

The Pennsylvania State University
The Graduate School
College of the Liberal Arts

**SPEAKING ON OUR BEHALF: ESSAYS ON POLITICAL
REPRESENTATION**

A Dissertation in
Political Science
by
Christopher Boylan

© 2017 Christopher Boylan

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

August 2017

The dissertation of Christopher Boylan was reviewed and approved* by the following:

Burt L. Monroe
Professor of Political Science, Social Data Analytics, and Information Science

Matt Golder
Professor of Political Science

Sona N. Golder
Professor of Political Science

Marie Hojnacki
Associate Professor of Political Science

John Gastil
Professor of Communication, Arts & Sciences and Political Science

James Piazza
Professor of Political Science
Director of Graduate Studies of the Department of Political Science

*Signatures are on file in the Graduate School.

Abstract

Politicians' representation of constituents' interests occupies a central role in modern democratic politics. Using new data sources and computationally intensive methods, I extend our understanding of political representation with a series of empirical essays examining the measurement of political representation and the influence of political factors on legislators' attention to different components of political representation. I show that legislators use legislative speeches, legislative questions, and tweets to emphasize different aspects of their legislative activities and demonstrate that legislators' tweets are a useful resource for generating new insights into legislative behavior. I also present evidence indicating that dynastic legislators in rural areas provide more constituency service than non-dynastic legislators in these areas and show that there is a strong positive association between election via an ethnic quota and legislators' attention to issues related to historically disadvantaged minority ethnic groups.

Contents

List of Figures	vii
List of Tables	ix
Acknowledgments	xi
1 Introduction	1
1.1 The Representation Literature	2
1.1.1 Substantive Representation	2
1.2 Outline	6
2 Measuring Legislators' Representation of Constituents	10
2.1 The limitations of roll call analyses	11
2.2 Beyond roll call votes	12
2.2.1 Legislator surveys and interviews	12
2.2.2 Audit studies	13
2.2.3 Legislative speeches	13
2.2.4 Legislators' press releases	14
2.2.5 Legislative questions	14
2.2.6 Legislators' tweets	16
2.3 Comparing representatives' speeches, questions, and tweets	17
2.3.1 A descriptive comparison of speeches, questions, and tweets	18

2.3.2	Comparing the content of legislators' speeches, questions, and tweets	26
3	Political Dynasties and Political Representation	30
3.1	Introduction	30
3.2	Dynastic Legislators and Representation	34
3.3	Research Design	38
3.3.1	Ireland	38
3.3.2	Data	39
3.3.3	Measuring service responsiveness	41
3.3.4	Model specification	47
3.4	Results	48
3.5	Discussion	54
4	Ethnic Quotas and Substantive Representation of Ethnic Minorities	57
4.1	Introduction	57
4.2	Electoral Quotas and Substantive Representation	59
4.3	Research Design	62
4.3.1	India	62
4.3.2	Data	63
4.4	Results	68
4.5	Discussion	73
5	Politweeting: Exploring Legislators' Use of Twitter	76
5.1	Introduction	76
5.2	Who tweets?	78
5.3	Understanding the content of legislators' tweets	83
5.3.1	Policy responsiveness on Twitter	87
5.3.2	Allocation responsiveness on Twitter	89

5.3.3	Service responsiveness on Twitter	91
5.3.4	Legislator advertising on Twitter	94
5.4	Discussion	96
6	Conclusion	98
	Bibliography	105
	Appendix	120
A.1	Technical Details	120
A.1.1	Text as data	120
A.1.2	Details of Fightin' Words model	125
A.1.3	Details of LASSO model	127
A.2	Chapter 3	130
A.2.1	Details of PR-STV	134
A.3	Chapter 4	135
A.3.1	Details on the assignment of ethnic quotas to districts in India	138
A.4	Chapter 5	139

List of Figures

2.1	Word counts of legislators' questions, speeches, and tweets	20
2.2	Distribution of number of speeches by TDs between March 2011 and February 2016	21
2.3	Distribution of number of questions by TDs between March 2011 and February 2016	22
2.4	Distribution of number of tweets by TDs between March 2011 and February 2016	23
2.5	Number of speeches, questions, and tweets per month by TDs	25
3.1	Dynastic legislators in Dáil Éireann over time	39
3.2	Number of questions submitted by legislators across Dáil Éireanns, 1973-2007	40
3.3	Number of service questions submitted by legislators across Dáil Éireanns, 1973-2007	46
3.4	Membership in a dynasty and the proportion of constituency service questions asked	49
3.5	Dynastic membership and service responsiveness in Ireland 1973-2007	50
3.6	Expected counts of constituency service questions	53
4.1	Distribution of number of questions by MPs between 1999 and 2009 .	65
4.2	Distribution of number of questions related to marginalized ethnic groups by MPs between 1999 and 2009	66

4.3	Bivariate relationship between ethnic quotas and questions related to ethnic groups	69
4.4	Ethnic quotas and policy responsiveness in India 1999-2009	70
4.5	Expected counts of questions related to ethnic minority groups	72
5.1	The adoption of Twitter by TDs in the 31st Dáil Éireann	80
5.2	Variation in Twitter usage by TDs in the 31st Dáil Éireann	81
5.3	The usage of Twitter by TDs in the 31st Dáil Éireann	82
5.4	Topic proportions and top terms associated with each topic from correlated topic model	85
5.5	Example tweets from different policy topics	87
5.6	Difference in government and opposition legislators' attention to policy issue	89
5.7	Example tweets from allocations topic	90
5.8	Difference in legislators' attention to allocations	91
5.9	Example tweets from constituency clinics topic	92
5.10	Differences in legislators' to constituency activity	93
5.11	Example tweets from advertising topics	94
5.12	Difference in urban and rural legislators' advertising behavior	95
6.1	Differences in words usage between Fianna Fáil and non-Fianna Fáil legislators	102
A.1	Example of a legislative text HTML page	121
A.2	Example of a tweet returned by API	123
A.3	Example of parliamentary question HTML page	124
A.4	Laplace Distribution	128

List of Tables

2.1	Total number of speeches, questions, and tweets by TDs between March 2011 and February 2016	18
2.2	Total number of TDs that made a legislative speech, asked a legislative question, or tweeted between March 2011 and February 2016	21
2.3	Top 50 words associated with legislators' speeches	26
2.4	Top 50 words associated with legislators' questions	27
2.5	Top 50 words associated with legislators' tweets	28
3.1	Results of constituency service classification model	43
3.2	Coefficients for most important words from LASSO logistic regression	44
4.1	Total number of questions asked by legislators in India between 1999 and 2009	64
4.2	Total number of legislators elected through ethnic quotas in India 1999-2009	67
5.1	TDs who served in the 31st Dáil Éireann Twitter status	78
A.1	Example questions classified as constituency service	130
A.2	Example questions classified as not constituency service	131
A.3	Coefficients from negative binomial models	133
A.4	Example questions related to ethnic minority groups	135
A.5	Example questions not related to ethnic minority groups	136
A.6	Coefficients from negative binomial models	137

A.7	Topic labels for Twitter topic model	139
-----	--	-----

Acknowledgments

I would like to thank my advisor and dissertation chair Burt Monroe for his advice, support, patience, and data during the dissertation process and my graduate studies more generally. I would also like to thank Matt Golder for his invaluable critical feedback during my graduate studies. Likewise, Sona Golder has consistently supported my academic development. I thank Marie Hojnacki for agreeing to join the committee at such a late stage. I would also like to acknowledge John Gastil, Ryan McMahon, Josh Semat, Laurie Tumaneng, Ted Chen, and Padraig Fitzgerald for their insightful comments on this project.

Chapter 1

Introduction

Politicians' representation of constituents' interests occupies a central role in modern democratic politics. Citizens in democracies elect a group of individuals to govern and represent their interests. In many electoral democracies, a legislator's representation of the interests of their constituency and individual constituents is therefore crucial for remaining in office. In this work, I attempt to extend our understanding of political representation with a series of empirical essays that provide new insights into the activities of legislators. Using new data sources and computationally intensive methods, I contribute to the literature on political representation by comparing what politicians discuss on different communication platforms, examining how institutional factors and the personal characteristics of politicians influence the policy and service responsiveness of legislators, and analyzing legislators' use of the social media platform Twitter. To frame the contribution of this work, I first discuss the literature pertaining to political representation in the following section.

1.1 The Representation Literature

The literature addressing legislator and constituent relations is one of the richest and largest in political science. Consequently, I primarily focus on the important work from the representation literature that is relevant to this project.¹ In her seminal work, Pitkin (1967) argues that political representation is a multifaceted concept comprising four dimensions, each of which can be utilized for analyzing representation. Specifically, Pitkin (1967) identifies these dimensions as formalistic, symbolic, descriptive, and substantive representation. Formalistic representation refers to the formal institutional procedures that precede and enable the act of representing such as elections. Symbolic representation refers to the degree that representatives stand for constituents, with a particular emphasis on constituents' perceptions about the representative. Descriptive representation refers to the degree to which the demographic distribution of representative institutions reflect the demographic distribution of represented constituents. Finally, substantive representation refers to the activities undertaken by a legislator to represent constituents (Pitkin, 1967). In line with Pitkin (1967), I consider the dimension of substantive representation to be particularly relevant and interesting, and therefore focus on research pertaining to substantive representation below.

1.1.1 Substantive Representation

In a similar vein to the work of Eulau and Karpis (1977) and Harden (2016), I consider substantive representation to be a multidimensional concept, comprising the three components of policy, allocation, and service responsiveness. Although this type of conceptualization has been present in the literature for a substantial period of time, it is important to reiterate that a more holistic comprehension of politicians' substantive

¹It would be somewhat infeasible to provide an exhaustive review of the work on political representation. For instance, Pitkin (1967), which is discussed in this section, has been cited nearly 7000 times, according to Google Scholar.

representation of constituents requires recognition of all three components (Eulau and Karps, 1977).

Policy Responsiveness

Substantive representation is perhaps most frequently studied by examining the congruence between constituents' policy preferences and the policy preferences of representatives (Eulau and Karps, 1977; Powell, 2004; Grimmer, 2013*b*).² This approach to studying representation follows the work of Miller and Stokes (1963), who found that the influence of constituents' preferences on the actions of Congressmen varied across issue areas. Although other scholars challenged the conclusions of this work (Achen, 1977, 1978; Erikson, 1978), a substantial number of similar studies exploring congruence between constituents' and representatives' preferences in America and other countries were conducted (e.g. Barnes, 1977; Converse and Pierce, 1986; Jackson and King, 1989; Clinton, 2006).

These single-country studies were followed by a number of works which attempted to explain cross-national variation in the level of congruence, with a particular emphasis on the influence of electoral rules (e.g. Huber and Powell, 1994; Powell and Vanberg, 2000; Powell, 2000; Golder and Stramski, 2010). Evidence from this research indicates that countries with proportional electoral rules boast legislatures that are more congruent to constituents' preferences than countries with majoritarian electoral rules (Golder and Stramski, 2010). While policy responsiveness is clearly a central component of representation, a representative may also respond to the policy preferences of other actors such as the executive, party leaders, and interest groups (Eulau and Karps, 1977; Urbinati and Warren, 2008). It is therefore important to recognize that an exclusive focus on congruence provides a somewhat limited understanding of

²When discussing policy responsiveness throughout this work, I am referring to the congruence between constituents' preferences and legislators' preferences. There are a number of other works which identify policy responsiveness as the response of policy to changes in public opinion (e.g. Page and Shapiro, 1983; Stimson, MacKuen and Erikson, 1995).

substantive representation (Hall, 1998; Butler, Karpowitz and Pope, 2012; Harden, 2013).

Allocation Responsiveness

Another important component of substantive representation is the allocation of pork barrel goods that are typically beneficial to constituents in a representative's district (Mayhew, 1974; Eulau and Karps, 1977; Shepsle and Weingast, 1981; Fiorina, 1989; Grimmer, Messing and Westwood, 2012; Golden and Min, 2013). Although such goods are often classified as local public goods, some allocations are specifically targeted to a small group of constituents and it is therefore difficult to distinguish public from private goods (Fiorina, 1977; Eulau and Karps, 1977). Allocations provided by representatives include grants for bridges, highways, and hospitals as well as funds for local police and educational programs (Fiorina, 1977; Milesi-Ferretti, Perotti and Rostagno, 2002). In the American context, it is often argued that sitting on relevant committees/subcommittees and membership in the majority party produces pork barrel benefits for a representative's district (e.g. Mayhew, 1974; Weingast and Marshall, 1988; Aldrich and Rohde, 2001; Cox and McCubbins, 2005). However, these claims are challenged by recent evidence which indicates that the partisan effect on the allocation on pork barrel goods is negligible and that only the chairs of the appropriations subcommittees benefit from their positions (Fowler and Hall, 2015; Berry and Fowler, 2015).

