0.09465)	(0.09894)		(0.09778)
Reelected Mayor	0.023			0.022	Reelected Mayor	-0.41896***			-0.40490**
	(0.061)			(0.061)		(0.15227)			(0.15788)
Dynastic Mayor:Margin of Victory		-0.083			Dynastic Mayor:Margin of Victory		-0.47277		
		(0.239)					(0.55016)		
Dynastic Mayor:Party Id with President		0.082		0.127	Dynastic Mayor:Party Id with President		-1.26300***		-1.32357***
		(0.155)		(0.152)			(0.40736)		(0.42356)

Source: The author.

Table 8 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Quality of Information Score (Ends here).

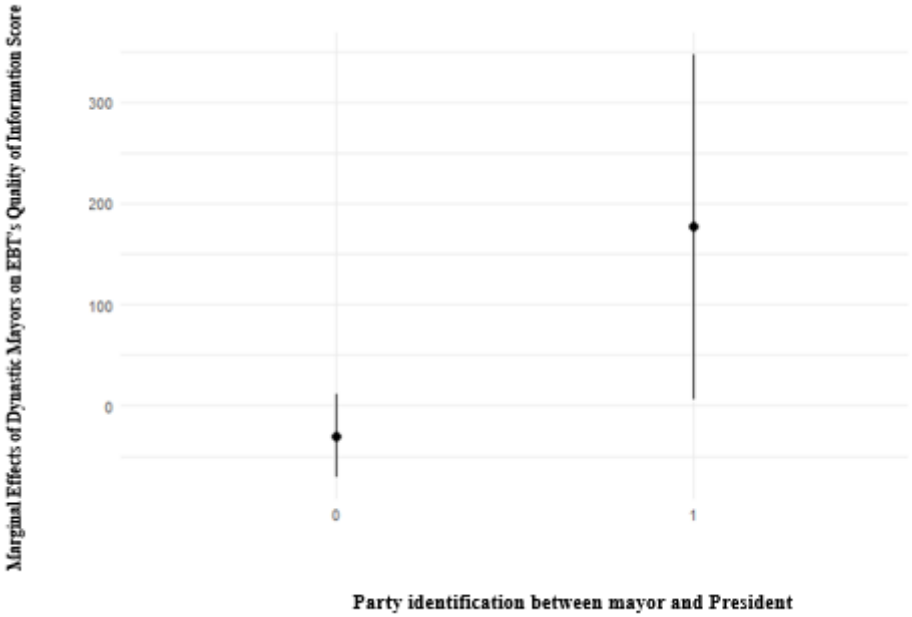
	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor:Party Id with Governor		0.009 (0.137)			Dynastic Mayor:Party Id with Governor		-0.34937 (0.34046)		
Log of GDP per capita			0.051** (0.023)	0.054** (0.023)	Log of GDP per capita			0.050757** (0.023162)	-0.77085*** (0.05529)
Log of Population			-0.026** (0.011)	-0.027** (0.011)	Log of Population			-0.025914** (0.010844)	-0.16807*** 0.03069
Proportion of Appointed Civil Servers			0.029 (0.227)	0.052 (0.227)	Proportion of Appointed Civil Servers			0.028776 (0.226695)	1.30932** (0.53380)
Dummy PT			0.021 (0.049)		Dummy PT			0.020763 (0.049488)	
Dummy PSDB			0.002 (0.047)		Dummy PSDB			0.002049 (0.047479)	
Constant	6.619*** (0.023)	6.620*** (0.023)	6.367*** (0.247)	6.328*** (0.247)	Constant	1.28058*** (0.05287)	1.24567*** (0.05344)	6.366984*** (0.246646)	10.18717*** (0.60583)
Observations	4,257	4,268	4,338	4,243					
Log Likelihood	-9,273.756	-9,274.075	-9,234.743	-9,125.437	Log Likelihood	-9274 on 13 Df	-9274 on 17 Df	-9235 on 13 Df	-9125 on 19 Df
Note:									*p<0.1; **p<0.05; ***p<0.01

Source: The author.

For the count model, only log of population and log of GDP per capita were significant. For a baseline of 582.2 e 559.8 score of quality of information, in model 3 and 4, respectively, one unit change in the size of the population reduces the predicted counts of quality scores in 0.97 in models 3 and 4, a small magnitude. For log of GDP, the coefficients were positive and increased scores in 1.05 also in both models.

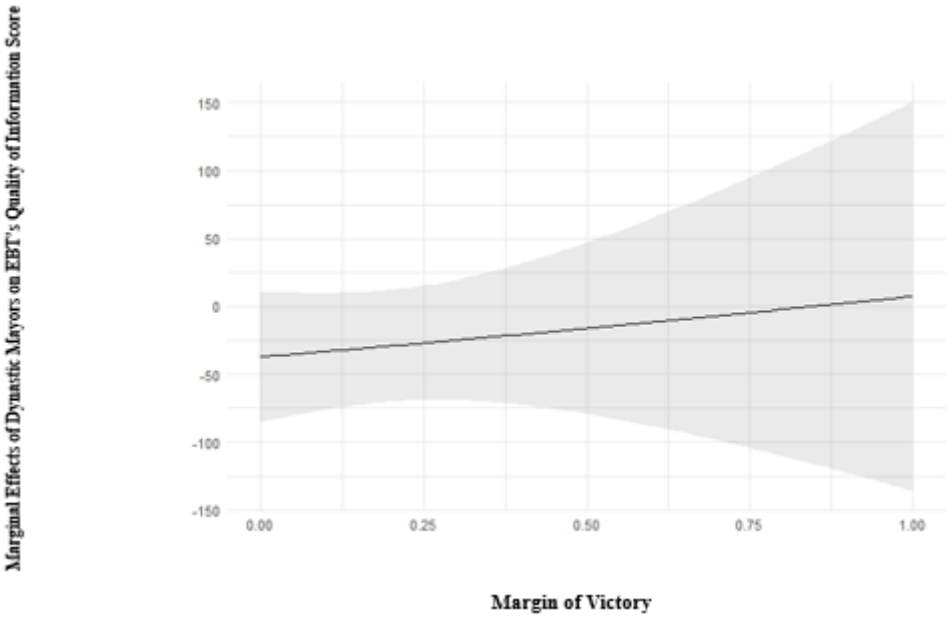
For the interaction terms, Figure 29, below, shows that dynastic mayors' marginal effect on EBT's quality of information is higher when the municipal leader is aligned with the president. For Margin of Victory, Figure 30 demonstrate that could increase marginal effects of dynastic politicians over quality of information, but the interaction is not significant for any value of margin of victory, since the interval of confidence, marked by the gray shade in Figure 30, indicates the 0 is part of the IC, thus, is not significant. For party identification with the governor, dynastic mayor's effect is slightly bigger when there is identification with the governor, but the results are not significant, as seen in Figure 31.

Figure 29 - Marginal effects of dynastic mayors conditional to Party Identification with the President – EBT Quality of Information.



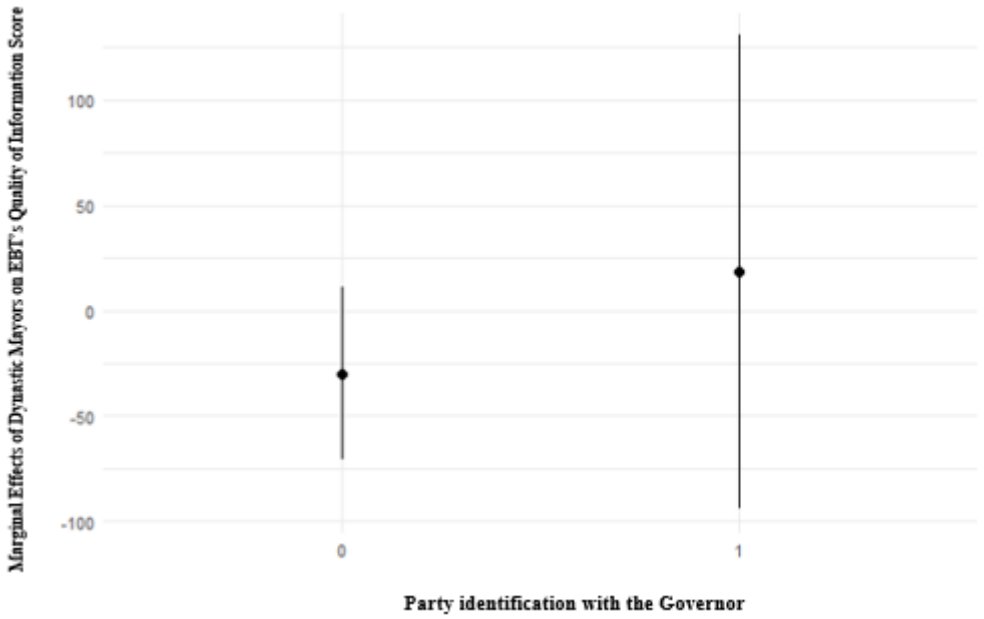
Source: The author.

Figure 30 - Marginal effects of dynastic mayors conditional to Margin of Victory – EBT Quality of Information.



Source: The author.

Figure 31 - Marginal effects of dynastic mayors conditional to Party Identification with the Governor – EBT Quality of Information.



Source: The author.

As for the zero model, once again the political variables are important for the cities that have not complied with AIA but have the chances to comply. Dynastic mayor is significant only for the model 2, with a positive coefficient of magnitude 1.43, for a baseline of 3.47. Margin of Victory is not significant. Party identification with the President is significant , with coefficients 1.28 and 1.33, in models 2 and 4. On the other hand, party identification with the governor is negatively associated with quality of information for the cities that are susceptible to compliance (coefficients 0.70 and0.87). The same happens for mayors in the second term (coefficients 0.65 and 0.66). In this model, dynastic mayors identification with the president's party os negatively associated with compliance. Coefficients are: 0.28 and 0.26. The other interaction terms are not significant.

Also, control variables seem to play a more important role in the case of quality of information: besides log of population and log of GDP per capita, proportion of appointed officials is also relevant, for the first time in the analysis. Its coefficient is positive, indicating the cities with chances of complaining would have greater odds to comply with a less politicized bureaucracy.

5.1.3 Results for EBT's Passive Transparency Scores – *De facto* Compliance

Table 9 presents the results for Passive Transparency Scores.

Table 9 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Passive Transparency Score (Continues).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor	0.054	-0.065		0.025	Dynastic Mayor	0.31920***	0.28072**		0.43219***
	(0.044)	(0.063)		(0.050)		(0.11219)	(0.15547)		(0.12604)
Margin of Victory	0.032	0.017			Margin of Victory	0.13794	0.06625		
	(0.055)	(0.056)				(0.14516)	(0.15271)		
Party Identification with President	-0.005	-0.026		-0.038	Party Identification with President	0.20184**	0.24463**		0.25173**
	(0.036)	(0.037)		(0.038)		(0.09770)	(0.10086)		(0.10271)
Party Identification with Governor	0.051*	0.039		0.036	Party Identification with Governor	-0.19541**	-0.17463**		-0.17506*
	(0.029)	(0.031)		(0.031)		(0.08538)	(0.08881)		(0.09241)
Reelected Mayor	0.127**			0.136***	Reelected Mayor	-0.04697			-0.03673
	(0.052)			(0.052)		(0.14337)			(0.14299)
Dynastic Mayor:Margin of Victory		0.344			Dynastic Mayor:Margin of Victory		0.67666		
		(0.225)					(0.48930)		
Dynastic Mayor:Party Id with President		0.300**		0.278**	Dynastic Mayor:Party Id with President		-0.71820**		-0.70709*
		(0.140)		(0.140)			(0.40594)		(0.40530)

Source: The author.