Studies on pork barrel politics in other countries indicate that marginal districts receive more allocations (Ward and John, 1999; Arulampalam et al., 2009; Vaishnav and Sircar, 2011), and that the ethnic and/or religious identity of constituents are important factors for allocations (Banerjee and Somanathan, 2007; Thachil and Teitelbaum, 2015). There is also a strong interest in the influence of a country's electoral rules on representatives' allocation of pork barrel goods (Cain, Ferejohn and Fiorina,

1987; Persson and Tabellini, 2002). This work suggests representatives serving in single member districts will supply more pork to their constituencies than their counterparts elected under proportional representation electoral rules (Lizzeri and Persico, 2001; Stratmann and Baur, 2002; Milesi-Ferretti, Perotti and Rostagno, 2002).

Service Responsiveness

A further key component of substantive representation is the service that representatives provide to individual constituents and groups in their constituency (Mayhew, 1974; Fenno, 1977; Cain, Ferejohn and Fiorina, 1987). This service component of representation, also known as casework, refers to benefits that representatives can provide to particular constituents. Services provided by representatives include responding to constituents' queries and helping constituents with issues with bureaucracies such as tax agencies and social welfare services (Eulau and Karps, 1977; Fenno, 1978). Providing service is also beneficial to representatives as a large number of constituents care more about service than policy issues and it can help representatives to cultivate a personal vote, thereby potentially alleviating the effect of changes in the national popularity of political parties (Cain, Ferejohn and Fiorina, 1987). Research attempting to decompose the sources of the incumbency advantage in American elections indicates that a substantial portion of the incumbency advantage can be ascribed to the personal vote of a legislator (Levitt and Wolfram, 1997; Ansolabehere, Snyder Jr and Stewart III, 2000; Carson, Engstrom and Roberts, 2007).

However, it is somewhat unclear whether such personal vote seeking strategies have a similar electoral benefit in other countries with single member districts (Mezey, 1994; Norton and Wood, 1993; Gaines, 1998). Turning once more to the influence of institutional factors, studies suggest that constituency service is more prevalent in single member district systems than in proportional systems with large multimember districts, although constituency service is promoted in the latter by intraparty

electoral competition (Scholl, 1986; Carey and Shugart, 1995; Grofman, 1999). In a similar vein, others posit that the interaction between district magnitude and electoral rules influences representatives' propensity to engage in activities which cultivate a personal vote under proportional representation (Carey and Shugart, 1995; Shugart, Valdini and Suominen, 2005). However, some studies empirically testing theories pertaining to the cultivation of a personal vote employ proxy measures such as the birthplace of legislators rather than measures that directly capture the service provided by legislators to constituents (e.g. Shugart, Valdini and Suominen, 2005).

1.2 Outline

In this work I build on the literature by examining how different factors impact the representation legislators provide to individual constituents and groups of constituents. I do this by first analyzing and comparing how different data sources can be utilized for measuring representation and its components. I then examine the empirical relationship between membership in a political dynasty and the service activities of legislators. I follow this with an examination of the influence of ethnic quotas on legislators' responsiveness to policies that pertain to ethnic minority groups. I then provide an exploratory study of how legislators utilize the new social platform Twitter, which has quickly become an important forum for political discussion. This dissertation therefore contributes to a number of different substantive topics in the political science literature. Moreover, by utilizing computationally intensive methods to create and analyze large datasets of legislative text, this work also makes a contribution to the computational social science literature.

In the next chapter I discuss and compare various ways of measuring and operationalizing legislators' representation of their constituents. I argue that roll call votes are not a particularly useful tool for understanding the representational styles of legislators in many parliamentary democracies and that the analysis of legislative

text data can provide a more complete and nuanced understanding of legislators' representation of constituents. I then conduct descriptive and content analyses on three different sources of legislative text — legislative speeches, questions submitted by legislators, and legislators' tweets — to assess how these sources can be utilized to measure the different components of representation discussed earlier. The results of these content analyses indicate that all three data sources can provide valuable insights into different aspects of representation. Specifically, it appears that speeches are particularly useful for understanding the policy responsiveness of legislators, that parliamentary questions can be utilized for measuring both the service and policy responsiveness of legislators, and that legislators' tweets are especially useful for gaining insight into the advertising activity of legislators.

In the following chapter, I investigate the relationship between membership in a political dynasty and the constituency service provided by legislators. While there is a growing literature on political dynasties, little is known about the representational styles and the service responsiveness of dynastic politicians. I posit that a unique political socialization in which their previously elected relatives emphasize the importance of the cultivation of a personal vote and the social networks associated with dynasties lead to dynastic legislators engaging in more constituency service than non-dynastic legislators. Using an original dataset comprising the dynastic ties of legislators and written questions submitted by legislators to ministers in Ireland from 1973 to 2007, I find that dynastic legislators in rural areas appear to provide more constituency service than non-dynastic legislators in these areas. My analysis also suggests that dynastic legislators in highly urbanized areas engage in similar service responsiveness to non-dynastic legislators in highly urbanized areas. These results suggest that different mechanisms may underlie the formation and persistence of dynasties in different geographical regions.

In the fourth chapter, I examine the influence of electoral quotas for ethnic minorities on legislators' attention to policies relevant to these underrepresented groups.

The introduction of electoral quotas in countries is typically based on the idea that quotas can enhance the political representation of minority groups. Although the majority of research indicates that quotas improve the descriptive representation for underrepresented sections of society and increase allocations to the minority groups, it is less clear whether quotas improve legislators responsiveness to minority groups' concerns and interests. Building on existing research, I assume that legislators elected via ethnic quotas understand that these institutional arrangements were introduced with the goal of promoting the substantive representation of their respective group. I therefore argue they work to achieve this goal by focusing on the interests of their fellow group members more than legislators not elected via quotas. Using an original dataset consisting of detailed legislator information and written questions submitted to ministers by legislators in India from 1999 to 2009, I find that there is a strong positive association between election via an ethnic quota and legislators' attention to issues related to historically disadvantaged minority ethnic groups.

In the penultimate chapter, I provide a detailed exploratory study of legislators' use of Twitter. Using an original dataset comprising all tweets by Irish legislators between 2011 and 2016, I contribute to the growing literature employing tweets to research legislators' activities by examining variation in the frequency of legislators' Twitter use and analyzing the content of legislators' tweets. I find that the age of a legislator, whether a legislator holds ministerial office, and whether a legislator serves in an urbanized geographic region are influential predictors of a legislator's Twitter usage. In my analysis of the content of legislators' tweets, I demonstrate that legislators tweet about both political and non-political issues, a number of the topics legislators tweet about can be utilized to study important concepts in political representation, and the covariance between these topics and political factors meet theoretical expectations. This suggests that legislators' tweets are a useful resource for studying political representation.

In the final chapter, I conclude by discussing the contribution of this dissertation

and potential future research projects related to this work.

Chapter 2

Measuring Legislators’ Representation of Constituents

As noted in my earlier review of the representation literature, a large share of studies examining representation follow the work of Miller and Stokes (1963) and solely focus on the congruence between constituents’ preferences and the policy preferences of representatives. In their work, Miller and Stokes (1963) surveyed constituents and explored how strongly representatives’ votes on different issues correlated with the opinions of constituents. This approach quickly became the default strategy for measuring representation in quantitative studies of the topic (Eulau and Karps, 1977; Grimmer, 2013*b*). Below I discuss the limitations of this approach and outline alternative approaches for measuring politicians’ representation of constituents. In this discussion, I pay particular attention to how text sources can be used for this purpose. Using original data that I collected, I then quantitatively compare three different text sources — legislative speeches, legislative questions, and legislators’ tweets — that are commonly available for parliamentary democracies to understand the utility of these sources for measuring different components of representation. I find that text of legislative questions can provide deep insight into the service responsiveness of

legislators, the text of speeches primarily helps us to better understand the policy responsiveness of legislators, and that tweets provide new information for understanding legislators' advertising of their activities.

2.1 The limitations of roll call analyses

While Miller and Stokes (1963) argued that their congruence study should only be a starting point for measuring and understanding representation, this model was quickly adopted as the predominant definition of representation (Eulau and Karps, 1977). This led to a large number of studies employing roll call votes or low dimensional summaries of voting histories to understand policy responsiveness and the behavior of representatives (Poole and Rosenthal, 1985; Lee, Moretti and Butler, 2004; Clinton, 2006). However, little evidence has been provided to demonstrate that roll call votes alone are sufficient for understanding legislators' representation of constituents (Grimmer, 2013*b*).

Roll call votes only inform us about the policy preferences of issues that make it through to this stage of the legislative process. Work on legislative agenda setting shows that leaders of parties control the legislative agenda to prevent important but potentially problematic issues from being voted on (Cox and McCubbins, 2005). Furthermore, while in the American context legislators often exhibit significant autonomy from their parties, in many parliamentary democracies there are extremely disciplined parties (Laver, Benoit and Garry, 2003; Spirling and McLean, 2007). In such disciplined parties failure to vote along party lines can result in legislators losing their party affiliation and greatly reduce their likelihood of re-election. Therefore, while methods have been developed to identify intraparty factions from roll call votes in such contexts (Spirling and Quinn, 2010), it is commonly the case that there is very little variation in how party members vote on issues and roll call votes are not particularly informative. Additionally, while roll calls can be a useful tool for under-

standing policy responsiveness in certain contexts, as noted earlier, representation is a multidimensional concept, and roll call votes provide limited information about allocation and service responsiveness. The study of these components of representation therefore necessitates the use of other data sources and empirical strategies. I discuss these different strategies below.

2.2 Beyond roll call votes

There are a number of methods and data sources that can supplement roll call votes and help arrive at a more complete understanding of political representation. Research on representation has employed surveys of legislators, conducted experiments on legislators and their offices, and, analyzed large corpora of legislative text such as representatives' speeches, press releases, parliamentary questions, and posts on the social media website Twitter (Cain, Ferejohn and Fiorina, 1987; Butler, Karpowitz and Pope, 2012; Quinn et al., 2010; Grimmer, 2013*a*; Martin, 2011*a*; Barberá et al., 2014). I discuss each of these approaches and their respective benefits and limitations individually below.

2.2.1 Legislator surveys and interviews

A straightforward approach to understanding the representational style of politicians is to use surveys or interviews to learn about their activities and their priorities (Fenno, 1978; Cain, Ferejohn and Fiorina, 1987; Norton and Wood, 1993; Farrell and Scully, 2007). This approach can yield detailed information about the tasks undertaken by representatives and their views on these tasks. However, it can be problematic to convince legislators to participate in surveys and it is obviously very difficult to assess the validity of the responses provided. These problems thereby produce uncertainty regarding the reliability of findings from studies employing this

research design.

2.2.2 Audit studies

Recently researchers have started to use field experiments to learn more about the responsiveness of legislators (Butler, Karpowitz and Pope, 2012; Harden, 2013; Grose, Malhotra and Parks Van Houweling, 2015; Kalla and Broockman, 2015). These experiments usually involve contacting legislators using letters or email and randomly assigning some treatment related to the characteristic of the putative constituent or the topic of the constituent request (Butler and Broockman, 2011; McClendon, 2016). These treatments allow researchers to measure the representational priorities of legislators and learn whether legislators discriminate against constituents. Although these studies boast stronger validity regarding the relationship between the treatment and outcome than a typical observational study (Trochim and Donnelly, 2001), similar to legislator surveys, experiments can suffer from low response rates from legislators. While this issue can be alleviated if the sample size is large, most national legislatures typically have a few hundred legislators. This also means that such experiments can be statistically underpowered, making it impossible to detect a treatment effect if the effect is small (Montgomery, 1991).

2.2.3 Legislative speeches

Legislative speeches have become increasingly popular data source for understanding the preferences of representatives (Laver, Benoit and Garry, 2003; Monroe and Maeda, 2004; Slapin and Proksch, 2010). In legislative speeches representatives generally discuss bills which propose new legislation and debate topical issues and policies that are of public relevance (Houses of the Oireachtas, 2016). In parliamentary democracies with strong parties, legislators are typically less constrained in speeches than in roll call votes. It appears that party leaders are willing to accept some vocal dissent in

speeches as long as legislators vote along party lines (Schwarz, Traber and Benoit, 2015). Speeches therefore provide researchers with a more nuanced understanding of representatives' policy preferences than roll call votes. However, there is also evidence party leaders will deny potentially dissident legislators access to the floor if there are incentives to maintain the reputation of a unified party (Proksch and Slapin, 2012). This suggests that, although offering an improvement over roll calls, speech data may also be affected by some selection bias.