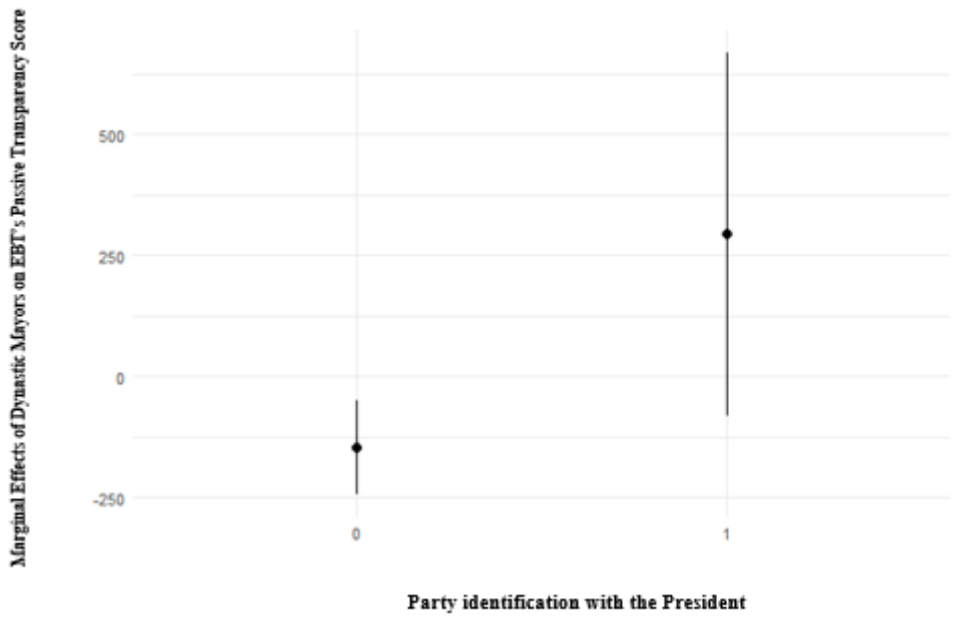
Table 9 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Passive Transparency Score (Ends here).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor:Party Id with Governor		0.124		0.131	Dynastic Mayor:Party Id with Governor		-0.16497		-0.10770
		(0.114)		(0.114)			(0.31460)		(0.31338)
Log of GDP per capita			0.179***		Log of GDP per capita			-0.642629***	
			(0.017)					(0.050337)	
Log of Population			0.026***	0.036***	Log of Population			-0.061631**	-0.08412***
			(0.009)	(0.009)				(0.029045)	(0.02753)
Proportion of Appointed Civil Servants			-0.104		Proportion of Appointed Civil Servants			-0.576690	
			(0.147)					(0.418572)	
Dummy PT			-0.007		Dummy PT			0.171350*	
			(0.035)					(0.098319)	
Dummy PSDB			0.033	0.027	Dummy PSDB			0.009692	-0.04615
			(0.036)	(0.032)				(0.104862)	(0.10792)
Constant	7.113***	7.126***	5.140***	6.776***	Constant	-0.51222***	-0.50879***	6.315369***	0.30116
	(0.017)	(0.017)	(0.180)	(0.084)		(0.04545)	(0.04620)	(0.550141)	(0.26233)
Observations	4,257	4,268	4,338	4,258					
Log Likelihood	-23,537.300	-23,577.710	-23,772.620	-23,525.890	Log Likelihood	-2.354e+04 on 13 Df	-2.358e+04 on 17 Df	2.377e+04 on 13 Df	-2.353e+04 on 19 Df
Note:								*p<0.1; **p<0.05; ***p<0.01	

Source: The author.

For the count model, the significant variables are mayors in the second term, which are positively associated with passive transparency, as shown in models 1 and 4, with coefficients 1.13 and 1.31; party identification with the governor (coefficient of 1.05), the interaction term between dynastic mayor and party identification with the president, with positive coefficients of 1.13 and 1.14. Among the controls, log of the population and log of the GDP per capita are also important. Figure 32 below shows that the interaction between dynastic mayor and identification with the President's party loses significance in the final model, when controlled for other variables. Yet, dynastic mayor's marginal effects on passive transparency is indeed higher with there is identification.

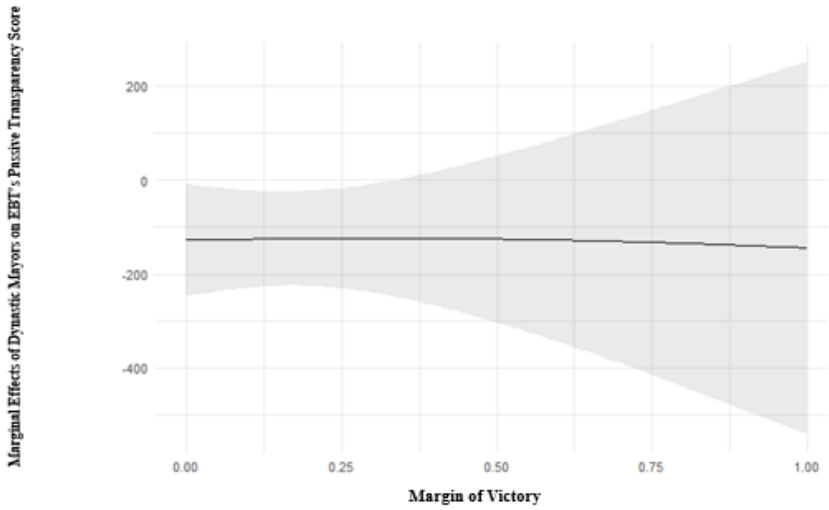
Figure 32 - Marginal effects of dynastic mayors conditional to Party Identification with the President – EBT Passive Transparency.



Source: The author.

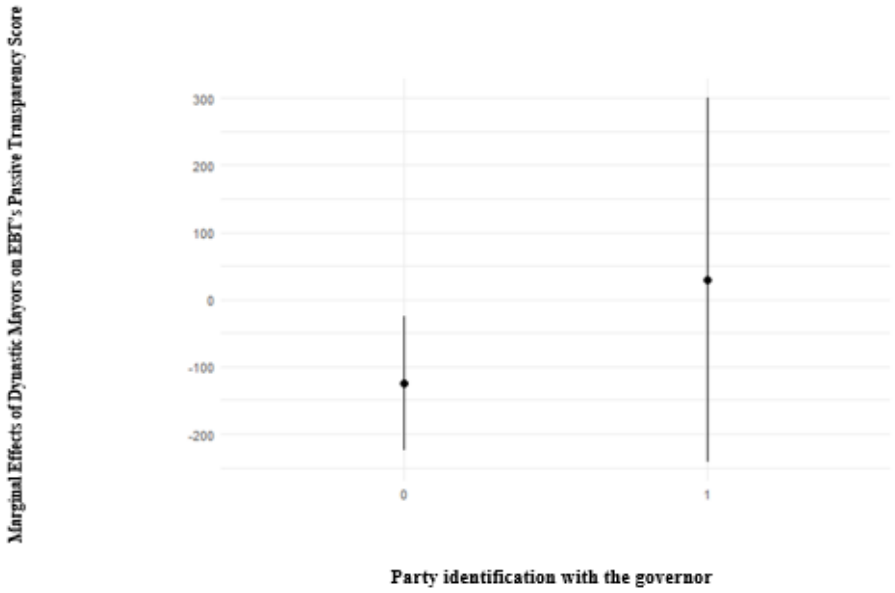
As seen in Figures 33 and 34, interaction terms between dynastic mayors and both margin of victory and party identification with the governor are non-significant for levels of the conditional terms.

Figure 33 - Marginal effects of dynastic mayors conditional to Margin of Victory – EBT Passive Transparency.



Source: The author.

Figure 34 - Marginal effects of dynastic mayors conditional to Party Identification with the Governor – EBT Passive Transparency.



Source: The author.

Concerning the zero model, again political variables are relevant. In this model, dynastic mayor is significantly positive for all the models in which it was included. Coefficients are 1.37 (for an intercept of 0.599), 1.32 (for a baseline of 0.601) and 1.54 (for a baseline of 1.35). Party identification with the President and with the governor are both significant as well, for all models. Variables such as margin of victory and mayors in the second term are not significant. Also, dynastic mayor interaction with party identification with the President is also significant for both models, with coefficients of 0.48 and 0.49. Among the controls, log of population and GDP per capita are significant, as well as the dummy for the mayor being affiliated to the PT. While the first two are negatively associated with passive transparency, Dummy PT association is positive. That means that political and social variables are relevant to influence the chances of a false zero city, that is, a municipalities that is susceptible to compliance turns to effectively comply with the law.

5.2 INTERPRETING THE ANALYSIS RESULTS

Before discussing the results presented so far, it is important to approach a limitation of this data. As discussed in the previous chapter, the automatic coding of dynastic mayors has as a caveat the bias of generating more false positives than true positives. Even with ensemble and calibration, a 70% of compatibility between predicted and manual classification still allow for a great margin of misclassification. As a consequence, misclassification could generate misleading results for the ZINB regressions.

To account for this possibility, I also ran the same models that used automatized coding as the measure of dynastic politicians with the manual coding. However, the measurement is not the only difference: automatized coding includes more than 4,000 cities, while the manual coding was only performed for cities from the first RBT edition, which generated a sample of more than 1,000 municipalities. Is important to have this in mind when observing the differences between the estimations. Tables of regression for manual coding are available in Appendix B.

For the first proxy of compliance, EBT's regulation, modes are very similar for automated and manual coding of political dynasties. For both of them, in the count model the interaction between dynastic mayors and party identification with the president is significant. Direction of the coefficients are also the same. In the zero model, dynastic mayor as significant and with negative effect was the same. Results are also similar in terms of interaction: only

dynastic mayors and party identification with the governor were significant. For quality of information proxy, the only difference is that for the count model in the automated measure of dynasties, dynastic mayor is not significant in any model, while for the manual coding it is significant in model 2.

For the last proxy of transparency, however, results a considerably different for the zero model. While in the automatized measure, dynastic mayor was significant for all the count models, in the analysis with manual classification no coefficient of dynastic mayors was significant. It Is not certain this indicates a problem with the coding, since fort the other two proxies the results were very similar, but it raises a flag indicating that the interpretation for dynastic effects on passive transparency should be cautious.

That being said, what does these results mean for the hypothesis of this research? The first hypothesis I suggested was *that (H1) Dynastic mayors will be negatively associated with local compliance with AIA*. For the three proxies of compliance, having a dynastic mayor was systematically significant for the prediction of counts of EBT's scores. However, for the zero model, that is, for the cities that have not complied with AIA yet, but have the chance of doing, dynastic mayors seem to negatively influence *de jure* compliance (regulation), but it is positively associated with *de facto* compliance (quality of information and passive transparency).

In that way, since results are not very consistent, and results for passive transparency differ from the manual robustness check, it is not possible to reject the null hypothesis. However, the results indicate that dynastic politicians are not that relevant once cities have already started to comply with the law, but political variables seem central for the chances of a city start complying. Future research could focus of exploring in what other situations dynastic politicians may affect transparency.

For the second hypotheses, *(H2) The higher the margin of victory, the higher the conditional marginal effect of dynastic mayors on local compliance with AIA*, for all the models in which the conditional marginal effects were verified, margin of victory did not significantly influence dynastic mayor's effect on compliance, not the direction was consistent. In that way, it is possible to accept to null hypothesis and confirm that margin of victory is not that relevant for dynastic mayors.

For the third and fourth hypothesis - *(H3) If the mayor and the governor are from the same party, the conditional marginal effect of dynastic mayors over local compliance with AIA will be higher* and *(H4) If the mayor and the President are from the same party, the conditional*

marginal effect of dynastic mayors over local compliance with AIA will be higher - the results seem to be supportive in some level. For regulation (*de jure* compliance) and passive transparency (*de facto* compliance), the interaction between dynastic mayors and significant and positive in consistent way. On the other hand, the interaction between dynastic mayors and party identification with the governor is consistently significant and negative.

Dynastic mayors who are aligned with the President are relevant for the cities that have already complied with the law (count model) and may help to understand how to increase compliance. On the other hand, dynastic mayors combined with their alignment with governor, is important for the cities that have not yet complied with the law but have the possibility of doing it (zero model). For the hypothesis, results indicate that alignment with the governor reduces, not increase dynastic mayors effect on transparency scores. On the other hand, results are aligned with the hypothesis, but only for cities that have already started to comply with the law.

Although promising is some aspects, these results have important limitations that can be explored in future studies. The first limitation is the classification of dynastic mayors that seem to have generated inconsistences for estimates for passive transparency. In this, improving the classification method and combing this method with other alternatives such as name matching to arrive at a more rigorous measurement is crucial for this research agenda. Other important point is the research design. For being an observational study, it can only inform about the association between dynastic mayors and compliance. In this way, testing this argument with more rigorous methods that enables causal inference is important.

Also, a qualitative approach would help to understand the differences in dynastic mayors behavior – for instance, the differences in their relationships with the central and the state-level governments. In the next section, I explore some of the cases manually classified in this research. Based on local blog posts, I try to understand important features related to political dynasties and how they organize in power. The goal is to observe whether there are relevant traits in the way dynasties organize themselves and reflect on whether these characteristics are relevant to understand the relationship of dynasties with access to information.

5.3 EXPLORING THE MAIN FEATURES OF POLITICAL DYNASTIES IN THE 119 MANUALLY CODED SAMPLE.

As explained in chapter 3, I manually coded 119 municipalities to understand the type of information I could extract from local blog posts for EBT's evaluation sample. I checked information on the mayor from 2000 to 2016 for each municipality. I also took note of the main local news, including interesting text excerpts. It was an explorative task, not guided by a research question per se. Still, it aimed to collect information about local political dynasties and observe the local news reports' main features and possible patterns.

This process resulted in a detailed yet brief description of each local dynasty. Looking at patterns among the family political groups may shed light on important aspects regarding political dynasties' strategies, which could be later better understood in the development of the research agenda in this field.

There are some interesting variations in the way political dynasties are organized. For instance, in the 58 cities in which political dynasties were present, political groups were generally identified by the group's leader and not necessarily by the parties its members are affiliated with. For instance, in Ararendá, CE, the Mourão family is known as the "Mourão Group" (In Portuguese, *grupo dos Mourão*). News from the *Dinâmica Política: Bastidores do Poder* blog, about the State Deputy (for the PI state) Janaína Marques, says the following:

> The *deputy's group* [Janaína's] is closed and does not give space to those who are not from the family nucleus. In Joca Marques [a city in Piauí state], the current mayor Edilberto Marques Filho is Janaínna's brother. The vice mayor Fernanda Marques is also the deputy's sister. A rare case where the mayor and vice are siblings. They were elected in 2016 in a massacre of the opponent. They had 75.85% of the votes (ALMEIDA, 2020, free translation, emphasis added).

See how the news author refers to the Marques family as *Janaína's group*, even though she is not the first member of the family to be elected. There is an emphasis on the political leader, in this case, Janaína Marques. Related to this is that political parties do not seem as relevant as political leaders in local politics, especially in small cities. This is aligned with the findings of Boas et al. (2018). They draw attention to the fact that partisanship at the municipal level is weak and political dynasties seem to receive the loyalty expected to be directed to parties. Of course, political parties are still significant since they are essential for politicians to enter the electoral competition. However, it may be the case that political dynasties use parties

as a tool to access power, but political agreements and bargains happen inside the political groups.

Hence, members of the same family can run for different parties that are not necessarily allies and still offer mutual political support. For instance, in Baixo do Guandu, a city of the Espírito Santo state, Neto Barros was elected mayor in 2016. His father, Chico Barros, was elected to command the city three times before that. While Neto is described as a communist, being a loyal affiliate of the PC do B (Initials for "Communist Party of Brazil"), Chico is a member of the MDB, the biggest centrist party in Brazil, and took part in the foundation of the party. Chico Barros supported his son's campaign and celebrated his victory (TAVEIRA, 2020, free translation).

Chico Barros' support for his son's campaign is also an example of how dynastic candidates can use kinship as political capital. It seems that candidates who are part of a family with a certain tradition in politics can appeal to a type of legitimacy that regular candidates would have difficulty emulating. The idea is that being raised in a family that has influenced politics gives the new candidate a certain amount of innate experience or a more legitimate motivation to enter politics. This apparent electoral strategy has been well explored by sociological studies that point to how familism uses symbols and collective affective memory in politics (OLIVEIRA et al., 2017). Zheng et al. (2017) present an electoral model of political dynasties in which the electoral success of dynastic candidates is related to how voters evaluate the actions of their relatives in politics. If a regular candidate is evaluated by her actions, a dynastic candidate would also be evaluated by her family's legacy.

Some news reports present excerpts of politicians' discourses that illustrate this idea. Onofre Franco, when launching his daughter, Fabiana, into Joca Marques elections, stated that "Fabiana is strong, is a leader, is a person of the people, and *since she was little, she has been witnessing our fights. She was in politics, even before she knew it.* I don't have a better name; not because she is my daughter, but because I know her strength" (ANDRÉAS, 2020, free translation, emphasis added). On the other hand, Fabiana declared: "*I will honor my father's name* and work for a Joca Marques free because the people want it. It's not us, it's the people's [wish]" (ANDRÉAS, 2020, free translation, emphasis added).

In the city of Lorena, in the state of São Paulo, there is the Marcondes dynasty. The more recent elected member of the family in politics was Fábio Marcondes. His great-grandfather Leopoldo Marcondes was an interventionist mayor, i.e., someone who governed the city before the first election was held, between 1927 and 1929. Benedito de Moura

Marcondes Sobrinho, Fábio's grandfather, was a city councilor several times between 1940 and 1950. Carlos Eugênio Marcondes, Fábio's father, was city councilor and mayor twice. In addition, one of Fábio's brothers, Carlos Eugênio Marcondes Filho, was city councilor between 1997 and 2000 and 2001 and 2004. In a blog post by a former political supporter of Fabio, he says: "(...) I watched a speech by him [Fábio] during the campaign in which, when he took the stage, the population shouted that he was the best mayor Lorena ever had and he corrected the audience: 'I can be the second-best *because the first one is called Carlos Marcondes*'" (LUZ, 2020, free translation, emphases added).

Another example is the case of the Nascimento family, in São Francisco, Sergipe. The family has several former mayors. The first of them was Maria Lúcia, mother of Altair Nascimento and Antônio Nascimento, and aunt of Ailton Nascimento - all former mayors of São Francisco. Maria Lúcia was elected mayor of São Francisco in 1988, having the nephew Ailton as the vice mayor. Ailton was also mayor of the city, from 2001 to 2009. Altair was elected mayor in 2016, but her cousin, Ailton, denounced her campaign for vote-buying to electoral justice. Altair's victory was revoked, and new elections were hold. In the supplementary elections, Ailton's wife, publicly known as "Alba de Ailton", was elected with the support of Altair, even though her cousin had denounced her. At the time, she declared: "The *group I joined* was the one that filed the appeal [for the cancellation of her mandate at the city hall], *but I understood that Ailton is my cousin and family is family. When I passed this onto the electorate, the electorate accepted it*" (PORTAL INFONET, 2019, free translation, emphasis added).

In addition to using political capital, political dynasties also face conflicts that, in some cases, can lead to the rupture of the group into different groups. In other cases, such as the Nascimento family, conflicts can be converted depending on the election - family members can decide it is more strategic to join forces once again. One example of a conflict that led to the suspension of candidacy is the Marques family in Joca Marques (PI). José Aguiar Marques filed a lawsuit in court against his niece, Janaína Marques, after he lost the 2008 elections to her. 8 of the 58 cities that had political dynasties identified in it had news reports indicating ruptures: Abreu e Lima (PE), Ararendá (CE), Barro Alto (GO), Joca Marques (PI), Lagoa Dourada (MG), Maribondo (AL), Santa Cruz (PE), and São Miguel dos Campos (AL). In the latter, the Jatobá family reportedly had three members moving independently in local politics: Rosinha, Pedoca, e Maria Helena Jatobá (POLÍTICA 82).

Another interesting features thar was common among the 58 cities is the fact that some political dynasties were present in more than one municipality. This is the case of São Miguel dos Campos (AL), for instance. Rosinha Jatobá, daughter of Nivaldo Jatobá, was mayor of Jequiá da Praia, which is a city that is situated 56,5 km from São Miguel dos Campos. Maria Helena Jatobá was mayor of Roteiro, located 21 km from São Miguel dos Campos.

In Joca Marques, there is also the presence of a political dynasty in more than one municipality. In this case, it is clear how they benefited from the emancipation of an old district that turned into a new city as an expansion of their electoral market. Joca Marques (PI) was emancipated from the city of Luzilândia (PI) in 1995. Its name was chosen in honor of João de Assis Marques, former mayor of Luzilândia, and Janaína's grandfather. Elections in Joca Marques started in 1996 when Janaína was elected mayor, being reelected in 2000. Posteriorly, Janaína changed to PTB, due to her loyalty to the politician João Vicente Claudino. She was then elected mayor of Luzilândia in 2004 and reelected in 2008. From 1996 to 2016, the Marques family was present in every executive election of Joca Marques.

Other cities benefited from the emancipation of districts that turned into cities. For instance, Iguaba Grande, a city from the state of Rio de Janeiro, was emancipated from São Pedro da Aldeia in 1996. At the time, Rodolfo José Mesquita Pedrosa was the mayor of São Pedro da Aldeia. His son, Rodolfinho, was elected mayor of Iguaba Grande in 2000 (IBGE, 2022; PORTAL IGUABA ONLINE, 2017). The apparent benefits enjoyed by political dynasties from the emancipation of new municipalities could be understood as an expansion of their electoral market. Establishing themselves in new municipalities can be strategic for political dynasties as they can take the opportunity to compete in the new jurisdiction that was created.

One last aspect present across the cases of political dynasties in the cities analyzed is the perpetuation of families in power. In 2012, in São Miguel dos Campos (AL), Nivaldo and his wife, Rosiane Santos, candidacies were revoked by the Electoral Justice after judicial action from Nivaldo's opponent, George Clemente, who alleged that the couple's candidacies for mayor and city council, consecutively, represented perpetuation in power, since Nivaldo have had been mayor of São Miguel for 2 terms and Rosiane for 2 terms as well.

The *Jornal Extra* explained that "(...) The magistrates understood that the former mayor [Rosiane Santos] maintains a relationship in a common-law marriage with her predecessor, Nivaldo Jatobá, which *characterizes the perpetuation in power of a member of the same family* since the mill owners were mayor for two executive terms and his wife in two more" (JORNAL

EXTRA, 2012, free translation, emphasis added). A judge said: “there was the intention to start a family, the intention to deceive the Justice, *the decision to keep a dynasty in power*. The evidence collected proves what was happening in São Miguel was a huge fraud” (JORNAL EXTRA, 2012, free translation, emphasis added).

In 2008, in Iguaba Grande (RJ), Oscar Magalhães was elected mayor of the city. In 2012, the Electoral Justice decided that he should step down. Because of that, Oscar’s daughter-in-law, Grasiella, run in the municipal elections that year. In 2016, Justice considered that Grasiella’s candidacy for reelection was illegal since her election would represent three consecutive terms of the same political group. In 2020, she tried to run for mayor once again, and the MPE (Electoral Public Ministry) filed an Action to Challenge her Campaign Registry (*Ação de Impugnação ao Registro* - AIRC, in Portuguese). Among the motives for the Action, MPE stated that

> (...) her [Grasiella’s] candidacy represents an attempt to exercise a third consecutive term by the same family group since her father-in-law was mayor of the city between 2009 and 2012 and the candidate held the position for two terms since 2013 until being impeached in 2019 (MPRJ, 2020, free translation).

MPE explains that the jurisprudence of the Superior Electoral Court, based on paragraphs 5 and 7 of article 14 of Brazil’s Constitution, decided to prohibit the exercise of a third elective term by the same family group. The idea is to avoid “(...) the perpetuation of family groups in power” (MPRJ, 2020, free translation). MPE referred to the “*Lei de Inelegibilidades*” (the Law of Ineligibility) from May 18, 1990. This legal dispositive establishes, by art. 14, § 9, of the Federal Constitution, cases of ineligibility, termination deadlines, and other measures (BRASIL, 1990).

Two concepts in this law are central to the debate on political dynasties: *inelegibilidade reflexa* (reflex ineligibility) and *prefeitos itinerantes* (itinerant mayors). The former is part of article 14, the 7th paragraph, of the Federal Constitution. This excerpt prevents relatives (either by marriage or by blood) of politicians currently in the Executive (i.e., mayor, governors, and president) from being elected in the respective jurisdiction. This article aims to protect the republican and democratic values of the Federal Constitution by preventing the perpetuation of families in power (JUSBRASIL, 2011). The article is written as follows:

> Art. 14, § 7 - Will be considered ineligible the spouse and relatives by blood or similar, up to the second degree [of kinship] or by adoption, of the President of the Republic, Governor of a State or Territory, Federal District, Mayor or of the person who has

> replaced them within the six months before the election unless they already hold an elective mandate and are a candidate for reelection (BRASIL, 1990).

On the other hand, the itinerant mayor describes the case in which a mayor the has been reelected and served for two terms in a certain city changes her jurisdiction to run for another locality. Politicians' relatives could also be considered itinerant mayors if they were running in the same jurisdiction or in a jurisdiction that was emancipated from the territory in which previous elections took place. This is the understanding of the Electoral Superior Court for a case in Barra de Santo Antônio (AL). Judge Luis Roberto Barroso explains that "(...) TSE jurisprudence is that the spouses and relatives of a reelected mayor are not ineligible for another electoral constituency, even in a neighboring municipality, *as long as it does not result from dismemberment and incorporation or merger* [of the territory]" (TSE, 2019, free translation, emphasis added).

Another example is the case of the Marques family in Joca Marques city. The Regional Eleitoral Court of Piauí (TRE-PI, according to the initials in Portuguese) suspended Janaína Marques's second term on May 4, 2009. She was charged with being part of a group of "itinerant mayors", and remained in office until April 29, 2011, when the Superior Electoral Court confirmed the verdict. News from G1 (2009) explains the situation:

> The mayor's impeachment by TRE-PI had the support of the Electoral Public Ministry in Piauí (MPE-PI), which in May issued a favorable opinion on the decision. In the analysis of regional prosecutor Marco Túlio Lustosa Caminha, Jannaina's re-election was illegal, and her management is the "perpetuation of a family oligarchy in power in the cities of Luzilândia and Joca Marques", criticized the prosecutor, referring to the power of the family Marques in the municipalities. "Both Janaína Marques and her uncle José Aguiar Marques (PSDB) are what there is of most retrograde", he added (G1, 2001, free translation).