2.2.4 Legislators' press releases

Representation studies have also employed legislators' press releases to understand the representational priorities of legislators (Grimmer, 2010; Grimmer, Messing and Westwood, 2012; Grimmer, 2013*a,b*). Grimmer (2013*b*) shows that press releases are widely used by US Senators to inform constituents of their work. Press releases are published at all times throughout the year and there are no formal constraints on their content. Grimmer (2013*b*) demonstrates that the content of press releases differs considerably from that of floor speeches, with Senators regularly using press releases to claim credit for appropriations. The analysis of press releases therefore helps to provide a broader understanding of how a legislator represents constituents than the sole use of either roll calls or floor speeches. However, although commonly published by legislators in the United States, press releases are not widely used in other legislatures and are often simply not an available data source for researchers studying other contexts.

2.2.5 Legislative questions

Another data source which can be utilized to learn about the representational styles of legislators is the text of legislative questions. Legislative questions are a central activity in most national legislatures in the world (Franklin and Norton, 1993; Martin,

2011a).¹ There is substantial variation in the rules regarding questions across national legislatures, some countries permit oral, written, and special urgent questions whereas others only have procedures for a subset of these. To get an idea of how the question process operates, I detail the formal procedures for submitting parliamentary questions in Ireland, which is the main country of interest in this work.

In Ireland, members of parliament can ask oral questions to government Ministers three days a week and ask oral questions to the Taoiseach one day a week during parliamentary sessions.² Ministers answer oral questions according to an agreed rota (Houses of the Oireachtas, 2016). This enables legislators to query Ministers on a regular basis. Irish TDs can also submit written questions which will receive a written reply from the relevant Ministry.³ Legislators can ask a maximum of two oral parliamentary questions to a minister per day whereas there is no limit on the number of written questions that can be submitted (Houses of the Oireachtas, 2016). Every day five oral questions are prioritized, with priority questions shared between parties based on the number of seats held by each party. Questions are drafted by either a legislator or a staff member of a legislator's office. Representatives are required to provide a notice of five days for oral questions and a notice of three days for written questions (Houses of the Oireachtas, 2016).

Parliamentary questions are predominantly utilized by opposition and backbench government legislators, thereby providing information on the large majority of representatives in the legislature.⁴ Moreover, in contrast to roll call votes and plenary speeches, party leaders exhibit very little control over the content of parliamentary

¹Notably, the United States Congress does not have a procedure enabling legislators to submit questions. This may help explain the understudy of legislative questions given the majority of representation research focuses on Congress (Martin, 2011a).

²The Taoiseach is the prime minister of Ireland.

³TD is an acronym for Teachta Dála, a legislator serving in the lower house of the Irish parliament.

⁴Backbench government legislators are government legislators that do not serve in the cabinet. The term originates from the tradition of cabinet members sitting on the front benches during parliamentary proceedings.

questions.⁵ Legislators therefore tend to employ questions to address a mixture of national and local constituency issues (Martin, 2011*b*; Saalfeld, 2011). This means that parliamentary questions can be utilized to measure both the policy and service responsiveness of representatives.

2.2.6 Legislators' tweets

In recent years the social media website Twitter has become an important platform for political communication. Millions of Twitter users express political opinions on Twitter everyday. Both ordinary citizens and influential political actors such as high ranking civil servants, leaders of international organizations, and elected representatives are active on Twitter. Most legislators in advanced democracies have Twitter accounts (Barberá, 2015). The extant Twitter literature indicates that Twitter network data can be utilised to accurately measure and compare the policy preferences of legislators, other political actors, and citizens, that tweets provide information on the campaign strategies of election candidates, and legislators utilize tweets to discuss a number of issues (Lassen and Brown, 2011; Barberá et al., 2014; Evans, Cordova and Sipole, 2014; Barberá, 2015). This suggests that Twitter can be a useful data source for helping understand the representational priorities of legislators. However, one obvious challenge to using Twitter data is that tweets cannot be longer than 140 characters. Tweets therefore necessarily contain less information than the typical speech, question, or press release of a legislator.

⁵Given that oral questions asked in a parliamentary session are more likely to receive public attention, it seems fair to assume that oral questions are much more strongly regulated by party leaders than written questions.

2.3 Comparing representatives' speeches, questions, and tweets

As noted above, the text of legislative speeches, parliamentary questions, and representatives' tweets is freely available for most parliamentary democracies.⁶ Although each of these data sources have been used in studies by a large number of researchers, to my knowledge the content of these data sources has not been systematically compared. This is important as knowledge of the strategies employed by legislators across these platforms also enables representation researchers to carefully choose the data source(s) that is the most appropriate match for their question. Therefore, to understand whether legislators utilize these communication platforms to emphasize different components of representation, I perform a content analysis on the text of all Irish TDs' speeches, questions, and tweets between March 2011 and February 2016, which covers the duration of the 31st Dáil Éireann.⁷ Before discussing the results of this analysis, to understand these data sources better, I first provide a detailed descriptive analysis of this data.

⁶However, as outlined in the Appendix, obtaining and organizing this data into a usable format is often not trivial.

⁷Dáil Éireann is the lower house of the bicameral Irish parliament.

2.3.1 A descriptive comparison of speeches, questions, and tweets

Table 2.1: Total number of speeches, questions, and tweets by TDs between March 2011 and February 2016

Source	Documents	Words
Speeches	271,059	37,081,000
Questions	197,314	9,170,422
Tweets	148,580	2,027,864
Total	616,267	48,292,770

I created this dataset by scraping the text of legislative speeches and questions from the Dáil Éireann website and accessing legislators’ tweets from the Twitter API.⁸ A speech is defined as any contribution to a parliamentary debate by a legislator that is not a question. A document is identified as a question if a legislator is querying the Taoiseach, the Tánaiste, a minister, or a junior minister about a given issue or policy.⁹ I utilized information contained in the webpages of speeches and questions to correctly distinguish speeches from questions.¹⁰ A tweet is simply the text of a given tweet by a legislator, which is easy to identify using the Twitter API. In Table 2.1, I present the total number of speeches, questions, and tweets by TDs for the period of the 31st Dáil Éireann. From this we can see that the combined corpus of speeches, questions, and tweets contains over 616,000 unique documents.¹¹ This corpus comprises over 270,000 speeches, close to 200,000 questions, and nearly 150,000 tweets. The size and

⁸More details on how this data was obtained and processed for analysis can be found in the Text as Data section of the Appendix.

⁹The Tánaiste is the deputy prime minister of Ireland.

¹⁰I do not differentiate between oral questions and written questions in this analysis. Both oral and written questions are included in the analysis under the title of questions. Although queries related to a legislators’ specific constituency appear to be more prevalent in written questions, legislators also use oral questions for localistic inquiries.

¹¹Throughout this work, I only focus on tweets by legislators that are not retweets, where a user re-posts another user’s tweet, or replies.

scope of this corpus provides a detailed insight into the behavior and activities of legislators.

Turning to the content of the corpus, I present the distribution of logged document word counts of legislators' questions, speeches, and tweets in Figure 2.1. From this plot, and in line with expectations, it is clear that tweets typically contain fewer words than either legislative questions or speeches. The mean word count of tweets is 14, while the median number of words used in a tweet is 14. From this plot, it is also noticeable that there is substantial variation in the word counts of legislative speeches. There are both very short and very long speeches. This is not surprising as legislators often engage in swift exchanges with their colleagues in addition to discussing bills, amendments, and legislation in great detail. The mean number of words spoken in a speech is 137, while the median word count of speeches is 28. Finally, we can see that legislative questions are typically longer than tweets and there is less variation in the length of questions compared to speeches. The mean number of words employed in a question is 46, with the median question containing 41. This basic descriptive analysis of the data suggests that there is wider variation in the level of information contained in speeches compared to both tweets and questions.

Figure 2.1: Word counts of legislators' questions, speeches, and tweets

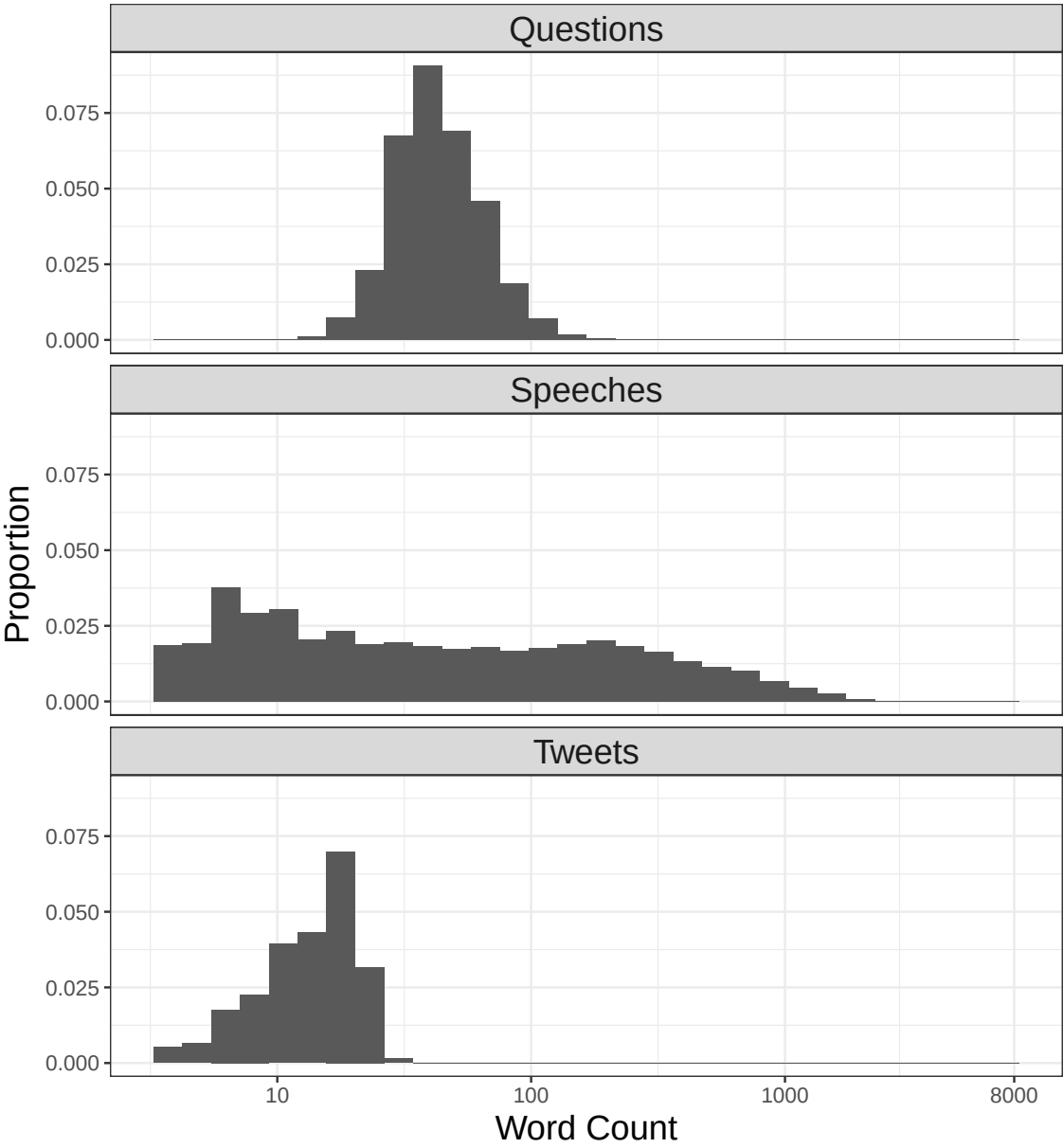
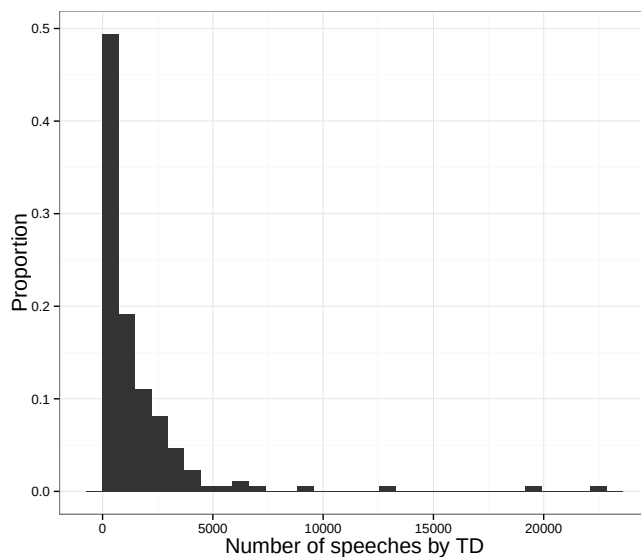


Table 2.2: Total number of TDs that made a legislative speech, asked a legislative question, or tweeted between March 2011 and February 2016

Communication platform	TDs that communicated using this platform
Speech	172
Question	167
Tweet	144

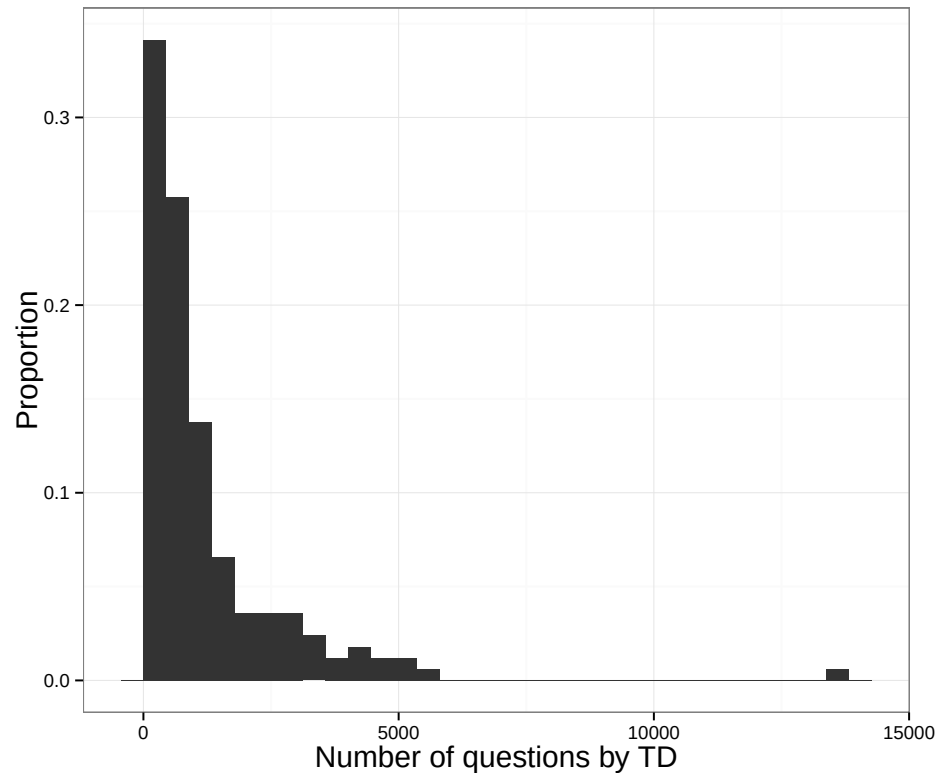
I next provide information on the variation in the number of speeches, questions, and tweets by legislators. In Table 2.2, I present the number of legislators that made a speech, asked a parliamentary question, or wrote a tweet. Out of the total 173 TDs that served in the 31st Dáil, 172, or 99 per cent, made a legislative speech, 167, or 97 per cent, submitted a legislative question, and 144, or 83 per cent, wrote a tweet. Although 153 of 173 TDs have Twitter accounts, nine didn't actually post a tweet in the relevant period.

Figure 2.2: Distribution of number of speeches by TDs between March 2011 and February 2016



Focusing on variation in the total number speeches made by legislators during the period of the 31st Dáil, from Figure 2.2 we can see that a small number of legislators made far more speeches than other legislators. The speaker of Dáil Éireann, the Ceann Comhairle, Sean Barrett, unsurprisingly, made the most speeches in the relevant period, with the Taoiseach, Enda Kenny, accounting for the second largest number of speeches. The mean number of speeches made by a legislator in this period is 1,572, with the median TD making 760 contributions to parliamentary debates. This suggests that speeches provide good coverage of legislators' activities in the parliament.

Figure 2.3: Distribution of number of questions by TDs between March 2011 and February 2016



Turning to the variation in the total number of questions asked by legislators

during the period of the 31st Dáil, from Figure 2.3 it is immediately obvious that one TD, Bernard Duggan, a veteran member of the majority government party Fine Gael known for his dedication to constituency work, asked far more questions than any of his colleagues.¹² By submitting over 13,000 questions, Deputy Duggan asked more than twice the number of questions submitted by any other legislator. The mean number of questions submitted by a legislator during this Dáil is 1,176, with the median TD making 682 enquiries. This descriptive examination indicates there is also substantial variation in legislators' utilization of parliamentary questions.

Figure 2.4: Distribution of number of tweets by TDs between March 2011 and February 2016

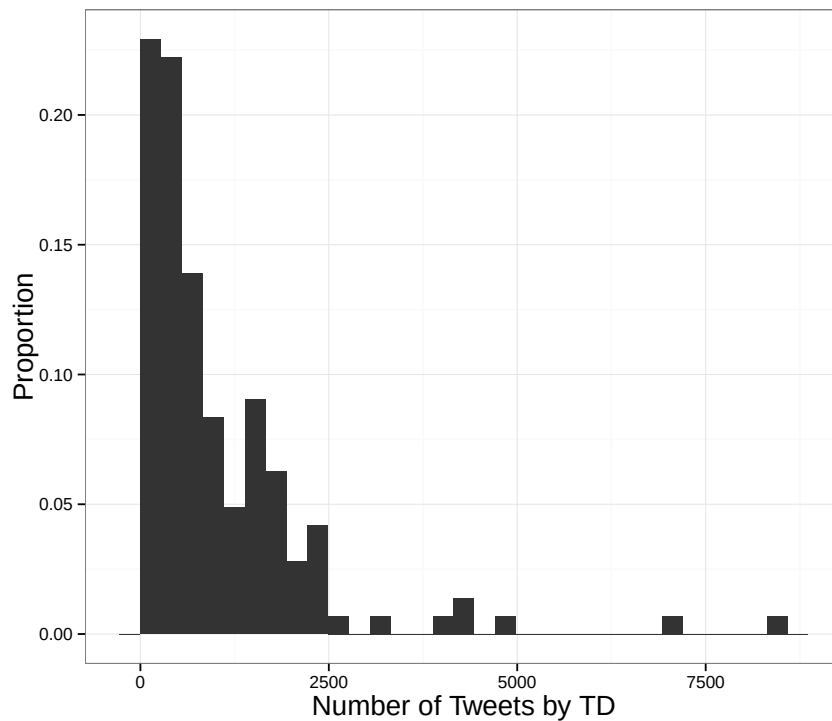


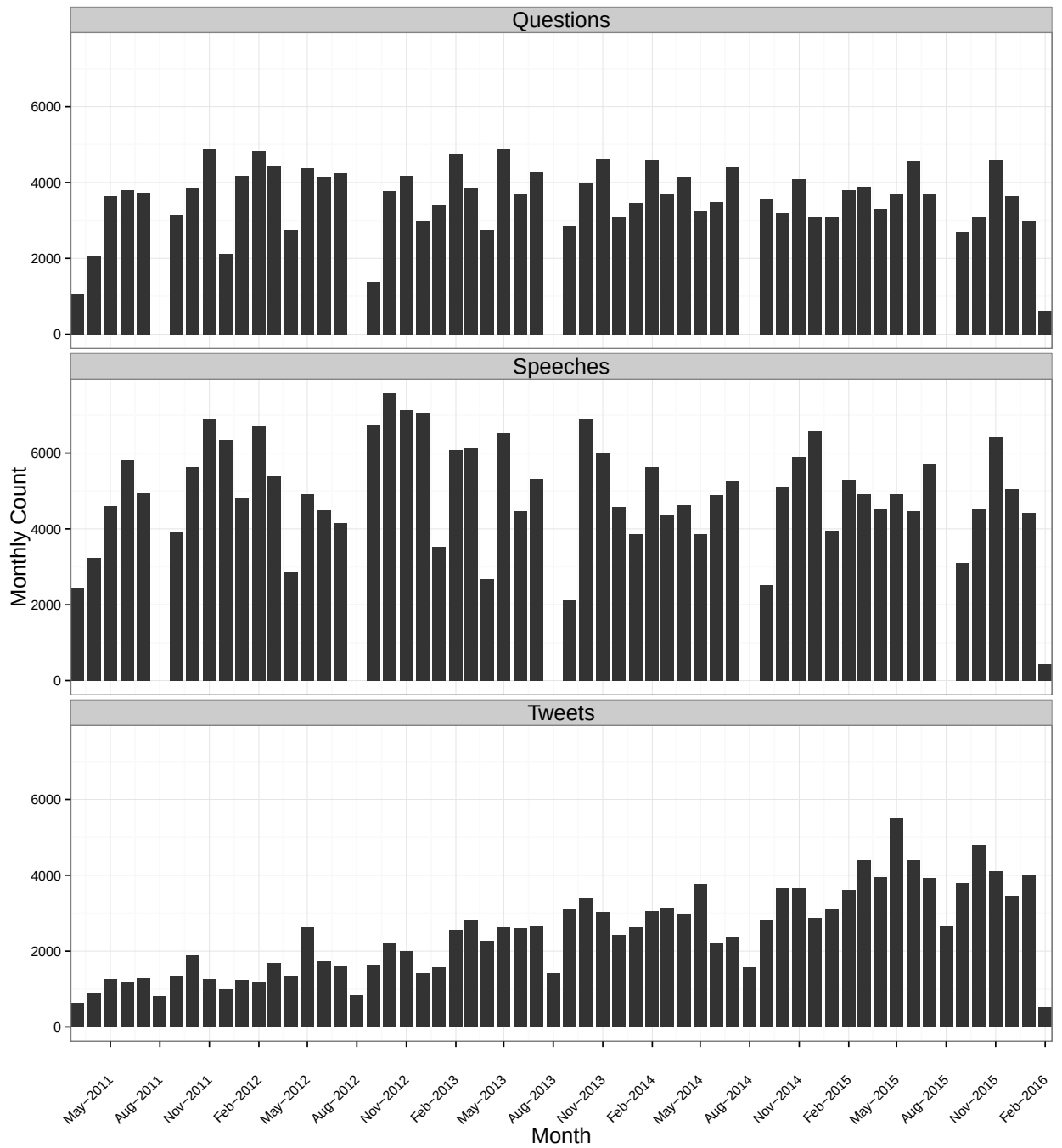
Figure 2.4 presents the variation in the total number of tweets posted by TDs

¹²Deputy Durkan was recently criticized for using governmental resources to produce and send over 1,000 Christmas cards to constituents (O'Connor, 2016).

during the 31st Dáil. This distribution is similar to that of speeches and questions, with a small number of TDs using this communication platform much more than their colleagues. Despite Twitter being in its infancy as a communication tool, we can see that multiple politicians write a significant number of tweets. For example, Aodhán Ó Ríordáin, a member of the government coalition partner Labour Party, posted 8,310 tweets in the duration of the 31st Dáil. This equates to him posting approximately 4 to 5 tweets every day. The mean number of tweets written by a legislator for this Dáil is 1,032, with the median TD tweeting 682 times. This variation suggests Twitter can also be a useful tool for understanding the activities of legislators.

A final way to descriptively compare legislators' legislative speeches, legislative questions, and tweets, is to examine the temporal coverage and variation of these data sources. In Figure 2.5, I present the total number of legislative speeches, legislative questions, and tweets by TDs in every month of the 31st Dáil. In line with the previous descriptive analysis, it is immediately noticeable that the number speeches made by TDs in a month is greater than the number of questions asked and the number of tweets posted. Another striking feature of Figure 2.5 are the recurring gaps for the questions and speeches panels for the month of August, which correspond to the parliament's summer recess. Although legislators still use Twitter during this period, there is a clear drop in the number of tweets posted. While the monthly counts of questions and speeches by TDs appear to be relatively constant across the time period, the number of tweets posted by TDs is linearly increasing over time. This suggests that legislators are increasingly utilizing this platform to broadcast information. This may result in tweets becoming a crucial data source for researchers interested in understanding the behavior and representational styles of legislators. From Figure 2.5 it is clear that all three data sources provide quite good temporal coverage of legislators' activities.

Figure 2.5: Number of speeches, questions, and tweets per month by TDs



2.3.2 Comparing the content of legislators’ speeches, questions, and tweets

I next compare the content of legislators’ speeches, questions, and tweets. To do this, I employ the informative Dirichlet prior model, also known as the “Fightin’ Words” model, developed by Monroe, Colaresi and Quinn (2009). This model enables the identification of words that are statistically overrepresented in speeches, question, and tweets as compared with the other platforms (Jurafsky and Martin, 2017).¹³ This model therefore can help us to understand whether legislators utilize different communication platforms to emphasise different activities and elements of representation.

Table 2.3: Top 50 words associated with legislators’ speeches

Word Category	Examples
Parliamentary Exchanges	<i>deputy, deputies, house</i>
Policy	<i>bill, legislation, amendment, people</i>
Rhetoric/Argument	<i>should, would, must, because, could, can, important, need, want, cannot, know</i>
Common/Stop Words	<i>are, have, this, not, we, that, it, is, to, do, what, but, be, as, they, was, there, some, am, about, many, however, one, also, very, does, our, which, them, were, those, had, said</i>

I first discuss the model comparing the content of legislative speeches to that of legislative questions and tweets. In Table 2.3, I present the top 50 words associated with legislators’ speeches based on the estimates of this model. From this table, we can see results of this model correspond to our prior expectations about the content

¹³I provide more details about this model in the technical details section of the Appendix.

of legislative speeches. These results suggest that a *deputy* often debates *bills* and *amendments* that affect the *people* of the country with other *deputies* in the lower *house* of the parliament. From Table 2.3, we can also see that the model correctly identifies strong differences between the oral language of speeches and written language of the more structured legislative questions, which I discuss in more detail below, and tweets. These results also provide support to the idea that speeches are primarily a tool for understanding policy responsiveness.