Despite the Law of Ineligibility, political dynasties not only continue to exist in Brazil, but new familiar groups keep being formed in politics. Ineligible candidates often indicate their spouse or relative, such as their child, nephews, or mother, to run the elections for them. In Joca Marques (PI), for instance, Onofre Franco was ineligible due to the *Ficha Limpa* law, which forbids the candidacy of individuals convicted by collegiate bodies. He then launched his daughter, Fabiana Franco, as a candidate (ANDRÉAS, 2020). In 2012, in the city of Boa Vista do Gurupi (MA), Walmir de Oliveira's victory was invalid since his registry of candidacy was revoked. A supplementary election was made, and Dilcilene de Melo Oliveira, Walmir's wife, was elected (BLOG PORTAL DO PINDARÉ, 2013).

Why would using relatives to circumvent ineligibility be a rational strategy if there is a legal framework that restricts politicians? Because the application of the law depends on how the judges are going to interpret it. A collection of jurisprudence on reflex ineligibility of the Superior Electoral Court shows that decisions vary from case to case. In some cases, decisions are favorable to the candidates (TSE)[6].

These characteristics are limited to the sample and cannot be taken as a comprehensive representation of political dynasties as a phenomenon in Brazil. Yet, it is useful to systematize important features found in the 119 cases and explore variations among distinct political sceneries. The following table summarizes the main characteristics and strategies of Political Dynasties found in 58 cities, from a total of 119 explored cities.

Table 10 - Main features and strategies of political dynasties in 58 municipalities in Brazil (Continues).

Topics	Description
Political groups	News reports on local politics usually refer to *grupo político* (political groups) when it comes to electoral competition. Political dynasties are nominated with the term "*grupo familiar*", that is, family groups.
The role of political parties	Political groups are associated with their main leader, rather than a political party. Parties are still important, but local political arrangements seem to be led by the leaders of the political groups, and not only or mainly by party leaders.
Transmission of political capital	Aligned with the evidence brought mainly by case studies of sociological research, political dynasties seem to use kinship to transmit and reinforce political capital. Some examples show that a new candidate, when dynastic, can signal her motivation comes from her family legacy, bringing legitimacy, and genuineness to her campaign. Dynastic candidates may also benefit from more credibility, even though they have not been elected before. One example is when Onofre Franco, talking about his daughter, a new candidate, claims that "*since she was little, she has been witnessing our fights. She was in politics, even before she knew it.*"
Conflicts/ruptures	Political dynasties can face conflicts that can either be ephemeral, i.e. relatives can be on opposite sides in one election and reunited in the other, or it can lead to a more definitive rupture when different members of the same family can represent distinct political groups. In the 58 cases in which political dynasties were present, only 8 had information indicating conflicts in the family.

[6] TSE offers a compilation of decisions in this link: <https://temasselecionados.tse.jus.br/temas-selecionados/inelegibilidades-e-condicoes-de-elegibilidade/parte-i-inelegibilidades-e-condicoes-de-elegibilidade/inelegibilidade-reflexa>.

Table 11 - Main features and strategies of political dynasties in 58 municipalities in Brazil (Ends here).

Topics	Description
Expansion of the electoral market	This expansion can happen through two mechanisms: 1) by emancipation when dynasties take advantage of the creation of a new municipality, which means a new jurisdiction to compete; 2) by the itinerant mayors, i.e. distinct members of the same family can compete in different cities and even perpetuate the family in power (e.g.: when the father is elected and reelected, and then the daughter also serve for two terms, leading to four terms of the same family group).
Perpetuation in power	Political dynasties can use distinct strategies to keep themselves in power. Two strategies were highlighted in the 58 cases of dynasties in the 119 cities analyzed: 1) reflex ineligibility; and 2) itinerant mayors. In the former, dynastic politicians launch relatives in politics, either to foment succession or to circumvent legal impediments to their own candidacy. In some cases, the Electoral Justice revokes candidacies based on the concept of reflex ineligibility. In other cases, politicians' strategies can be successful. In what concerns itinerant mayors, altering members of the family in consecutive elections is also a strategy to remain in power.

Source: The author.

These features reveal that political dynasties seem to have specific strategies to remain in power. That is aligned with my argument that political dynasties, at a certain level, have their own way to organize themselves and exert power. Even in a modernized public sphere, with a well-established electoral system, dynastic candidates can use kinship to overcome institutional constraints, such as detailed in Table 11. This tension between the private and the public spheres may be a key to understand why dynastic mayors, opposing initial expectations, are positively associated with transparency and, when aligned with the federal government, significantly increase compliance with AIA.

The reasoning behind a negative association between dynastic politicians and transparency is that these politicians would be more likely to commit malfeasance and, therefore, would be more averse to transparency than non-dynastic politicians. I argued this association holds, but when political dynasties are more consolidated, they will be more prone to enforce long-term policy outcomes and would be more likely to promote transparency because they value legacy and reputation more than regular politicians. Therefore, for political dynasties that are more electorally secure and have alliances with national and state-level

politicians (which I consider a proxy for consolidation), the negative association with transparency would be smaller.

What the results in the previous section show is that, in cities that have already started to comply with AIA, dynastic politicians are in fact positively related to access to information and, when consolidated - i.e., when the dynastic mayors are in the same party of the President - they are significantly associated with an expected increase in compliance *de jure* (the regulation score) and *de facto* (the passive transparency score). However, for cities that have not complied with access to information, dynastic mayors seem to negatively influence that chances of these municipalities starting to comply. Based on these main features of political dynasties, it may be the case that dynastic politicians are resistant to transparency, but once they comply with the law, they do not oppose to it because 1) they might have means to circumvent sanctions; 2) they know to explore the benefits of complying in the relationship with the federal government.

To check if this reasoning makes sense, a qualitative approach will help to identify the perspectives of the dynastic politicians regarding access to information. The literature shows that the politicians' expectations over transparency consequences are important to understand how access to information are implemented. Schnell (2017), for instance, find evidence that politicians implement strong access freedom of information laws because the enact FOI with the expectations of keeping transparency as a window dressing, but public attention pressured for compliance with the law.

For a limitation of time and resources, testing the mechanism that dynastic politicians are actually positively associated with transparency because they value reputation and legacy, and are less worried with access to information consequences is not possible in this dissertation. As a research agenda, the following studies could explore dynastic mayor's attitudes and perspectives on transparency, and their relationship with state-level and federal government qualitatively.

6 CONCLUSION

Throughout this dissertation, I tried to answer the following question: Why do some local governments comply with transparency, while others do not? To do so, I discussed the main findings of the specialized literature regarding two types of compliance: *de jure* and *de facto*. The first refers to activities made by public officials concerning the implementation of legal guidelines for the implementation of access to information. The second type of compliance, *de facto* compliance, regards how legal measures are translated into citizens' lives (MICHENER; NICHTER, 2022).

According to the literature, variables such as margin of victory, reelection, party alignment, per capita GDP, and the size of the populations, among others, may affect compliance with access to information at the municipalities. I argue that the literature has overlooked a variable in many Brazilian municipalities: the political dynasties. Drawing on international and national evidence about the effects of political dynasties on variables such as good governance, policy continuity, and corruption, I suggest that dynastic politicians will be negatively associated with the local compliance with access to information. Additionally, I propose that the association between political dynasties and access to information depends on how the dynasty is inserted into the formal political context.

This includes the level of electoral security and party identification, or alignment dynastic politicians may have with state and national politicians. I state that these aspects are related to the level of consolidation of a political dynasty. Based on previous arguments about how dynastic politicians value reputation and legacy, focusing on long-term returns, I expected that consolidated political dynasties would be more inclined to comply with transparency. In contrast, dynastic politicians from decaying or new dynasties tend to deliver immediate results, focusing on the policy cycle and leaning toward rent-seeking behavior.

I proposed the following hypotheses: *(H1) Dynastic mayors will be negatively associated with local compliance with AIA, (H2) The higher the margin of victory, the higher the conditional marginal effect of dynastic mayors on local compliance with AIA, (H3) If the mayor and the governor are from the same party, the conditional marginal effect of dynastic mayors over local compliance with AIA will be higher, (H4) If the mayor and the President are from the same party, the conditional marginal effect of dynastic mayors over local compliance with AIA will be higher.*

To test this hypothesis, I used three proxies for compliance, all derived from the *Escala Brasil Transparente* (EBT), the Brazil Transparent Scale: the score for the regulation of the Access to Information Act (*Lei de Acesso à Informação*), for the *de jure* compliance, and the score for passive transparency, and another score for the quality of the answers to questions made by CGU to evaluate the municipalities, both scores are related to *de facto* compliance. I also used data from the TSE and the IBGE for the control variables.

Regarding the main independent variable of this research, dynastic politicians, I proposed a new approach to identify dynastic ties between mayors. I suggest using google search results to collect local news about elected mayors and perform supervised learning for binary classification with the text collected from the web pages. To do so, I collected google search results with Google's Custom Search JSON API. I also used Beautiful Soup in Python to scrape the web pages rendered by the API. I then filtered those texts with a function in Python to reduce the number of web pages not related to the main topic of interest, kinship. After cleaning the dataset, I manually classified the Google results, focusing on the municipalities of the first edition of the EBT.

With this classified data, I trained several models until I found the four models with the best fit. These models did not identify the positive class (text indicating dynasty) as expected. This is likely to be a consequence of the nature of the data since I scraped the web pages' entire content, which made the data noisy for the algorithm. Also, binary classification is probably not the best call for this data, which has a class imbalance. 9% of the sample was coded as 1 (indicating dynasty). Using multiclass classification and separating the "non-dynastic" texts into different categories can make the algorithms more accurate.

To circumvent the model's bias toward the majority class, besides using techniques for sampling and model tuning, I used an ensemble approach. I combined the four models to arrive at a final classification. If 3 of the 4 models agreed with the classification, I accepted their prediction. If not, I manually classified the data. I compared automatized coding with manual coding for 119 municipalities extracted from the EBT's sample to validate this classification. After verifying that this classification still generated many false positives, I used a conservative measure: increasing the cutoff of google results indicating dynasties a candidate should have to be considered dynastic. I raised the cutoff from 1 to 3, as this made the measure's compatibility with the validation data increase to more than 70%.

Because the distributions of the EBT scores contain a significant number of zeros, I used zero-inflated negative binominal regression (LONG, 1997). I also included interactive terms to

test the conditional hypothesis. To account for any misclassification results from the dynastic politician's measure, I fitted the zero-inflated regression to two datasets: 1) a dataset containing all the observations from EBT, using automated coding of dynastic politicians; and 2) a dataset containing only the three editions of the municipalities evaluated in EBT 1, hence using manual coding of political dynasties.

Results indicate that dynastic mayors alone do not seem to influence compliance for the cities that have already complied with the law. But when combined with party identification with the President, dynastic mayor's marginal effects over compliance *de jure* (regulation) and *de facto* (passive transparency). Quality of information is not particularly affected by political variables in general. As for the cities that have not yet complied with AIA, dynastic mayor's seemed to reduce the chances that these cities will comply in the future when they are for the same party of the governor. In this way, there is not confirmation for H1, H2 and H3 can be rejected (the first because is consistently not significant and the second because the direction is opposite to the one hypothesized), and h4 has supporting evidence, but it should submit to rigorous inferential testes for confirmation.

Other variables that are important to understand the variation in compliance with access to information are the log of the population, which is consistently significant for all the models if the mayor is in the second term (reelected), which seems to be negatively associated with compliance, diverging from the political insurance mechanism proposed by Berliner (2014), and lop of GDP per capita. Results for margin of victory are mixed, in terms of significance and direction of the association. The proportion of appointed servers did not have any significant coefficient, unless for the quality of information score.

Even though the results do not confirm the main hypotheses of this research, it suggests that political dynasties is a relevant variable to understand compliance with AIA in the municipalities better. Exploring the relationship between local leaders and politicians at the federal level may be a good way of designing initiatives that may generate incentives for mayors to promote access to information. Moreover, these results also indicate that studying political dynasties may help us explain other relevant social and political phenomena, such as good governance, policy outcome, corruption, representation, and so on.

The topics structured in table 11 may indicate some subjects to enter the flourishing research agenda on political dynasties in Brazil. Also, the suggest that dynastic mayors have specific strategies to deal with institutional constraints. It is possible that dynastic politicians are positively associated with transparency, for cities that have already started to comply,

because the not only value more reputation and legacy than regular politicians, but they also be less concerned with sanctions. I do not test this mechanism, but I suggest it to be explored in future research.

Regardless of the limitations of these results, for instance, the observational nature of the study prevent any causal inference or the improvements that need to be made to refine the measurement of political dynasties, this dissertation contributes to the literature on transparency and political dynasties in two main ways. The first is testing a new political variable to understand compliance with transparency at the local level. The second contribution is proposing a novel way of identifying dynastic politicians that can be improved and combined with other strategies, such as name matching, to achieve a more comprehensive and rigorous measure.

As research agenda, studies could further explore the influence of political dynasties on compliance, including other proxies for compliance. Also, it would be interesting to investigate the relationship between local political dynasties and national and state-level politicians and how this can be related to the outcomes that dynastic politicians deliver. Qualitative research could also bring greater knowledge about dynastic politicians attitudes towards access to information and other topics. In-depth interviews with local politicians or other group of individuals with greater knowledge about local politics would help researchers to learn about dynastic politicians perspectives about transparency.

REFERENCES

ABRAMSON, S.; RIVERA, C. V. Political dynasties and party strength: Evidence from Victorian Britain. **Journal of Politics**. v. 78, n. 4, 2016.

ACEMOGLU, D.; JOHNSON, S.; ROBINSON, J. Persistence of power, elites, and institutions. **American Economic Review**. v. 98, p. 267–293, 2008.

ALESSI, G. Cem anos de proteção a Bolsonaro. 08 Aug. 2021. Available at: <https://brasil.elpais.com/brasil/2021-08-08/cem-anos-de-protecao-a-bolsonaro.html>. Accessed in: 22 Jul. 22.

ALMEIDA, G. O PODER DA IRMANDADE. 10 Feb. 2020. Avaliable at: <https://www.politicadinamica.com/colunas/gustavo-almeida/o-poder-da-irmandade1581364848-13988.html>. Accessed in: 12 Jul. 22.

ALT, J. E.; LASSEN, D. The Causes of Fiscal Transparency: Evidence from the U.S. States. **IMF Staff Papers.** v. 53, n. 2, 2006.

AMUNDSEN, I. Democratic dynasties? Internal party democracy in Bangladesh. **Party Politics**. 2016

ANDREÁS, P. Onofre Franco lança a filha Fabiana para concorrer a Prefeitura de Joca Marques. Available at: <https://clicacocais.com.br/noticia/5199/onofre-franco-lanca-filha-fabiana-para-concorrer-eleicoes-em-joca-marques>. Accessed in: 22 Jul. 22.

ANGELO, T. Relembre casos em que o governo decretou sigilo de 100 anos. 14 Apr. 2022. Available at: < https://www.poder360.com.br/governo/relembre-casos-em-que-o-governo-decretou-sigilo-de-100-anos/>. Accessed in: 12 Jul. 22

ARRUDA, L. R. V. de.; KERBAUY, S. T. M. Estudos sobre elites políticas e poder local. **Política Hoje.** v. 25, n. 1, 2016.

ASAKO, Y.; IIDA, T.; MATSUBAYASHI, T.; UEDA, M. Dynastic Politicians: Theory and Evidence from Japan. **Japanese Journal of Political Science**. 2015.

BARRO, R. J. The Control of Politicians: An Economic Model. **Public Choice**. v. 14, p. 19-42, 1973.

BATISTA, M.; ROCHA, V.; ALVES, J. L. A. dos. Transparência, corrupção e má gestão: uma análise dos municípios brasileiros. **Rev. Adm. Pública**. v. 54, n. 5, 2020.

BATISTA, M.; ROCHA, V.; NASCIMENTO, P. Atar as mãos do sucessor ou seguir o vizinho? Difusão do acesso à informação nos municípios brasileiros. **Rev. Adm. Pública**. v. 56, n. 3, 2022.

BATISTA, M. **A difusão da Lei de Acesso à Informação nos municípios brasileiros: fatores internos e externos.** Brasília: Enap, 2017.

BERLINER, D. The Political Origins of Transparency. **Journal of Politics**. v. 76, n. 2, p. 479–491, 2014.

BERLINER, D. Sunlight or Window Dressing? Local Government Compliance with South Africa's Promotion of Access to Information Act. **Governance**. v. 30, 4, p. 641-66, 2017.

BERLINER, D.; ERLICH, A. Competing for Transparency: Political Competi-tion and Institutional Reform in Mexican States. **American Political Science Review**. v. 109, 1, p. 110–128, 2015.

BESLEY, T. Political Selection. **Journal of Economic Perspectives**. v. 19, n. 3, 2005.

BESLEY, T.; REYNAL-QUEROL, M. The logic of hereditary rule: theory and evidence. **Journal of Economic Growth**. 2017.

BLOG DO FINFA. Category Archives: Brejinho. Available at: < https://blogdofinfa.com.br/category/brejinho >. Accessed in: 11 Jul. 22

BOAS, T; HIDALGO, D. MELO, M. A. Norms versus Action: Why Voters Fail to Sanction Malfeasance in Brazil. **American Journal of Political Science**. v. 63, n. 2, p. 385-400, 2018.

BORGES, M. Vote Buying and Impression Management:
Performing Largesse in Northeast Brazil. APSA. July 30, 2018.

BRAGANÇA, A.; FERRAZ, C.; RIOS, J. "Political Dynasties and the Quality of Government". Manuscrito não publicado. 2015.

BRAMBOR, T.; CLARK, W. R.; GOLDER, M. Understanding Interaction Models: Improving Empirical Analyses. **Political Analysis**. p. 63-82, 2006.

BRASIL. Lei de Inelegibilidade – Lei Complementar nº 64, de 18 de maio de 1990. Estabelece, de acordo com o art. 14, § 9º, da Constituição Federal, casos de inelegibilidade, prazos de cessação e determina outras providências.

BRASIL, Lei nº 12.527. Lei nº 12.527, de 18 de Novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º , no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências.

CONTROLADORIA GERAL DA UNIÃO. Escala Brasil Transparente - Transparência Passiva. 2020. Available at: <https://www.gov.br/cgu/pt-br/assuntos/transparencia-publica/escala-brasil-transparente>. Accessed in: 20 sep. 2020.

CONTROLADORIA GERAL DA UNIÃO. Programa Brasil Transparente. Available at: <https://www.gov.br/cgu/pt-br/governo-aberto/a-ogp/planos-de-acao/2o-plano-de-acao-brasileiro/controladoria-geral-da-uniao/programa-brasil-transparente>. Accessed in: 2 Mar 22.

CHHIBBER, P. Dynastic parties: Organization, finance and impact. **Party Politics**. 2013.

CLUBOK, A. B.; WILENSKY, N. M.; BERGHORN, F. J. Family Relationships, Congressional Recruitment, and Political Modernization. **The Journal of Politics**. v. 31, n. 4, 1969.

CROWLEY, G. R., & REECE, W. S. Dynastic Political Privilege and Electoral Accountability: the Case of U.S Governors, 1950 – 2005. **Economic Inquiry**. v. 51, n. 1, p. 735 – 746, 2013.

DAL BÓ, E., DAL BÓ, P.; SNYDER, J. Political dynasties. **Review of Economic Studies**. v. 76, n. 1, p. 115 – 142, 2009.

DANIELE, G; GEYS, B. Organised Crime, Institutions and Political Quality: Empirical Evidence from Italian Municipalities. **The Economic Journal**. v. 125, n. 586, 2015.

DREHMER, A. F.; RAUPP, F. M. Do The Information Platforms of The Branches of Governments of Brazilian States Contribute to Build Passive Transparency? **Rev. Adm. UFSM**, v. 13, n. 1, p. 72-89, 2020.

FEINSTEIN, B. D. The Dynasty Advantage: Family Ties in Congressional Elections. **Legislative Studies Quarterly**. v. 35, n. 4, 2010.

FEREJOHN, J. Incumbent Performance and Electoral Control. **Public Choice**. 1986.

FERRAZ, C.; FINAN, F.; MARTINEZ-BRAVO, M. Political Power, Elite Control, and Long-Run Development: Evidence from Brazil. **NBER Working Papers 27456, National Bureau of Economic Research, Inc**. 2020.

FIGUEIREDO, V. S.; GAZONI, R. L.Transparência nos Sites dos Municípios do Estado do Rio de Janeiro. **Revista de Contabilidade e Controladoria**. v. 8, n. 2, 2016.

FIGUEIREDO, A.; LIMONGI, F. **Relações Executivo Legislativo na Nova Ordem Constitucional**. Editora FGV: Rio de Janjeiro, 2001.

FIVA, J.H.; SMITH, D.M. Political dynasties and the incumbency advantage in party-centered environments. **American Political Science Review**. v. 112, n. 3, p. 706 – 712, 2018.

FOLHA DE SPa. Gestão Bolsonaro muda lei, e assessores poderão impor sigilo de dados do governo. 2019. Available at: <https://www1.folha.uol.com.br/poder/2019/01/gestao-bolsonaro-muda-lei-e-assessores-poderao-impor-sigilo-a-dados-do-governo.shtml>. Acessed in: 11 Jan 22.

FOLHA DE SPb. Para entidades, mudança na Lei de Acesso à Informação é 'deplorável' e um retrocesso. 2019. Available at: < https://www1.folha.uol.com.br/poder/2019/01/para-entidades-mudanca-na-lei-de-acesso-a-informacao-e-deploravel-e-um-retrocesso.shtml >. Acessed in: 11 Jan 22.

FOLHA DE SPc. Após derrota na Câmara, Bolsonaro revoga decreto que mudava Lei de Acesso. 2019. Available at: < https://www1.folha.uol.com.br/poder/2019/02/apos-derrota-na-camara-bolsonaro-revoga-decreto-que-mudava-lei-de-acesso.shtml >. Acessed in: 11 Jan 22.

FOLHA DE SP. 2001. SENADO EM CRISE: Baiano ainda tenta adiar ao máximo decisão sobre seu caso. Available at: < https://www1.folha.uol.com.br/fsp/brasil/fc1705200104.htm >. Acessed in: 9 Mar 20.

FOLKE, O.; RICKNE, J.; SMITH, D. M. **Gender and dynastic political recruitment: theory and evidence**. Unpublished Manuscript. 2016.

G1. ‘Pretendo beneficiar filho meu, sim’, diz Bolsonaro sobre Eduardo em embaixada nos EUA. 2019. Available at: < https://g1.globo.com/globonews/jornal-das-dez/video/pretendo-beneficiar-filho-meu-sim-diz-bolsonaro-sobre-eduardo-em-embaixada-nos-eua-7776393.ghtml>. Acessed in: 11 Jan 22.

GEYS, B. Political dynasties, electoral institutions, and politicians’ human capital. **Economic Journal**. v. 127, 605, p. 474 – 494, 2017.

GEYS, B.; SMITH, D. M. Political Dynasties in Democracies: Causes, Consequences and Remaining Puzzles. **The Economic Journal**. v. 127, p. 446 – 454, 2017.

GOOGLE. Available at: < Custom Search JSON API: Introduction. 2022. https://developers.google.com/custom-search/v1/introduction#background-operations >. Acessed in: 10 Jun 21.

GOTTLIED, J.; KOSEC, K. The Countervailing Effects of Competition on Public Goods Provision: When Bargaining Inefficiencies Lead to Bad Outcomes. **American Political Science Review**, v. 113, n. 1, 2019.

GRIMMER, J.; STEWART, B. M. Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. **Political Analysis**. v. 21, p. 267–297, 2013.

HAGOPIAN, F. “State capitalism and politics in Brazil . 1986. International Studies, University of Notre Dame.” Theses. 1986. Available at: <https://catalog.library.vanderbilt.edu/discovery/fulldisplay/alma991025338819703276/01VAN_INST:vanui>.

HEALD, D. Varieties of transparency. In C. Hood, & D. Heald (Org.), **Transparency: the key to better governance?** Oxford, UK: Oxford University Press. 2006.

HOLLYER, J. R., ROSENDORFF, P. B.; VREELAND, J. R. Measuring transparency. **Political Analysis**. v. 22, p. 413-434, 2014.

HVITFELDT, E. ISLR Tidymodels Labs. Chapter 9: Support Vector Machine. 2021. Available at: <https://emilhvitfeldt.github.io/ISLR-tidymodels-labs/index.html >. Acessed in: 11 Jun 22.

HVITFELDT, E.; SILGE, J. 2021. Supervised Machine Learning for Text Analysis in R. Available at: <https://smltar.com/ >. Acessed in: 11 Jun 22.

IBGE. Cidades e Estados. Accessed in 20 jul. 2020. Available at: <https://www.ibge.gov.br/cidades-e-estados/sp/americo-de-campos.html>. Acessed in: 11 Jun 21.

JACKMAN, S. (2020). pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory. United States Studies Centre, University of Sydney, Sydney, New South Wales, Australia. R package version 1.5.5. Available at: < https://github.com/atahk/pscl/>. Acessed in: 15 Jul 22.