Table 2.4: Top 50 words associated with legislators’ questions

Word Category	Examples
Question Structure	<i>asked, minister, for, if, he, she, will, make, statement on, the, matter</i>
Casework/Service	<i>person, persons, county, details, supplied, community payment, application, scheme, skills</i>
Ministerial Accountability	<i>plans, regarding, position, reason, views provide, number, date, department</i>
Ministerial Titles	<i>tourism, marine, affairs, transport, health social, protection, equality, environment finance, education, local, sport agriculture, food, justice</i>
Common/Stop Words	<i>and, this, each</i>

In Table 2.4, I present the top 50 words associated with legislators’ questions from the model comparing the content of legislative questions to that of legislative speeches and tweets. The model correctly identifies that all questions share a similar structured format. In a submitted question, a deputy could have *asked the minister for agriculture and food/education/finance/social protection/etc.* about a *payment* or *application* of a *person* in their constituency or a *scheme* relevant to their *county*

or their *plans, position, or views* on a *department* policy, and *if he/she will make a statemnent on the matter*. That is, the results suggest that questions, as compared to speeches and tweets, and again in line with prior expectations, are utilized for casework/constituency service on behalf of constituents in addition to holding ministers and government departments accountable for policy. These results therefore provide support to the arguments that the content of parliamentary questions can be employed to study both service and policy responsiveness (Martin, 2011*b*).

Table 2.5: Top 50 words associated with legislators’ tweets

Word Category	Examples
Dates/Times	<i>morning, today, day, th, st, tonight, tomorrow night, now, night, week, evening, on, last</i>
Locations	<i>dail, here, dublin, cork, centre, at</i>
Events/Activities	<i>speaking, meeting, launch, campaign</i>
Current Events	<i>vote, irish, team, news</i>
Interaction	<i>please, thanks, you, your, me, my</i>
Positive Sentiment	<i>great, good, delighted, best, well, done</i>
Twitter Language	<i>re, via</i>
Common/Stop Words	<i>just, new, from, all, with</i>

In Table 2.5, I present the top 50 words associated with legislators’ tweets from the model comparing the content of tweets to that of legislative speeches and questions. The content of tweets differs substantially from the content of legislative speeches and questions and tweets are often used to inform fellow Twitter users of their activities. Legislators tweet about *speaking* or attending a *meeting at* locations such as the *dail* and taking part in *launch* and *campaign* events. The results also indicate that legislators utilize Twitter to discuss *news* and interact directly with other Twitter

users. It also appears that legislators' language is more positive in tweets than in formal legislative settings, perhaps due to an effort to cultivate a positive image.

Taken together the results of these content analysis models suggest that legislators utilize these communication platforms to perform different activities. Legislators appear to primarily utilize speeches to emphasize their policy preferences, questions appear to be utilized for both constituency service and policy responsiveness, while Twitter appears to be a platform which is primarily employed to advertise legislators' activities. These results fit neatly with the seminal model of representatives' activities proposed by Mayhew (1974), in which legislators motivated by re-election focus on the activities of advertising, position taking, and credit claiming. Although further systematic analyses of the content of legislators' speeches, questions, and tweets is clearly required, the results provided here suggest that each data source can help us better understand legislators' activities and could be a promising resource for testing hypotheses pertaining to the activities of position taking, credit claiming, and advertising respectively.

Chapter 3

Political Dynasties and Political Representation

“The most simple fundamental thing is that to be a national politician you have to be elected locally ... I care for the people in my constituency. I dearly do.” - Michael Healy-Rae (2016).¹

3.1 Introduction

In February 2016, Danny Healy-Rae, son of former TD Jackie Healy-Rae, and brother of incumbent South Kerry TD Michael Healy-Rae, announced his intention to join his brother in seeking election to Dáil Éireann in the newly expanded constituency of Kerry. During the campaign the brothers noted their father’s legacy of representing the south Kerry electoral area and pledged to “to expand the first-class political service that the Healy-Raes gave to south Kerry over many years” (Lucey, 2016). Later that month, the Healy-Raes secured over 38 per cent of first preference votes and were the first two candidates elected in the constituency. In doing so, the brothers became

¹Quote from (McGee, 2016).

the first siblings to simultaneously represent the same Dáil Éireann constituency. The Healy-Raes are just one of a number of high profile families in Irish politics.² Political dynasties, however, are not unique to Ireland and dynastic legislators are prevalent in the legislatures of many democracies.³ In countries such as India, Japan, and the Philippines, dynastic legislators have held over twenty per cent of the seats in the national legislature during a single legislative session (Chandra and Umaira, 2011; Smith, 2012; Asako et al., 2015; Mendoza et al., 2012), while approximately seven per cent of leaders in democracies since the mid-nineteenth century have been members of political dynasties (Besley and Reynal-Querol, 2015).

The prevalence of dynastic politicians in democracies has led political commentators to speculate on the political and economic impact of political dynasties. Political dynasties have been accused of destroying the economy of Greece and linked to a culture of cronyism in Ireland (Lopez, 2011; Turner and Crimmins, 2011). Dynasties are also cited as one of the primary reasons for governmental corruption and electoral violence in the Philippines (Conde, 2007), leading Filipino legislators to introduce a bill which proposes to outlaw relatives of legislators from serving in successive, simultaneous, or overlapping terms.

Given this attention to dynastic politicians, it is unsurprising that political scientists and economists have begun to analyze the causes and consequences of political dynasties in recent years. This research suggests that higher proportions of dynastic legislators are more likely to be found in parliamentary democracies with candidate-centered electoral systems and in political parties using decentralized candidate selection procedures (Smith, 2012). Regarding the emergence of political dynasties, existing studies suggest the longer that legislators serve the more likely they are to

²Other high profile political families in the Ireland are the Lenihans, the Haugheys, and the Springs (Fallon, 2011).

³I consider a legislator to be dynastic if a close relative has served in the same chamber of the legislature prior to their election. Close relatives are legislators who are related by blood or marriage such as parents, siblings, children, aunts, uncles, grandparents, sons-in-law, etc.

have a relative subsequently serve in the same chamber (Laband and Lentz, 1985; Dal Bó, Dal Bó and Snyder, 2009; Rossi, 2014; Querubin, 2016). Membership in a political dynasty also seems to provide candidates with an electoral advantage over non-dynastic candidates in certain situations (Feinstein, 2010; Smith, 2012). Specifically, dynastic candidates appear to have an advantage over first-generation candidates in open seat races due to ‘brand name advantages’ (Feinstein, 2010). It is also claimed that dynastic candidates benefit from an ‘inherited incumbency advantage’. That is, first time dynastic candidates are purported to inherit a considerable electoral advantage over non-dynastic candidates, similar to the electoral advantage that incumbents enjoy over challengers (Smith, 2012).⁴ Once in office, dynastic politicians whose relatives served in the cabinet also appear to have a greater chance of obtaining a cabinet position than their non-dynastic colleagues (Smith and Martin, 2016).

Turning to the economic and political consequences of dynasties, there is disagreement about the economic impact of dynastic politicians. While some research finds that dynastic politicians hinder economic outcomes due to an inefficient use of resources (Asako et al., 2015; Braganca, Ferraz and Rios, 2015), other work concludes that dynastic leaders actually improve economic performance when the constraints on the executive are weak (Besley and Reynal-Querol, 2015). Focusing on the political impact of dynasties, a potential unintended consequence of dynasties is that they appear to have a positive influence on the descriptive representation of females (Geys and Smith, 2016). Dynastic membership helps female candidates to be selected by parties and overcome disadvantages typically experienced by female candidates in elections (Folke, Rickne and Smith, 2016; Labonne, Parsa and Querubin, 2015).

In this growing literature remarkably little attention has been paid to dynastic legislators’ representation of constituents.⁵ Do dynastic and non-dynastic politicians

⁴‘Brand name advantages’ are again cited as the primary mechanism driving this effect.

⁵Rahman (2013) finds a negative association between membership in a dynasty and the parliamentary attendance of legislators in Bangladesh.

have different representation styles? Understanding how dynastic legislators represent constituents is important as it can help explain the survival of political dynasties in democratic nations, it sheds light on dynasties impact on the quality of representation in democracies, and it expands our understanding of the normative consequences of large proportions of dynastic legislators. This paper contributes to the literature on dynasties by investigating whether there are systematic differences in how dynastic and non-dynastic politicians represent constituents.

I expect that dynastic legislators will pay particular attention to service responsiveness. Politics occupies a central role in the home life of political families and members of political families are therefore likely to experience a political socialization that others bereft of similar ties will typically not experience. I anticipate that this involvement in politics will allow them to develop social ties with constituents in their district prior to their political career. I also assume that politicians who have relatives follow them into office benefit from a substantial personal vote and these politicians emphasize the importance of the personal vote for succeeding in politics. I additionally posit that the social network associated with the dynasty will help voters to more easily request service from dynastic legislators. I therefore expect that dynastic legislators will prioritize and engage in considerable constituency service. Given that personal connections to politicians and constituency service are more prevalent in rural areas, I expect that the influence of membership in a dynasty on service responsiveness is greatest in more rural geographic regions.

I examine the empirical relationship between membership in a political dynasty and service responsiveness using an original dataset, that I constructed, comprising the dynastic ties of legislators and written questions submitted by legislators to ministers in Ireland from 1973 to 2007. Ireland is a good case to investigate the relationship between membership in a political dynasty and the service provided by legislators as, similar to other countries with substantial proportions of dynastic legislators, it has a candidate-centered electoral system that promotes the cultivation

of a personal vote. The results of my analysis indicate that dynastic legislators in rural areas engage in more constituency service than non-dynastic legislators in these areas, although, somewhat surprisingly, there is little difference in the service responsiveness of dynastic and non-dynastic legislators in urban areas. This indicates that different mechanisms may underlie the formation and persistence of dynasties in rural and urban regions.

3.2 Dynastic Legislators and Representation

The relationship between voters in a constituency and their legislator can be understood within a principal-agent framework, with voters assuming the role of principal and the legislator as their agent (e.g. Ferejohn, 1986). Voters in a constituency elect a politician to engage in legislative behavior which they collectively want. In the event that the performance of legislators does not meet the expectations of constituents, then these legislators can be punished and voted out of office by the constituents at election time (Fiorina, 1981; Ferejohn, 1986). A legislator's electoral connection to their constituents is therefore of critical importance for re-election (Mayhew, 1974; Fenno, 1978). I therefore assume that retaining office is the immediate goal of legislators and legislators employ certain strategic behaviors in pursuit of this goal (Downs, 1957; Mayhew, 1974; Carson and Jenkins, 2011).

The personal vote is a key concept in the electoral connection between legislators and their constituents. The personal vote can be defined as the proportion of the vote received by a legislator attributable to their personal qualities, record, and activities as opposed to the proportion of the vote received due to partisan affiliation (Cain, Ferejohn and Fiorina, 1987; Carey and Shugart, 1995). The existing literature proposes that the personal vote is based on certain types of legislative behavior such as casework and allocations that benefit constituents and a legislator's constituency (Mayhew, 1974; Fiorina, 1989; Carey and Shugart, 1995; Ansolabehere, Snyder Jr and

Stewart III, 2000). Research attempting to decompose the sources of the electoral advantage that incumbents have over challengers has noted that a substantial portion of the incumbency advantage can be credited to the personal vote of a legislator (Levitt and Wolfram, 1997; Ansolabehere, Snyder Jr and Stewart III, 2000; Carson, Engstrom and Roberts, 2007). Specifically, this research estimates that the personal vote accounts for between one-half to two-thirds of the electoral advantage enjoyed by incumbents.

Although research on dynastic politicians is still in its infancy, existing work suggests that legislators with successors typically have longer tenures in office and are less likely to be defeated in an election than legislators that do not have a relative follow them into office (Laband and Lentz, 1985; Smith and Martin, 2016). Similarly, research has shown the longer the tenure of a legislator the greater the probability that they will have a relative subsequently serve in the same legislature (Dal Bó, Dal Bó and Snyder, 2009). Given that legislators who had relatives subsequently obtain legislative office are likely to have served for multiple terms, I assume that such legislators cultivated and benefited from a personal vote, engaged in significant constituency service, and are trusted by their constituents.⁶

In the family of a politician, politics is very likely to occupy a central role in the home.⁷ Dynastic politicians have noted that politics was discussed every day in the household during their formative years (Collins, 1993). Individuals growing up in such an environment will be exposed to political episodes that others bereft of similar relations will typically not experience.⁸ Relatives of a legislator will likely attend the

⁶Unfortunately, I do not have sufficient data to test this assumption. However, anecdotal evidence regarding the legislative behavior of dynastic legislators supports this assumption. For instance, Breeda Moynihan-Cronin, an Irish dynastic legislator whose father also served in the Irish Dáil, noted that her wedding had to be delayed as her father was running late due to meetings with constituents (O’ Shea, 2011).

⁷To be clear, this is true for all politicians and not just the case for dynastic politicians.

⁸There will be exceptions to this. For instance, politics is likely to be central in the household of a trade union leader. However, in general, compared to non-dynastic legislators who do not grow up in the household of a politician, dynastic legislators will likely have a relatively unique socialization.

local meetings of party organizations, accompany the legislator to events, canvass in electoral campaigns, and regularly witness the legislator interacting with constituents both inside and outside their home (O’ Shea, 2011).⁹ This active involvement in politics means that it is likely that close relatives of politicians will also cultivate strong ties with individuals in their constituency from a relatively young age.

Dynastic politicians are assumed to inherit brand name advantages and local political connections from their family ties (Dal Bó, Dal Bó and Snyder, 2009; Feinstein, 2010; Smith, 2012). It is also probable that members of political families are the recipients of a ‘political education’ from a very young age (Kurtz, 1989). Given that politicians who have relatives follow them into the legislature have above average tenures in office, I assume that these individuals understand how to succeed in politics and that this knowledge will be transferred to dynastic legislators. The senior family members will likely advise their relatives about actions and behavior that they perceive to be beneficial for their political career. Given their relatively lengthy tenures in office, I assume that these individuals will emphasize the importance of the constituency service for maintaining constituents’ trust and remaining in office.

I also assume that a political dynasty makes it easier for constituents to receive service from a legislator. Typically, before a legislator can provide service to a constituent, the constituent needs to contact the legislator or somebody closely connected to the legislator and request assistance (Cruz, Labonne and Querubin, 2017). The relative of a dynastic legislator that previously obtained office will likely have built a broad social network that includes many constituents during their tenure. In addition to the social ties cultivated by the dynastic legislator themselves, I expect the brand name, trust, and broad social network associated with a political dynasty to assist constituents make requests and receive service from dynastic legislators.

⁹Anne McEllistrim, sister of Tom McEllistrim Jr and daughter of Tom McEllistrim Sr both former legislators in the Irish Dáil, has noted that ‘queues of people’ would regularly be sitting in the living room of the family home at nighttime waiting to speak to her father (O’ Shea, 2011).

From the discussion above, I expect that dynastic legislators will pay considerable attention to service responsiveness. Politics occupies a central role in the formative years of dynastic legislators. I posit that this exposure to political life enables dynastic legislators to develop social ties with constituents in the district from a young age. I also expect that advice from previously elected relative places an emphasis on the provision of service to constituents as key to succeeding in politics and the social network associated with a dynasty provides constituents with greater access to service. I propose that due to political socialization processes and the social networks of political dynasties, dynastic legislators will engage in substantial constituency service activities.

Building on the literature highlighting considerable variation in the politics of urban and rural areas, I also expect the activities of dynastic legislators to be influenced by whether they serve in an urban or rural area. A number of works have shown that there are substantial differences in the political processes of urbanized and rural regions (Fenno, 1977; Clarke, 1978; Ellickson and Whistler, 2001; Monroe and Rose, 2002). Rural areas are typically more culturally homogeneous and less affluent than urban areas, although urban areas also often contain large numbers of low income individuals. Social networks in rural areas also tend to be far denser than urban social networks, with kin playing a more central role in rural regions (Höllinger and Haller, 1990; Komito, 1992). It is therefore more likely that constituents will have personal connections with a legislator in a rural constituency. In line with this, existing evidence indicates that legislators in rural areas provide more constituency service than legislators in urban areas (Clarke, Price and Krause, 1975; Freeman and Richardson Jr, 1996; Ellickson and Whistler, 2001). Consequently, I anticipate that the influence of dynastic membership on service responsiveness to be especially strong in rural areas.

I therefore hypothesize that dynastic membership is positively related to legislators' attention to constituency service, and this relationship is strongest in rural

geographic regions.

3.3 Research Design

In this section I explain why I examine the service responsiveness of legislators in Ireland, describe my data, outline my strategy for measuring service responsiveness, and specify my empirical model.

3.3.1 Ireland

Ireland is a good country to investigate the relationship between membership in a political dynasty and the localistic nature of legislators' activities. For the past forty years approximately twenty five per cent of legislators in Dáil Éireann, the lower house of the Irish parliament, have had a relative serve in the Dáil prior to their own election. Figure 3.1 plots the proportion of dynastic legislators in Dáil Éireann across time. The high proportion of dynastic legislators in the Irish parliament helps determine whether any systematic differences exist in the activities of dynastic and non-dynastic legislators. Similar to other countries with high proportions of dynastic legislators, Ireland has a candidate centered electoral system which encourages the cultivation of a personal vote. The single transferable vote system (STV) used in Ireland enables voters to choose between both parties and candidates. This system provides voters with a high degree of freedom in selecting between candidates on the ballot and multiple candidates from the same party often compete in the same constituency (Marsh, 2007).¹⁰ Legislators in Ireland therefore spend a significant amount of time engaging in constituency service and actively campaign for personal preference votes (Gallagher and Komito, 2006). By focusing on Ireland we can obtain a clear idea of the influence of membership in a political dynasty on the service responsiveness of

¹⁰A more extensive discussion of STV can be found in the section dedicated to this chapter in the Appendix.

legislators in parliamentary democracies with a high proportion of dynastic legislators.

3.3.2 Data

To examine the empirical relationship between membership in a political dynasty and service responsiveness, I have constructed an original dataset comprising the dynastic ties of legislators that served in Dáil Éireann and written questions submitted by legislators to ministers between 1973 and 2007. There are ten Dáil Éireanns in this period. The maximum term length of a Dáil is five years.

Figure 3.1: Dynastic legislators in Dáil Éireann over time

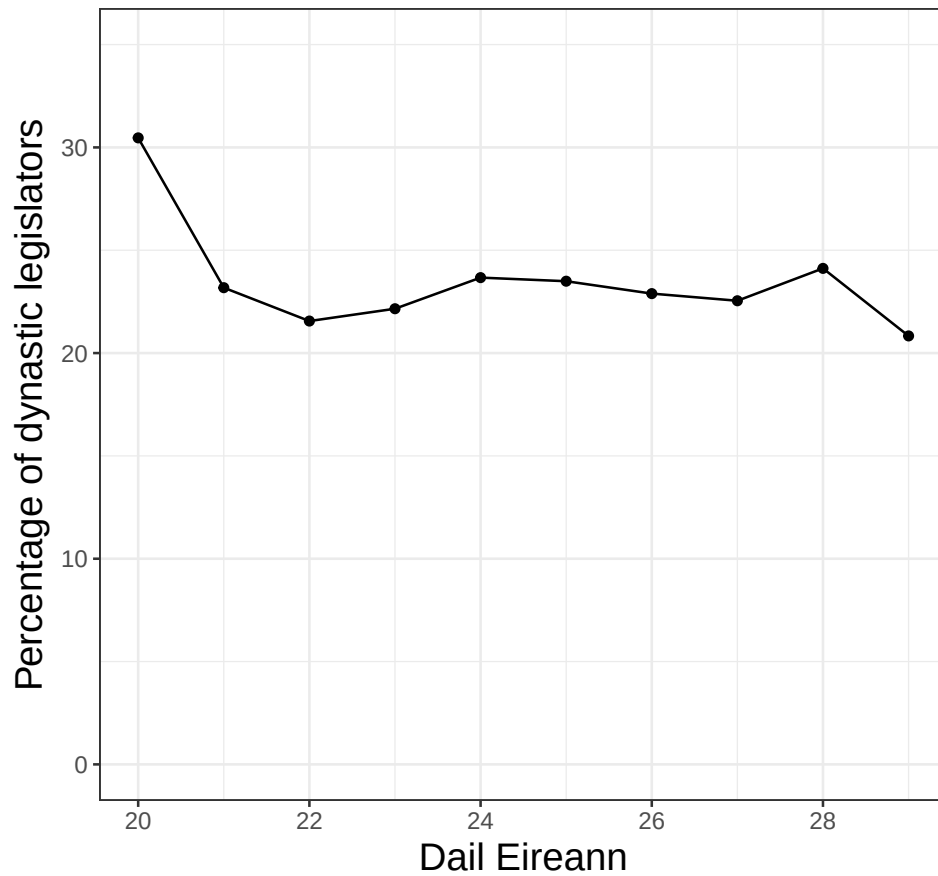
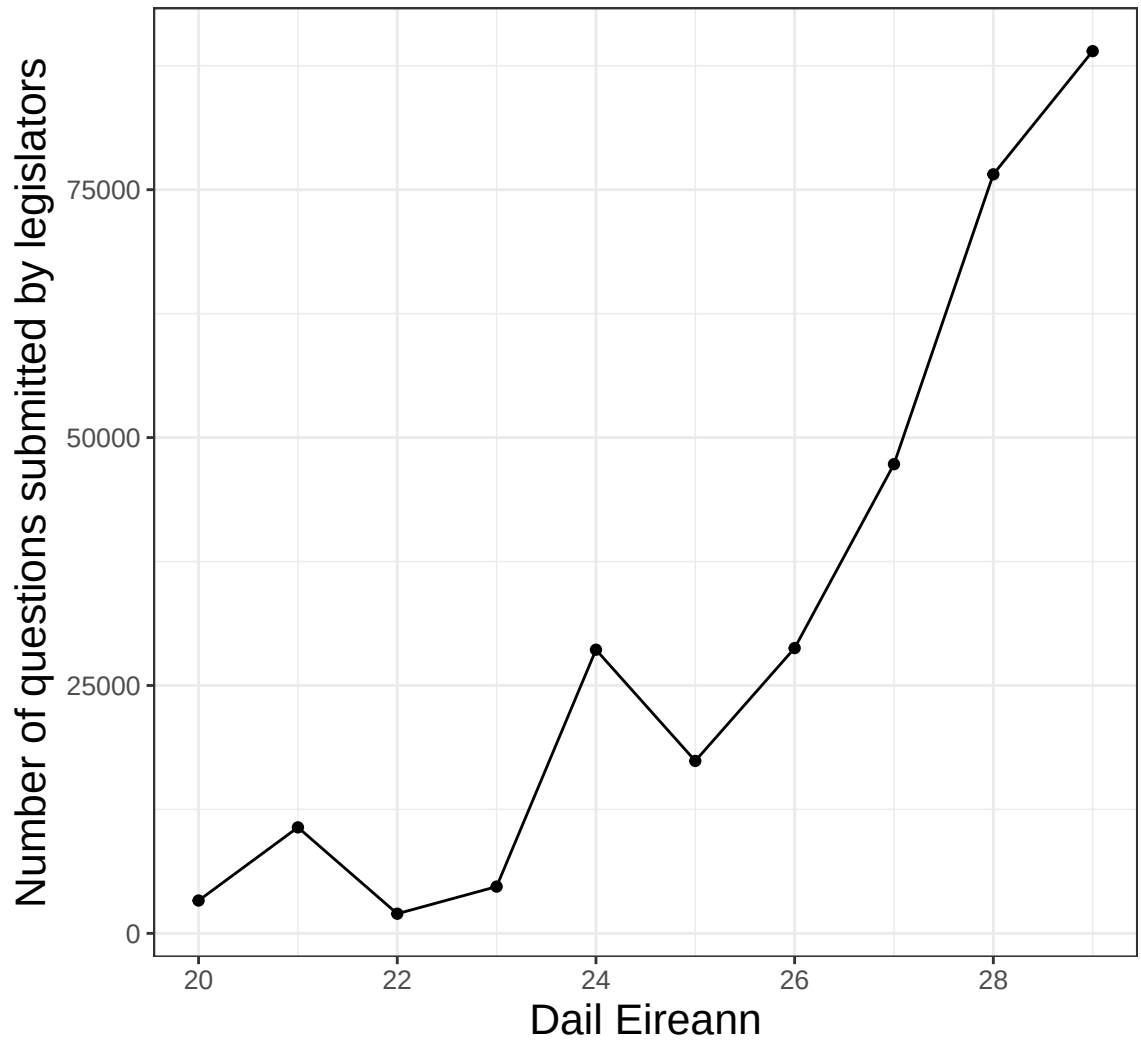


Figure 3.2: Number of questions submitted by legislators across Dáil Éireanns, 1973-2007



As noted earlier in this work, parliamentary questions are an important feature of the majority of national legislatures (Norton and Wood, 1993; Martin, 2011a). As shown in Chapter 2, parliamentary questions are employed by both government and oppositions legislators to question ministers in a relation to a mixture of local constituency and national issues. Compared to other forms of recorded legislator

behavior, and given the lack of constraints on the content of submitted questions, parliamentary questions are particularly useful for understanding the primary concerns of legislators. Analysis of parliamentary questions helps to measure both the policy and service responsiveness of legislators and can therefore improve our understanding of the preferences and role orientation of legislators. Between 1973 and 2007 Irish legislators submitted approximately 308,000 written questions to ministers. In Figure 3.2, I present the number of questions submitted by legislators across Dáil Éireanns.