JORNAL EXTRA, 2012. Justiça põe fim à dinastia de Nivaldo Jatobá em São Miguel dos Campos. Available at: <https://novoextra.com.br/noticias/geral/2012/09/3670-justica-poe-fim-a-dinastia-de-nivaldo-jatoba-em-sao-miguel-dos-campos>. Acessed in: 15 Jul 22.

JUSBRASIL, 2011. O que se entende por inelegibilidade reflexa? Available at: <https://lfg.jusbrasil.com.br/noticias/2585945/o-que-se-entende-por-inelegibilidade-relativa-reflexa-denise-cristina-mantovani-cera> Acessed in: 11 Jun 22.

KITSCHELT, H. Linkages Between Citizens and Politicians in Democratic Polities. **Comparative Political Studies**. v. 33, n. 6/7, p. 845 – 879, 2000.

KOSACK, S.; FUNG, A. Does Transparency Improve Governance? **Annual Review of Political Science.** v. 17, p. 65-87, 2014.

KUHN, M.; JOHNSON, K. **Applied Predictive Modeling**. Springer: New York. 2013.

KUHN, M.; WICKHAM, H. Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. 2020. Available at: <https://www.tidymodels.org >. Acessed in: 15 Jul 22.

LABONNE, J.; PARSA, S.; QUERUBIN, P. Political dynasties, term limits and female political empowerment: Evidence from the Philippines. **NBER Papers**. 2019.

LAMBERT, D. Zero-inflated Poisson regression with an application to defects in manufacturing. **Technometrics**. v. 34, p. 1–14, 1992.

LEAL, V. N. **Coronelismo, Enxada e Voto: O município e o regime representativo no Brasil**. Companhia das Letras; Rio de Janeiro. 2012.

LONG, S. J.; FREESE, J. 1997. **Regression Models for Categorical and Limited Dependent Variables**. Sage Publications: London, 1997.

LOWATCHARIN, G.; MENIFLELD, C. E. Determinants of Internet-enabled Transparency at the Local Level: A Study of Midwestern County Web Sites. **State & Local Government Review**. v. 47, n. 2, p. 102-115, 2015.

LUNKES, R; LYRIO, M. V. L.; NUNES, G. S. de F.; GIUSTINA, K. A. D; SILVA, K. ds. Transparência no setor público municipal: uma análise dos portais eletrônicos das capitais brasileiras com base em um instrumento de apoio à decisão. **Revista da Controladoria-Geral da União**. v. 7, n. 10, p. 88-108, 2015.

MEIJER, A. Understanding the Complex Dynamics of Transparency. **Public Administration Review**. v. 73, n. 3, 2013.

MEIJER, A. Government Transparency in Historical Perspective: From the Ancient Regime to Open Data in the Netherlands. **International Journal of Public Administration**. 2015.

MELO, M. A., PEREIRA, C., FIGUEIREDO, C. M. Political and Institutional Checks on Corruption. **Comparative Political Studies**. v. 42, n. 9, p. 1217 – 1244, 2009.

MENDOZA, R. U.; BEJA, L. E.; VENIDA, V. S.; BARUA, D. Y. Inequality in Democracy: Insights from an Empirical Analysis of Political Dynasties in the 15th Philippine Congress. **Philippine Political Science Journal.** v. 33, n. 2, p. 132–145, 2012.

MICHENER, G.; CONTRERAS, E.; NISKIER, I. Da opacidade à transparência? Avaliando a Lei de Acesso à Informação no Brasil cinco anos depois. **Revista de Administração Pública.** v. 52, n. 4, 2018.

MIICHENER, G.; COELHO, J.; MOREIRA, D. Are governments complying with transparency? Findings from 15 years of evaluation. **Government Information Quarterly**. v. 38, n. 2, 2021.

MICHENER, G. FOI laws around the world. **Journal of Democracy.** v. 22, n. 2, p. 145 – 159, 2011.

MICHENER, G. Assessing Freedom of Information in Latin America a Decade Later: Illuminating a Transparency Causal Mechanism. **Latin American Politics and Society.** v. 51, n. 3, 2015a.

MICHENER, G. How Cabinet Size and Legislative Control Shape the Strength of Transparency Laws. **Governance**. v. 28, n. 1, 2015b.

MICHENER, G.; BERSCH, K. Identifying transparency. **Information Polity**. v. 18, n. 3, p. 233-242, 2013.

MICHENER, G.; NICHTER, S. Local compliance with national transparency legislation. **Government Information Quality**. 2022.

MPRJ, 2020. Ministério Público Eleitoral ajuíza ação contra candidata à prefeitura de Iguaba Grande que teve contas rejeitadas quando era prefeita da cidade. Available at: <https://www.mprj.mp.br/home/-/detalhe-noticia/visualizar/93801?p_p_state=maximized>. Accessed in: 12 Apr. 22.

NORTON, E. C.; WANG, H.; AI, C.. Computing Interaction Effects and Standard Errors in Logit and Probit Models. **The Stata Journal**. v. 4, n. 2, p. 154-167, 2004.

OLIVEIRA, R. C., GOULART, M. H. H. S., VANALI, A. C., MONTEIRO, J. M. Família, parentesco, instituições e poder no Brasil: retomada e atualização de uma agenda de pesquisa. **Revista Brasileira de Sociologia**. v. 5, n. 11, p. 165 – 198, 2017.

OLIVEIRA, V. I**nstituições Políticas Brasileiras**. Brasília, DF: Senado Federal. 1999.

OLIVEIRA, R. C. DE. **Na Teia do Nepotismo: sociologia política das relações de parentesco e poder político no Paraná e no Brasil**. 1. ed. Curitiba: Editora Insight. 2012.

OLIVEIRA, S. L. da C. O.; SOUZA, W. P. S. de F. Political Dynasties and Corruption: Evidence from Brazil. ANPEC. 2022.

PERISSINOTTO, R. M. Why is it important to study political elites? **Perspectivas**. v. 53, p. 137-158, 2019.

PORTAL INFONET, 2019. Cassada se une a adversários e elege nova prefeita em São Francisco. Available at: <https://infonet.com.br/noticias/politica/cassada-se-une-a-adversarios-e-elege-nova-prefeita-em-sao-francisco/ >. Accessed in: 12 Apr. 22.

PORTAL TERRA. Available at: < <http://noticias.terra.com.br/brasil/noticias/0,,OI1704517-EI306,00-Lembre+o+escandalo+que+afastou+ACM+do+Senado.html >. Accessed in: 12 Apr. 22.

QUERUBIN, P. Political Reform and Elite Persistence: Term Limits and Political Dynasties in the Philippines. Unpublished manuscript. 2011.

QUERUBIN, P. et al. Family and politics: Dynastic persistence in the philippines. **Quarterly Journal of Political Science**. v. 11, n. 2, p. 151–181, 2016.

RAUPP, F. M.; PINHO, J. A. G. Review of passive transparency in Brazilian city councils. **Revista de Administração**. v. 51, n. 3, p. 288-298, 2016.

Revista Isto é. Arruda é condenado por violar painel do Senado, 11 anos após escândalo Available at: < https://istoe.com.br/231136_ARRUDA+E+CONDENADO+POR+VIOLAR+PAINEL+DO+SENADO+11+ANOS+APOS+ESCANDALO/ >. Accessed in: 12 Apr. 22.

RICHARDSON, L. Beautiful soup documentation. 2007.

ROBERTS, M. E. Introduction to the Virtual Issue: Recent Innovations in Text Analysis for Social Science. **Political Analysis**. v. 24, n. 10, p. 1-5, 2016.

RODRIGUES, K. F. Desvelando o conceito de transparência: seus limites, suas variedades e a criação de uma tipologia. **Cadernos EBAPE.BR**. v. 18, n. 2, p. 237-253, 2019.

ROSSI, M.A. Self-perpetuation of political power. **Economic Journal**. v. 127, n. 605, p. 455 – 473, 2017.

ROUPP, M. D.; PERKINS, N. J.; WHITCOMB, B. W.; SCHISTERMAN, E. F. Youden Index and Optimal Cut-Point Estimated from Observations Affected by a Lower Limit of Detection. **Biom J**. v. 50, n. 3, 2008.

RPUBS. By Rstudio. Models for excess zeros using pscl package (Hurdle and zero-inflated regression models) and their interpretations. Available at: < https://rpubs.com/kaz_yos/pscl-2>. Accessed in 10 Jul. 22.

SAMUELS, D. Does Money Matter? Credible Commitments and Campaign Finance in New Democracies: Theory and Evidence from Brazil. **Comparative Politics**. v. 34, n. 1 , 2001.

SCHNELL, S. Cheap talk or incredible commitment? (Mis)calculating transparency and anti-corruption. **Governance**. v. 31, n. 3, 2017.

SETYANINGRUM, D.; SARAGIH, H. A. Political Dynasties and the Moderating Role of Good Public Governance. **Signifikan: Jurnal Ilmu Ekonomi**. 2019.

SILVA, OW. A. de O.; BRUNI, A. L. Variáveis socioeconômicas determinantes para a transparência pública passiva nos municípios brasileiros. **Rev. Adm. Pública**. 2019.

SMITH, D. M.; MARTIN, S. Political Dynasties and the Selection of Cabinet Ministers. Legislative Studies Quarterly. v. 42, n. 1, 2017.

STIGLITZ, J. E. On Liberty, the Right to Know, and Public Discourse: The Role of Transparency in Public Life. Oxford Amnesty, Lecture
Oxford, U.K. January 27, 1999.

TRIBUNAL SUPERIOR ELEITORAL. Inelegibilidade Reflexa. Available at: < https://temasselecionados.tse.jus.br/temas-selecionados/inelegibilidades-e-condicoes-de-elegibilidade/parte-i-inelegibilidades-e-condicoes-de-elegibilidade/inelegibilidade-reflexa >. Accessed in 14 apr. 2022.

TUSALEM, R. F.; PE-AGUIRRE, J. J. The Effect of Political Dynasties on Effective Democratic Governance: Evidence From the Philippine. **Asian Politics & Policy**. v. 5, n. 3, 2013.

VAN COPPENOLLE, B. Political Dynasties in the UK House of Commons: The Null Effect of Narrow Electoral Selection. **Legislative Studies Quarterly**. v. 42, n. 3, 2017.

WILLIAMS, R. Marginal Effects for Continuous Variables. 2021. Available at: <https://www3.nd.edu/~rwilliam/stats3/Margins02.pdf>. Accessed at Jul. 22.

WILLEMS, E. The Structure of the Brazilian Family. **Social Forces**. v. 31, n. 4, p. 339 –345, 1953.

YANG, S.; HARLOW, L. L.; PUGGIONI, G.; REDDING, C. A. A Comparison of Different Methods of Zero-Inflated Data Analysis and an Application in Health Surveys. **Journal of Modern Applied Statistical Methods**. v. 16, n. 1, 2017.

YOUDEN, W. J. Index for rating diagnostic tests. **American Cancer Society Journals**. v. 3, n. 1, 1950.

ZHENG, B.; BATINTI, A.; GHOSH, S.; WEN, Q. An electoral model of political dynasties. July 1, 2017. Draft.

APPENDIX A – CODEBOOK OF THE MANUAL CLASSIFICATION.

id: Number to identify the municipality

city: Name of the municipality

uf: Acronym of the state

description_dynasty: Brief description of the political context of the city, based on electoral data and news of local political blogs. This is the central information to code municipalities as dynastic and each one of the elected mayors (from 2000 to 2016) as dynastic or not.

The process to search for information of each municipality was the following:

1. Search for name and state of the municipality (e.g., *Toledo PR*) on Google, followed by the term "*lista de prefeitos*" (list of mayors). Usually, pages of Wikipedia contain the list of mayors of the city, which enabled me to check for commonalities among elected mayors' surnames.

2. Search for "name state" followed by the ballot name of the elected mayor for each one of the five local elections. E.g.: *toledo PR beto lunitti*.

3. Choose news which title indicated potential useful information in its content.

4. When few results were generated using the candidate's ballot name, repeat the search but with the candidate's complete name.

5. Read news and identify if there was any indication of kinship among the mayor and any other politician (city council, federal deputy, state deputy, senator, previous mayor, mayor in neighboring cities, and so on).

6. The number of news could vary from 15 to 25 for a municipality. A few cities had truly little information. When information was insufficient, cities received "NA", because it was not possible to check for the presence of political dynasties.

7. Also, in some cases candidates shared the same surname, indicating possible kinship (in cases in which the surname was not that common). However, the news did not mention any kinship. In these cases, there was the only indication of a possible political dynasty. These municipalities were scored as "NC", as in "Not Confirmed".