3.3.3 Measuring service responsiveness

To measure the service responsiveness of legislators, I use a supervised machine learning approach to classify whether a legislator is using a question for constituency service or some other purpose (Grimmer and Stewart, 2013). In this approach, I use 2,500 hand coded documents as a training set. The training documents were drawn using a stratified random sample where the Dáil Éireanns served as the strata. To code whether a question constitutes constituency service, I follow the rules used in previous research utilizing questions to measure the constituency service behavior of legislators by Martin (2011*a,b*). This work classifies a question as constituency service if one or more of the following conditions are met:

1. The legislator refers to their geographic constituency in the question.
2. The legislator refers to a location that is in the legislator’s constituency in the question.
3. The legislator refers to a constituent or an individual who we can assume to be a constituent in the question.
4. The legislator refers to a building or facility that is in the legislator’s constituency in the question.

5. The legislator refers to an organization or business that is in the legislator’s constituency in the question.
6. The legislator refers to an event that is occurring in the legislator’s constituency in the question.

This hand coded training data is used to learn about the relationship between the content of questions and the outcome of interest, whether a question constitutes constituency service or not (Grimmer and Stewart, 2013). I then utilize what is learned in this training process to classify the remaining unlabelled questions in the dataset as constituency service or not.

I represent each document as a bag of words. This representation ignores the order of words and simply counts the frequency of term occurrences in a document. Each document in the training set is represented as a row, while each word that occurs in the training set is represented as a column. Each cell therefore represents the count of a given word in a given document.¹¹ To statistically model the relationship between the words and the outcome classes of the training documents, I use three different statistical methods which are commonly used for text classification, the least absolute shrinkage and selection operator (LASSO) logistic regression model, the multinomial naive Bayes classifier, and the support vector machine classifier (Grimmer and Stewart, 2013; D’Orazio et al., 2014; Minhas, Ulfelder and Ward, 2015; Jurafsky and Martin, 2017).¹²

As each of these models have tunable hyperparameters which can impact their predictive accuracy, I further randomly split the 2500 hand coded documents set into two sets consisting of 2000 documents and 500 documents respectively. I use the larger

¹¹There is a total of 5,913 unique terms in the 2,500 documents in the training set. Consequently, the resulting matrix created from these documents consists of 5,913 columns and 2,500 rows.

¹²Further information about the LASSO logistic regression model is provided in the technical details section of the Appendix. See Grimmer and Stewart (2013) for more details on the naive Bayes classifier and D’Orazio et al. (2014) for more information on the support vector machine. See Hastie, Tibshirani and Friedman (2009) for an excellent introduction to machine learning methods.

set to select the hyperparameters that optimize predictive classification accuracy, while using the smaller set to determine the out-of-sample predictive performance. To tune the hyperparameters, I utilize 10-fold cross-validation which also helps to prevent overfitting. Overfitting refers to the case where a model has high predictive accuracy on the data used for training, but predicts poorly on new data. In 10-fold cross-validation, the data is randomly split into 10 approximately evenly sized subsets. Nine of these subsets are used to estimate a model and the predictions for the held out validation subset are generated from the estimated parameters of this model. This process is then repeated until all of the 10 subsets have been used as a hold out validation set (Ward, Greenhill and Bakke, 2010; Hill and Jones, 2014).

Table 3.1: Results of constituency service classification model

Model	Accuracy	Precision	Recall	F1-Score
LASSO Logistic Regression	0.91	0.92	0.84	0.88
Multinomial Naive Bayes	0.91	0.88	0.85	0.87
Support Vector Machine	0.91	0.88	0.85	0.87

I present the results of this classification task in Table 3.1. This table presents a variety of commonly employed predictive performance metrics calculated on the 500 held out documents that were not used in the cross-validation model selection procedure described above (Beger, Dorff and Ward, 2016; Jurafsky and Martin, 2017; Cranmer and Desmarais, 2017). The first metric in the table, accuracy, represents the proportion of questions that were correctly classified by the corresponding model. The second metric, precision, represents the proportion of the questions that were classified as constituency service in the hand coding that were also correctly classified as constituency service by the model. Out of the total number of questions classified as constituency service by the model, the third metric, recall, represents the proportion of the questions that are correctly classified as constituency service. The final metric,

F1-score, is the mean of the precision and recall metrics. From Table 3.1, we can see that all models have quite good predictive accuracy, with the LASSO logistic regression model slightly outperforming the multinomial naive Bayes and support vector machine classifiers. As the LASSO logistic regression model has the highest F1-score on the hand coded test data, I utilize the estimated parameters of this model to classify the remaining unlabelled questions in the corpus.

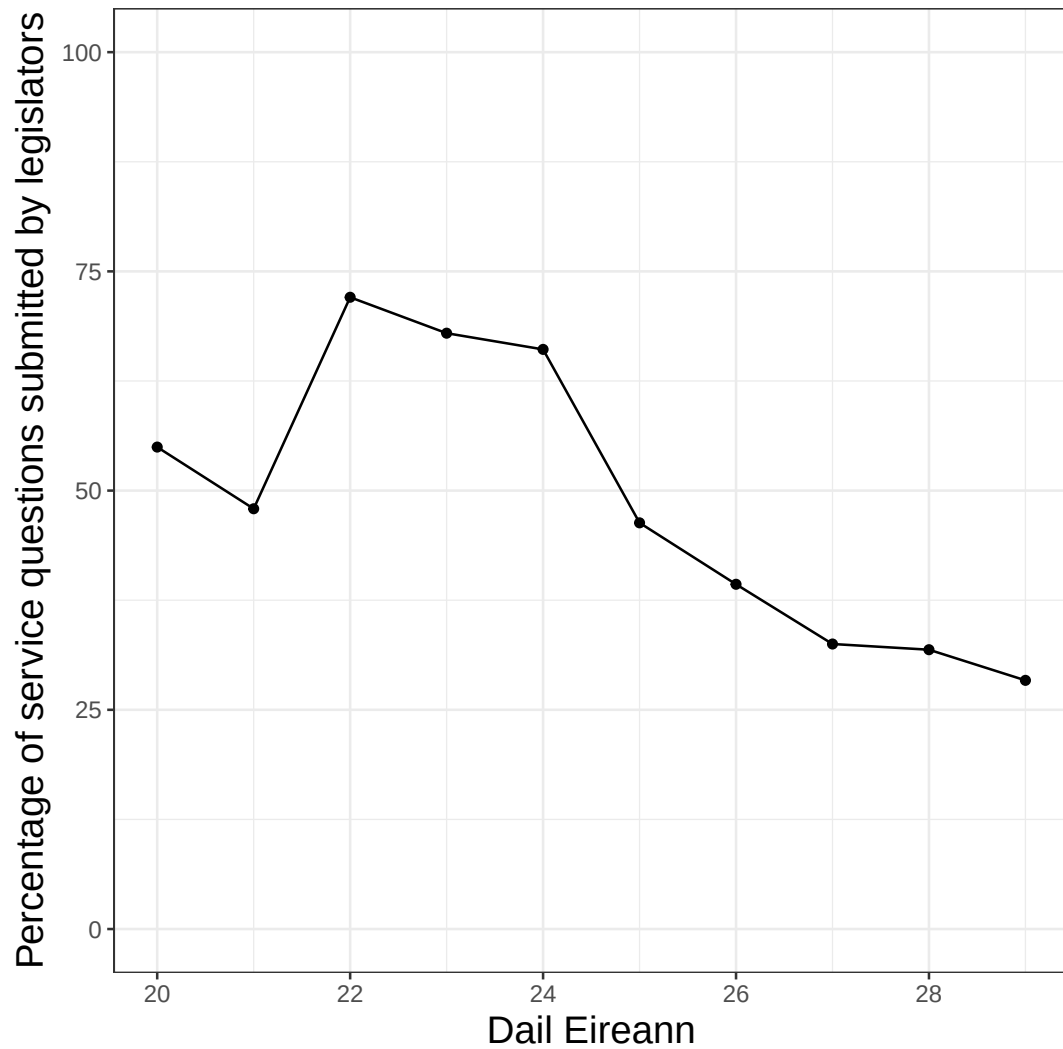
Table 3.2: Coefficients for most important words from LASSO logistic regression

<i>Constituency service</i>		<i>Not constituency service</i>	
Coefficient	Word	Coefficient	Word
2.35	dublin	-0.78	ireland
2.32	county	-0.74	irish
1.88	cork	-0.68	introduce
1.76	person	-0.63	pupils
1.34	supplied	-0.63	or
1.33	galway	-0.63	recent
0.99	waterford	-0.62	act
0.80	donegal	-0.62	basis
0.74	sanction	-0.60	company
0.73	mayo	-0.58	carlow
0.72	wexford	-0.57	board
0.71	extension	-0.56	sport
0.69	city	-0.56	educational
0.69	kerry	-0.55	asked
0.66	position	-0.54	marine
0.66	tralee	-0.54	reports
0.62	house	-0.54	registration
0.61	station	-0.53	affairs
0.60	road	-0.53	minister
0.60	hospital	-0.50	foreign

To provide more clarity about how the LASSO logistic model classifies questions, I present the 20 words with the largest negatives coefficients and the 20 words with the largest positive coefficients in Table 3.2. In this context, coefficients indicate how a unit increase in the use of a word in a question influences the likelihood of a

question being classified as constituency service. We can see that words most strongly associated with the constituency class are county and city names such as *county*, *dublin*, *cork*, *galway*, *waterford*, *donegal*, *kerry*, terms that refer to local buildings or infrastructure such as *house*, *station*, *hospital*, *road*, and the terms *person* and *supplied* which are typically used when a legislator is making a query related to a case of an individual constituent. We can also see that the words that are most strongly associated with the “not constituency service” class include the terms *ireland*, *irish*, which are typically used when discussing national issues, and the terms *foreign* and *affairs* which are used in queries pertaining to international issues.

Figure 3.3: Number of service questions submitted by legislators across Dáil Éireanns, 1973-2007



Of the unlabelled questions, 113,918, or 37 per cent, are classified as constituency service, while 191,886, or 63 per cent, of these questions are not classified as constituency service. In Figure 3.3, I break this down further by plotting the percentage of total questions that are classified as service responsiveness in each Dáil Éireann.

3.3.4 Model specification

To understand the influence of membership in a political dynasty on service responsiveness, I use the count of the service questions submitted by a legislator in a Dáil Éireann as my outcome variable. The key predictor of interest is whether a legislator is a member of a political dynasty or not. I classify a legislator as dynastic if a close relative has served in the same chamber of the legislature prior to their election. Close relatives include parents, siblings, children, aunts, uncles, grandparents, sons-in-law, etc. Information on the dynastic status of legislators was determined from the biographical sketches of legislators contained in the Nealon’s Guide to the Dáil series published after each Dáil Éireann election (Nealon, 1973, 1977, 1981, 1982, 1983, 1987, 1993, 1997; Kennedy, 2002). I then verified these family ties using the Elections Ireland website. As it is not possible to obtain constituency level estimates of urbanization, to measure potential differences in the service responsiveness of legislators serving in urban and rural regions, I employ an indicator capturing whether the legislator serves in the Dublin area, by far the largest urban area in Ireland. As noted by Komito (1989) “politics in Dublin differs considerably from rural politics”. Furthermore, statistics from the census of Ireland in 2011 demonstrate that the population density of Dublin is nearly three times greater than that of any other city or town in the country and seventeen of the twenty most densely population areas in Ireland lie in the Dublin region (Central Statistics Office Ireland, 2011).¹³

I also account for a set of other covariates that are likely predictors of legislators’ service responsiveness. As questions are predominantly employed by opposition legislators, I include an indicator variable representing whether a legislator is a member of the government or a member of the opposition. To account for potential differences in the service responsiveness of male and female legislators, I also include an indicator

¹³Population density measures the number of people per square kilometre in a region. The population density of the Dublin region is 3,498 while the average population density of other cities in Ireland is 1,311. The population density of Ireland is 67 (Central Statistics Office Ireland, 2011).

variable denoting the gender of a legislator. I further include a dichotomous variable capturing whether a legislator is affiliated with Fianna Fáil, the dominant party in Irish politics in the period of interest. Lastly, to account for the role that electoral performance can have on legislators' propensity to engage in constituency service, I include the percentage of first preferences that legislators received in the preceding general election for a given Dáil. Data on these predictors was also obtained from the *Nealon's Guide* series.