8. In these cases of suspicion of a dynasty, I checked DivulgaCand "*certidões*" to verify whether possible relatives had the same parents. When this was not the case, there was still the possibility of political dynasties, since they could have other types of kinship.

9. More than bringing additional evidence on dynasties, checking news was important because some mayors were convicted for electoral crimes or misbehavior while in office and then impeached. In this case, I searched for the politician who occupied the seat after the impeachment (either the deputy mayor or the winner of supplementary elections).

Following, more details are given for each variable.

dynastic_city: This variable identifies whether there are one or more dynasties present in the municipalities. In some cases, even though all mayors from 2000 to 2016 had no kinship with politicians, the city was still coded as dynastic because a previous dynasty was present. The same could happen whether the deputy mayor or the mayor elected in a supplementary election was dynastic. However, I *did not* code the city as "dynastic" (= 1) if the mayor elected 2020 had a relative in politics, because that would confuse my analysis, which is restricted to a time frame up to 2016.

I consider as dynastic those who have any kind of kinship (mother/father and daughter/son; uncle/aunts; siblings; cousins, husband, and wife; and "in-law" kinships, which were much less common – the first two kinships were the most frequent ones).

Cities were coded as "NA" when there was no or truly little information to verify whether there was a dynasty in the city or not (which hardly ever was the case).

Cities were coded as "NC" when there was a strong "suspicion" of kinship among two or more politicians, but there was no indication of kinship in the collected news, nor there was available information on DivulgaCand (depending on the election year). There were cases in which I was able to verify that possible relatives did not have the same parents, but this did not exclude the possibility of other types of kinship.

dynasty2000: Code 0 if the elected mayor in 2000 was not dynastic or 1 if s(he) was.

dynasty2004: Code 0 if the elected mayor in 2004 was not dynastic or 1 if s(he) was.

dynasty2008: Code 0 if the elected mayor in 2008 was not dynastic or 1 if s(he) was.

dynasty2012: Code 0 if the elected mayor in 2012 was not dynastic or 1 if s(he) was.

dynasty2016: Code 0 if the elected mayor in 2016 was not dynastic or 1 if s(he) was.

count_dynasties: Continuous variable that sums up the total number of families identified in local politics for each municipality.

split_dynasty: A dummy variable that indicates whether there was available information on any split in the dynasty. For instance, let us say there is a family called "*Silva*" that is present in a given city and they represent one political group. However, a hypothetical politician, called *Carlos Silva*, is a city council and had a disagreement with his brother, whom we will call *Felipe*, who is the city's mayor and is running for reelection. Let us say *Carlos* decided to support the hypothetical candidate of the opposition, *Paula Soares*. I would consider this a case of division in the political group of the Silva family.

When there was no indication of conflict and family members were reported as allies in the news, I coded the variable as 0.

When there was no mention of any kind of alliance or division, I coded it as "NA".

count_dynasticterms: Sums up the total amount of terms (2000-2016) in which the mayor was dynastic, regardless of being of a unique dynasty identified in the city or more than one. For instance: city A could have three mayors from the same family. City B, on the other hand, could have two dynastic mayors: one from family 1 and the other one from family 2. Both cities would score 3 terms.

One way to capture the difference between cities A and B is to consider both "*count_dynasties*" and "*count_dynasticterms*" variables in the analysis.

prop_dynasticterms: Is the proportion of dynastic terms considering a universe of five terms (2000, 2004, 2008, 2012, and 2016 elections). If city A had 2 dynastic terms, then it would have a proportion of 0,6 dynastic terms (3/5 = 0,6).

dynasty_neighbor: Indicates whether a dynasty is present in more than one city, *according to the news collected about that family* (which means I have *not* done any other kind of checking for this variable, such as crossing electoral data, and this information is solely based on the

collected news). For instance, there are three brothers and one of them is a federal deputy, the other is a mayor in city A and the third one is a mayor in city B.

surname_dyn1: Family name for dynasty 1. In some cases, it could be a nickname that was used to identify family members.

surname_dyn2: Family name for dynasty 2, when there was more than 1 dynasty.

surname_dyn3: Family name for dynasty 3, when there were more than 2 dynasties.

kinship_dyn1: Identifies the type of kinship between family members. When there were several family members in politics and various types of kinship, they were separated by "/", e.g.: pai-filho/tio-sobrinho/irmãos/primos/matrimonial.

The types of kinship were the following:

pai-filho(a) or *mãe-filho(a)* = father-son/daughter or mother-son/daughter

tio(a)-sobrinho(a) = uncle/aunt-nephew/niece

irmãos = siblings

primos = cousins

matrimonial = marital

sogro(a)-nora or *sogro(a)-genro* = father/mother-in-law - son/daughter-in-law

*padrinho-afilhado(a)** = godfather-goddaughter/godson

*In baptism, which in Brazil is usually a strong relationship, even though is not an immediate (or blood) one. This is different from what is called "apadrinhamento político", which means to lends one reputation and influence to help or to elect another candidate. Former President Dilma Rousseff, in her first presidential election, for instance, had the blessing and support from then President Lula. This is *not* the case a consider in this dataset. A godfather/godmother is someone responsible for a person and has usually a close relationship with the kid and their parents.

When it was not possible to identify that kind of kinship (where there was only a suspicion of the presence of a dynasty, for instance), the variable was coded as “NA”.

kinship_dyn2: Idem, but for dynasty 2, if there was any.

kinship_dyn3: Idem, but for dynasty 3, if there was any.

nepotism_news: Indicates if there was any news reporting nepotism in any mayor´s terms. For instance, if a mayor has appointed his/her daughter to be the head of the social assistance municipal department.

When there was no mention of nepotism, the variable was coded as NA.

kinship_leg_fedstate: Identifies, *solely based on the collected news*, if there is a family member of the dynasty who is or has been a Federal Deputy, State Deputy, or a Senator.

This variable does not differentiate the dynasties, such as “kinship_dyn2” does. If at least one dynasty has or has had a family member with a seat in the Federal or the State level Legislative, the variable received value “1”.

When there was no indication of members in the federal or state level Legislative, the variable was coded as NA.

name_candidate2000: The name of the elected candidate in 2000 elections.

name_ballot2000: The elected candidate’s name used in the “ballot” displayed in the electronic voting machine in 2000 elections.

cpf_2000: CPF (*Cadastro de Pessoas Físicas*) number of the elected candidate in 2000 elections.

party_2000: Party to which the elected candidate in 2000 elections was affiliated.

name_candidate2004: The name of the elected candidate in 2004 elections.

name_ballot2004: The elected candidate’s name used in the “ballot” displayed in the electronic voting machine in 2004 elections.

cpf_2004: CPF (*Cadastro de Pessoas Físicas*) number of the elected candidate in 2004 elections.

party_2004: Party to which the elected candidate in 2004 elections was affiliated.

name_candidate2008: The name of the elected candidate in the 2008 elections.

name_ballot2008: The elected candidate's name used in the "ballot" displayed in the electronic voting machine in 2008 elections.

cpf_2008: CPF (*Cadastro de Pessoas Físicas*) number of the elected candidate in 2008 elections.

party_2008: Party to which the elected candidate in 2008 elections was affiliated.

name_candidate2012: The name of the elected candidate in the 2012 elections.

name_ballot2012: The elected candidate's name used in the "ballot" displayed in the electronic voting machine in the 2012 elections.

cpf_2012: CPF (*Cadastro de Pessoas Físicas*) number of the elected candidate in 2012 elections.

party_2012: Party to which the elected candidate in 2012 elections was affiliated.

name_candidate2016: The name of the elected candidate in the 2016 elections.

name_ballot2016: The elected candidate's name used in the "ballot" displayed in the electronic voting machine in the 2016 elections.

cpf_2016: CPF (*Cadastro de Pessoas Físicas*) number of the elected candidate in 2016 elections.

party_2016: Party to which the elected candidate in 2016 elections was affiliated.

obs_supplementaryelection: Describes, for the cities in which there was a supplementary election - due to death, resignation, impeachment, or irregularities in the regular election, such as vote-buying - who was the elected mayor.

This did not affect the coding of dynasties for regular elections (2000, 2004, 2008, 2012, and 2016), but it influences the coding of the variable "*dynastic_city*". For example: if a city had a mayor elected in supplementary elections or a deputy mayor who turned into the mayor and this person was a member of a dynasty, then the city can be considered a "dynastic city", which means there is a political dynasty in its politics.

APPENDIX B – REGRESSION RESULTS FOR MANUAL CODIFICATION OF DYNASTIC MAYORS

Table 11 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Regulation Score (Continues).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor	0.015	0.004		-0.014	Dynastic Mayor	-0.49448***	-0.1635		-0.08945
	(0.025)	(0.037)		(0.031)		(0.15537)	(0.2246)		(0.19962)
Margin of Victory	0.022	0.025		0.026	Margin of Victory	-0.15986	-0.2374		-0.15485
	(0.048)	(0.053)		(0.047)		(0.28411)	(0.3200)		0.30374
Party Identification with President	-0.017	-0.056		-0.054	Party Identification with President	0.15542	0.2003		0.29683
	(0.033)	(0.040)		(0.040)		(0.21194)	(0.2447)		(0.25184)
Party Identification with Governor	-0.020	-0.021		-0.023	Party Identification with Governor	-0.64723***	-0.3095		-0.31315
	(0.025)	(0.030)		(0.030)		(0.16346)	(0.1903)		(0.20394)
Reelected Mayor	-0.082**			-0.070*	Reelected Mayor	-1.08021***			-1.14543***
	(0.040)			(0.040)		(0.27618)			(0.29842)
Dynastic Mayor: Margin of Victory		-0.094			Dynastic Mayor: Margin of Victory		-0.1912		
		(0.110)					(0.6765)		

Source: The author.

Table 11 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Regulation Score (Ends here).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor: Party Id with President		0.127*		0.101	Dynastic Mayor:Party Id with President		-0.2118		-0.36608
		(0.073)		(0.073)			(0.5348)		(0.56547)
Dynastic Mayor:Party Id with Governor		0.014		0.012	Dynastic Mayor:Party Id with Governor		-1.4645***		-1.26621***
		(0.054)		(0.053)			(0.4130)		(0.44760)
Log of GDP per capita			-0.009	-0.004	Log of GDP per capita			-0.66801***	-0.62157***
			(0.015)	(0.016)				(0.09548)	(0.09781)
Log of Population			0.010*	0.012**	Log of Population			-0.43132***	-0.42108***
			(0.006)	(0.006)				(0.05500)	(0.05670)
Proportion of Appointed Civil Servers			-0.244		Proportion of Appointed Civil Servers			-0.74323	
			(0.164)					(0.86798)	
Dummy PT			-0.028		Dummy PT			0.03442	
			(0.033)					(0.21617)	
Dummy PSDB			-0.012		Dummy PSDB			-0.21483	-0.05409
			(0.032)					(0.21271)	(0.23084)
Constant	6.668***	6.667***	6.683***	6.593***	Constant	1.56186***	1.4520***	11.91519***	11.45581***
	(0.016)	(0.017)	(0.163)	(0.157)		(0.09864)	(0.1020)	(1.09136)	(1.08953)
Observations	1,416	1,419	1,443	1,416					
Log Likelihood	-2,730.713	-2,730.648	-2,720.738	-2,663.157	Log Likelihood	-2731 on 13 Df	-2731 on 17 Df	-2721 on 13 Df	-2663 on 23 Df
Note:								*p<0.1; **p<0.05; ***p<0.01	

Source: The author.

Table 12 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Quality of Information Score (Continues).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor	0.015	0.062		0.012	Dynastic Mayor	-0.24105	0.01201		-0.02790
	(0.062)	(0.091)		(0.068)		(0.16642)	(0.23814)		(0.19574)
Margin of Victory	-0.001	0.029			Margin of Victory	-0.01958	-0.14939		
	(0.107)	(0.119)				(0.30279)	(0.33770)		
Party Identification with President	0.056	0.075		0.071	Party Identification with President	0.04915	0.30579		0.44473
	(0.081)	(0.100)		(0.100)		(0.04915)	(0.26897)		(0.44473)
Party Identification with Governor	0.098	0.117		0.096	Party Identification with Governor	-0.16224	0.02265		0.02981
	(0.064)	(0.078)		(0.065)		(0.18007)	(0.21048)		(0.18836)
Reelected Mayor	0.155			0.162	Reelected Mayor	-0.69098**			-0.67561**
	(0.099)			(0.100)		(0.29068)			(0.30570)
Dynastic Mayor:Margin of Victory		-0.098			Dynastic Mayor:Margin of Victory		0.03307		
		(0.296)					(0.72150)		
Dynastic Mayor:Party Id with President		-0.038		0.023	Dynastic Mayor:Party Id with President		-1.07413**		-1.58643***
		(0.177)		(0.174)			(0.52291)		(0.54217)

Source: The author.

Table 12 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Quality of Information Score (Ends here).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor:Party Id with Governor		-0.045			Dynastic Mayor:Party Id with Governor		-0.78878*		
		(0.150)					(0.42594)		
Log of GDP per capita			-0.048	-0.046	Log of GDP per capita			-0.93155***	-0.93311***
			(0.039)	(0.039)				(0.10244)	(0.10445)
Log of Population			-0.025*	-0.025*	Log of Population			-0.26382***	-0.24434***
			(0.015)	(0.015)				(0.05175)	(0.05291)
Proportion of Appointed Civil Servers			-0.169	-0.110	Proportion of Appointed Civil Servers			0.12239	0.16211
			(0.366)	(0.359)				(0.95773)	(0.97185)
Dummy PT			0.093		Dummy PT			0.12257	
			(0.082)					(0.22768)	
Dummy PSDB			0.007		Dummy PSDB			0.26871	
			(0.087)					(0.24148)	
Constant	6.634***	6.633***	7.398***	7.346***	Constant	1.56175***	1.49805***	12.98455***	12.85472***
	(0.038)	(0.041)	(0.408)	(0.407)		(0.10060)	(0.10493)	(1.13460)	(1.14490)
Observations	1,416	1,419	1,443	1,413					
Log Likelihood	-2,608.279	-2,607.498	-2,575.507	-2,542.804	Log Likelihood	-2608 on 13 Df	-2607 on 17 Df	-2576 on 13 Df	-2543 on 19 Df
Note:								*p<0.1; **p<0.05; ***p<0.01	

Source: The author.

Table 13 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Passive Transparency Score (Continues).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor	0.118**	0.005		0.044	Dynastic Mayor	0.01399	-0.08728		0.12391
	(0.055)	(0.081)		(0.059)		(0.13672)	(0.18942)		(0.14608)
Margin of Victory	0.140	0.103			Margin of Victory	0.42151*	0.21890		
	(0.102)	(0.108)				(0.24208)	(0.26768)		
Party Identification with President	0.011	-0.058		-0.054	Party Identification with President	0.17847	0.23458		0.15201
	(0.069)	(0.077)		(0.076)		(0.17561)	(0.19279)		(0.18567)
Party Identification with Governor	0.042	-0.004		-0.084	Party Identification with Governor	-0.41226***	-0.33125**		
	(0.054)	(0.060)		(0.030)		(0.14753)	(0.16223)		
Reelected Mayor	0.382***			0.418***	Reelected Mayor	0.09267			0.11334
	(0.111)			(0.110)		(0.26745)			(0.26597)
Dynastic Mayor:Margin of Victory		0.366			Dynastic Mayor:Margin of Victory		1.17219*		
		(0.324)					(0.64055)		
Dynastic Mayor:Party Id with President		0.304*		0.367**	Dynastic Mayor:Party Id with President		-0.22887		-0.43792
		(0.178)		(0.172)			(0.48041)		(0.45489)

Source: The author.

Table 13 - Results for Zero-Inflated Negative Binomial (ZINB) Regression - EBT's Passive Transparency Score (Ends here).

	Count Model					Zero Model			
	(1)	(2)	(3)	(4)		(1)	(2)	(3)	(4)
Dynastic Mayor:Party Id with Governor		0.156			Dynastic Mayor:Party Id with Governor		-0.55224		
		(0.140)					(0.40140)		
Log of GDP per capita			0.159***		Log of GDP per capita			-0.72664***	
			(0.032)					(0.08365)	
Log of Population			0.045***	0.046***	Log of Population			-0.11916**	-0.11913***
			(0.015)	(0.014)				(0.04841)	(0.04542)
Proportion of Appointed Civil Servers			0.025		Proportion of Appointed Civil Servers			-0.38839	
			(0.273)					(0.67719)	
Dummy PT			-0.005		Dummy PT			0.09750	
			(0.069)					(0.17550)	
Dummy PSDB			-0.026	-0.012	Dummy PSDB			-0.06237	
			(0.068)	(0.032)				(0.18161)	
Constant	7.070***	7.106***	5.158***	6.687***	Constant	-0.11682***	-0.09867	8.05764***	0.98530**
	(0.032)	(0.034)	(0.337)	(0.135)		(0.07798)	(0.08144)	(0.94058)	(0.42176)
Observations	1,416	1,419	1,443	1,416					
Log Likelihood	-6,802.394	-6,819.336	-6,841.909	-6,796.551	Log Likelihood	-6802 on 13 Df	-6819 on 17 Df	-2721 on 13 Df	2684 on 19 Df
Note:								*p<0.1; **p<0.05; ***p<0.01	

Source: The author.

APPENDIX C – TERMS TO CHECK FOR MENTIONS OF KINSHIP IN TEXT.

Table 14 - Terms included in functions to identify mention to kinship (Continues).

Type of kinship the function aims to detect	Terms in Portuguese included in the function
son/daughter	filho, filho do, filho de, filho da, filha do, filha de, filha, filha da, filh d, filhe d
father/mother	pai do, pai de, pai da, pái do, pái de, pái da, pãi do, pãi de, pai, pãi da, pâi do, pâi de, pâi da, mae, mãe, mae de, mae de, mae de, mae da, mae do, mãe de, mãe da, mãe do, mâe de, mâe da, mâe de, máe de, máe da, máe do, maede, maeda, maedo
matrimonal/in-law	casado com, casada com, casada , esposa da, esposa do, esposa, esposa de, espôsa da, espôsa do, espoza da, espoza do, espoza de, exposa da, exposa do, exposa de, ésposa da, ésposa do, ésposa de, espôza da, espôza do, espôza de, mulher da, mulher do, mulher de, múlher da, múlher do, múlher de, ex-mulher da, ex-mulher do, ex-mulher de, ex -mulher da, ex -mulher do, ex -mulher de, ex- mulher da, ex mulher do, ex mulher, ex mulher de, exmulher da, exmulher do, exmulher de, ex-esposa da, ex-esposa, ex-esposa do, ex-esposa de, exesposa da, exesposa do, exesposa de, ex -esposa da, ex -esposa do, ex -esposa de, ex- esposa da, ex- esposa do, ex- esposa de, ex esposa da, ex esposa do, ex esposa de, ex-espôsa da, ex-espôsa do, ex-espôsa de, ex-espoza da, ex-espoza do, ex-espoza de, ex espôsa da, ex espôsa do, ex espôsa de, ex -espôsa da, ex -espôsa do, ex -espôsa de, ex- espôsa da, ex-espôsa do, ex- espôsa de, esposo da, esposo, esposo do, esposo de, espôso da, espôso do, espôso de, espozo da, espozo do, espozo de, exposo da, exposo do, exposo de, ésposo da, ésposo do, ésposo de, espôzo da, espôzo do, espôzo de, marido da, marido, marido do, marido de, márido da, márido do, márido de, ex-marido da, ex-marido, ex-marido do, ex-marido de, ex marido da, ex marido do, ex marido de, ex- marido da, ex- marido do, ex- marido de, ex -marido da, ex -marido do, ex -marido de, ex-esposo da, ex-esposo do, ex-esposo de, ex -esposo da, ex -esposo do, ex -esposo de, ex- esposo da, ex- esposo do, ex- esposo de, ex esposo da, ex esposo do, ex esposo de, ex esposo, ex -esposo, cunhado da, cunhado, cunhado do, cunhado de, cunhado da, cunhado de, cunhada da, cunhada, cunhada do, cunhada de, sogra da, sogra, sogra do, sogra de, sogro da, sogro, sogro do, sogro de
siblings	irmao do, irmao, irmão, irmao de, irmao da, irmão do, irmão de, irmão da, irmáo do, irmáo de, irmáo de, irmáo da, irma do, irma de, irma, irmã, irma da, irmã do, irmã de, irmã da, irmá do, irmá de, irmá da, irmâ da, irmâ do, irmâ de, irmâo do, irmâo de, irmâo da

Source: The author.

Table 14 - Terms included in functions to identify mention to kinship (Ends here).

Type of kinship the function aims to detect	**Terms in Portuguese included in the function**
grandparents/grandchild	avo do, avo de, avo, avó, avô, avo da, avó do, avó de, avó da, avô do, avô de, avô da, avõ do, avõ de, avõ da , ávó do, ávõ de, ávô da, neto do, neto de, neto, neto da, neta do, neta, neta de, neta da, néto do, néto de, néto da, néta do, néta de, néta da
uncle/aunt	tio da, tio do, tio, tio de, tia, tia da, tia do, tia de, tío da, tío do, tío de, tía da, tía do, tía de, sobrinho da, sobrinho, sobrinho do, sobrinho de, sobrinha da, sobrinha, sobrinha do, sobrinha de
cousins	primo do, primo de, primo, primo da, prima do, prima, prima de, prima da, primá do, primá de, primá da, primâ do, primâ de, primâ da, primã do, primã de, primã da, primó do, primó de, primó da, primô do, primô de, primô da, primõ do, primõ de, primõ da, príma do, príma de, príma, príma da, prímo do, prímo de, prímo da

Source: The author.

APPENDIX D – FUNCTION TO IDENTIFY MENTION TO POLITICAL OFFICE.

```
# identify
def identifyPolitician(text):

    if  ' prefeito '  in text:
        new_text = 'politician'
    elif  ' prefeita '  in text:
        new_text = 'politician'
    elif  ' ex prefeit'  in text:
        new_text = 'politician'
    elif  '  ex-prefeit'  in text:
        new_text = 'politician'
    elif  ' ex -prefeit'  in text:
        new_text = 'politician'
    elif  ' ex- prefeit'  in text:
        new_text = 'politician'
    elif  ' deputad'  in text:
        new_text = 'politician'
    elif  ' ex-deputad'  in text:
        new_text = 'politician'
    elif  ' ex -deputad'  in text:
        new_text = 'politician'
    elif  ' ex- deputad'  in text:
        new_text = 'politician'
    elif  ' senador'  in text:
        new_text = 'politician'
    elif  ' ex-senador'  in text:
        new_text = 'politician'
    elif  ' ex -senador'  in text:
        new_text = 'politician'
    elif  ' ex- senador'  in text:
        new_text = 'politician'
    elif  ' vereador'  in text:
        new_text = 'politician'
    elif  ' ex-vereador'  in text:
        new_text = 'politician'
    elif  ' ex -vereador'  in text:
        new_text = 'politician'
    elif  ' ex- vereador'  in text:
        new_text = 'politician'
    elif  ' governador'  in text:
        new_text = 'politician'
    elif  ' ex-governador'  in text:
        new_text = 'politician'
```

```
    elif  ' ex -governador'  in text:
        new_text = 'politician'
    elif  ' ex- governador'  in text:
        new_text = 'politician'
    elif  ' president'  in text:
        new_text = 'politician'
    elif  ' ex-president'  in text:
        new_text = 'politician'
    elif  ' ex -president'  in text:
        new_text = 'politician'
    elif  ' ex- president'  in text:
        new_text = 'politician'
    else:
        new_text = 'non-politician'

    return new_text

# Add dyn_cand text to df as column
df['politician'] = df['page_text_clean'].apply(identifyPolitician)
df.info()
```

Source: The author.

APPENDIX E – MODEL RECIPE FOR SUPERVISED CLASSIFICATION.

```
model_recipe <-
  recipe(dynasty_classification ~ page_text_new_clean +
           sentence_words_categorical +
           mayor_name_is_in_text +
           kinship_sentences_count +
           type_of_search_tag +
           categorical_kinship_function_python, data = df_train) %>%
  # Use textrecipes to deal with page_text_new_clean
  # First, tokenize the text to words
  textrecipes::step_tokenize(page_text_new_clean) %>%
  # Keep only the 1000 most frequent tokens - Avoid creating too many variables
  textrecipes::step_tokenfilter(page_text_new_clean, max_tokens = 1e3) %>%
  # Compute tf-idf
  textrecipes::step_tfidf(page_text_new_clean) %>%
  # Combine low frequency factor levels
  #step_other(all_nominal(), threshold = 0.05) %>%
  # Remove no variance predictors which provide no predictive information
  step_nzv(all_predictors()) %>%
  # Normalize numeric variables to have a standard deviation of one and a mean of zero
  step_normalize(all_numeric()) %>%
  # Convert nominal variables to numeric binary variables
  # -all_outcomes leave class_dyn out
  step_dummy(all_nominal(), -all_outcomes(), one_hot = TRUE) %>%
  # Imput mean value to missing
  #step_impute_mean(sentence_words_count) %>%
  # Remove rows of a data set to make the occurrence of levels in a specific
  #factor level equal
  themis::step_downsample(dynasty_classification, under_ratio = 1)
```

```
# Downsample (w/ under ratio = 1) was the best strategy to deal w/ imbalanced
# data, as I tested several other possibilities (eg.: usample, smote etc.)
```