I estimate negative binomial models of the counts of the constituency service questions submitted by legislators. To account for variation in the total number of questions submitted by different TDs I include the log of the total number of questions submitted by a legislator in a Dáil Éireann as an offset variable. For some of the model specifications, I use a multilevel regression framework to estimate the relationship between membership in a political dynasty and the service responsiveness of a legislator. In these specifications, to help account for potential variation in service responsiveness across Dáil Éireanns, potential variation across constituencies, and potential variation across individual legislators, I also employ parliament, constituency, and legislator random effects.

3.4 Results

Before discussing the results of the various models outlined above, I first descriptively explore the conditional relationship between membership in a dynasty and service responsiveness. In Figure 3.4, I present the proportion of constituency service questions for both dynastic and non-dynastic legislators that serve in Dublin as well as in other geographic areas in the country. From this plot, we can see that the proportion of constituency service questions submitted is highest for the category representing dynastic legislators that serve in less urbanized areas. It is also noticeable that the proportion of constituency service questions submitted appears lowest for the cat-

egory representing dynastic legislators in Dublin. This basic comparison suggests that dynastic legislators in less urbanized areas engage in more constituency service than all other legislators, while, contrastingly, dynastic legislators in Dublin appear to provide slightly less service than Dublin's non-dynastic legislators.

Figure 3.4: Membership in a dynasty and the proportion of constituency service questions asked

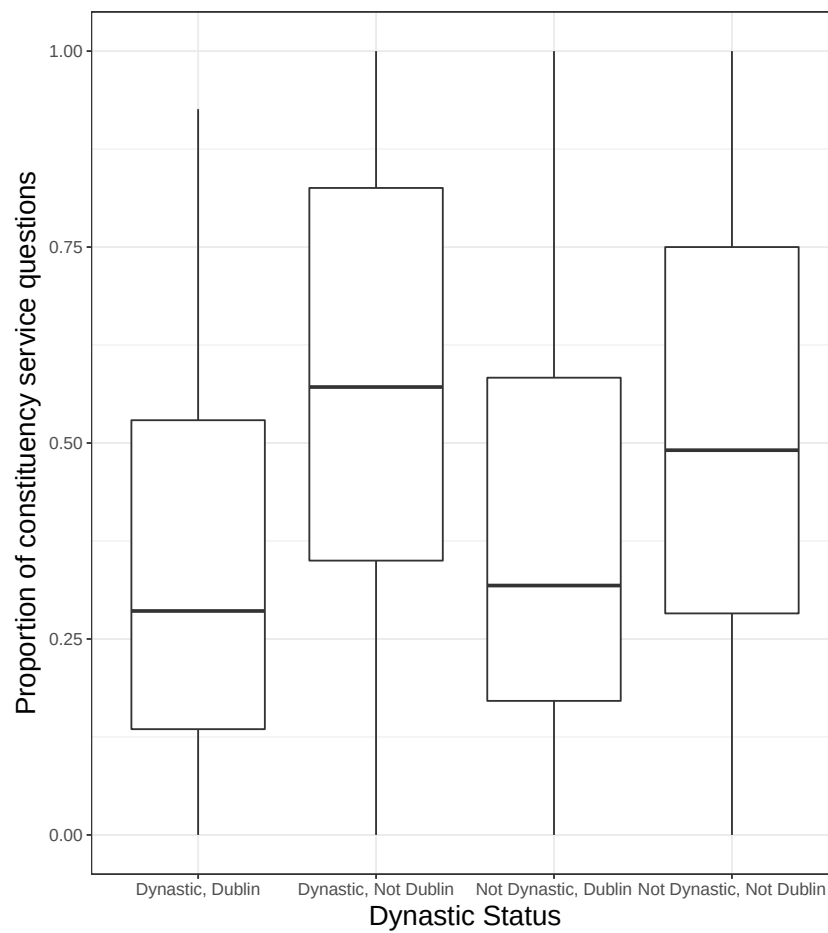
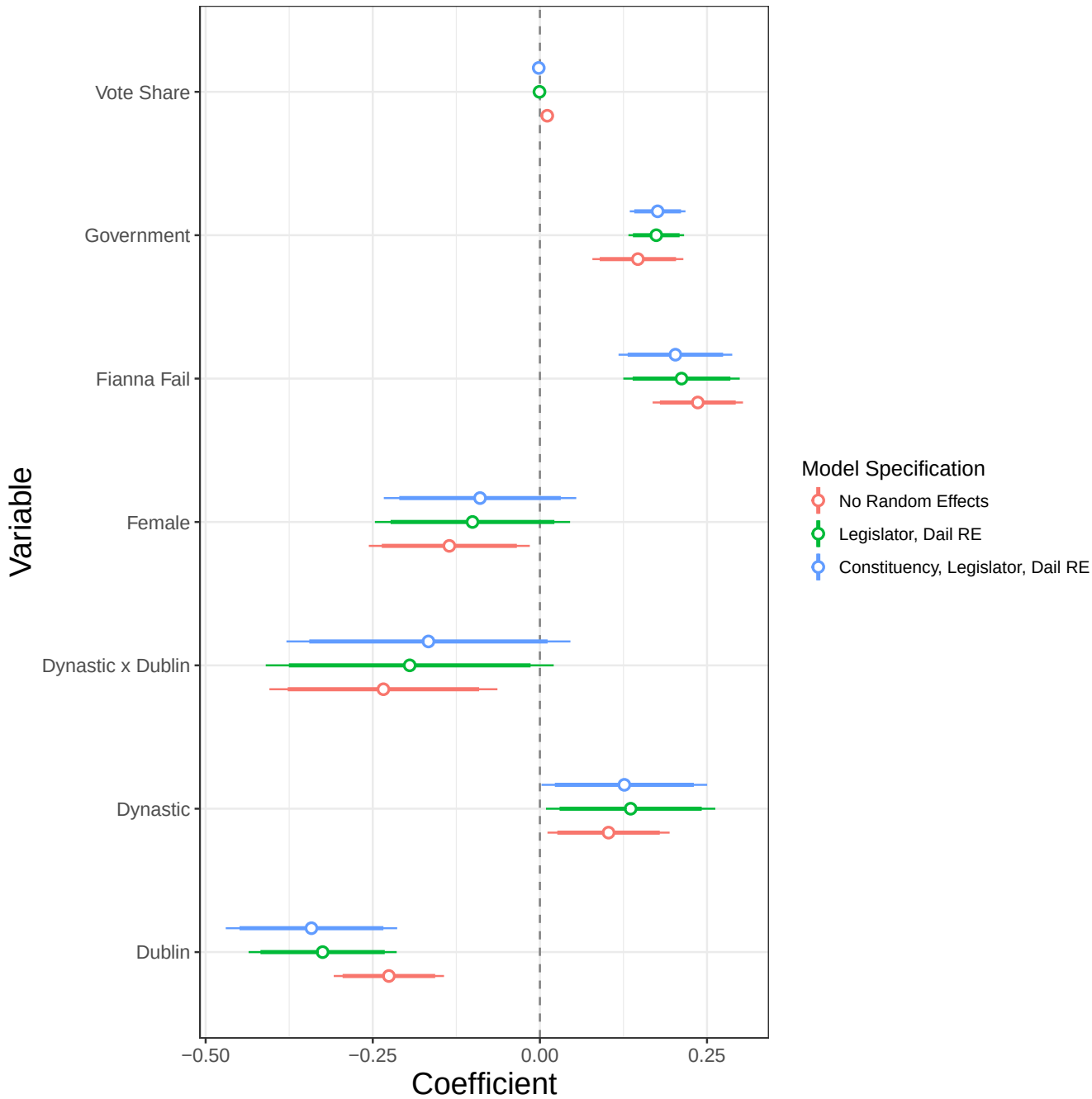


Figure 3.5: Dynastic membership and service responsiveness in Ireland 1973-2007



I present the results of the various estimated negative binomial count models in Figure 3.5.¹⁴ Given the inclusion of an interaction term in these models, when interpreting the coefficients pertaining to dynastic membership and geographic location, it is important to remember that the baseline category is non-dynastic legislators serving outside of Dublin. In line with the previous examination of the conditional relationship between dynastic membership and legislators' service responsiveness, we can see that the coefficient indicating a legislator is a member of a political dynasty is positive and different from zero at the conventional level of statistical significance. This indicates that there is a positive relationship between dynastic membership and constituency service in more rural regions. We can also see that the coefficient for the interaction term is negative although the coefficient is not consistently different from zero at the conventional level of statistical significance across model specifications. This suggests that dynastic legislators do not ask more constituency service questions than non-dynastic legislators in more urbanized areas. We can also see that the coefficient for the Dublin geographic indicator is quite large and negative. This suggests that legislators in more rural areas generally pay greater attention to service responsiveness.

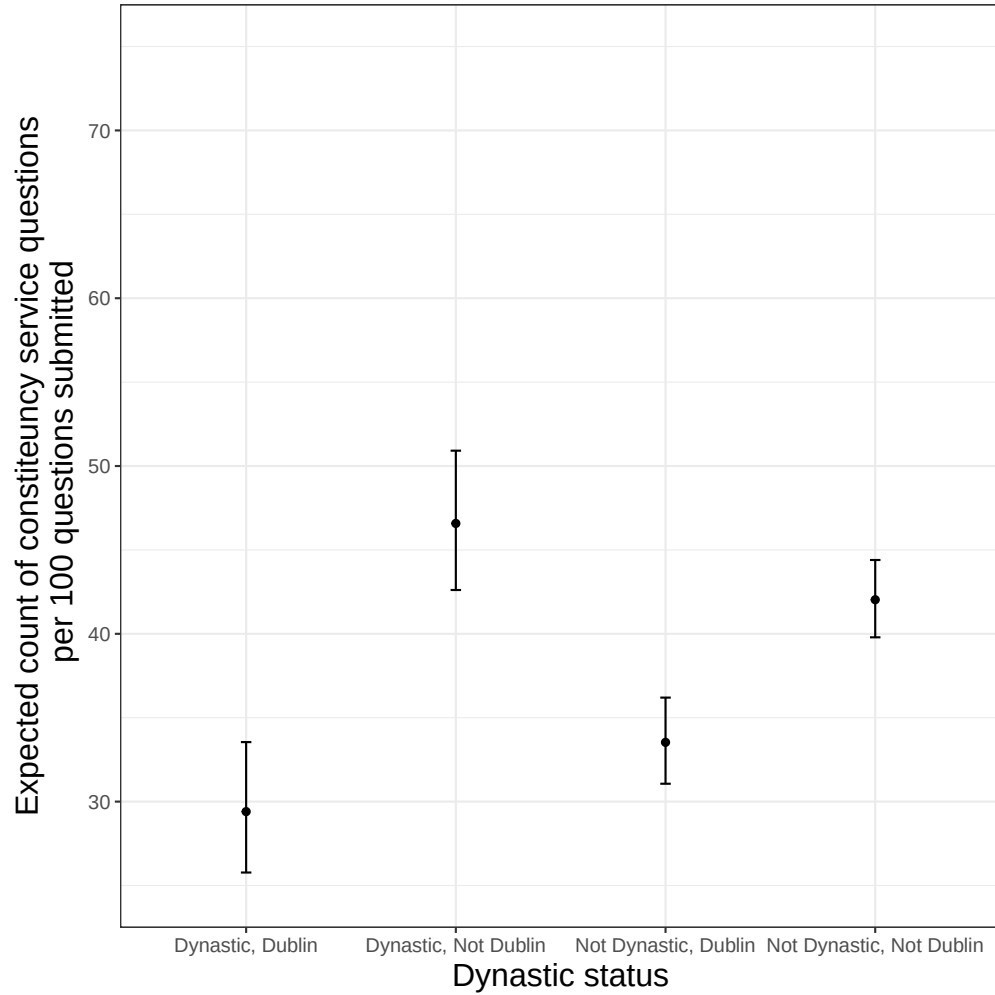
To increase the interpretability of these results, I also calculate the expected counts of constituency service questions for every 100 questions asked for both dynastic and non-dynastic legislators serving in the Dublin region and less urbanized areas.¹⁵ For these estimates, I set the other indicator variables at their modal values and set the continuous vote share measure to its mean value. I present these expected counts in Figure 3.6. From this plot, it is clear that the expected count of constituency service questions submitted by dynastic legislators in less urbanized areas have a higher expected count of constituency service questions than any other group of legislators.

¹⁴A table of coefficients and other model statistics can be found in the section dedicated to this chapter in the Appendix.

¹⁵The expected counts were generated using the estimated parameters from the negative binomial model specification that did include any random effects.

More specifically, for every 100 questions submitted by a dynastic legislator serving in a more rural region, we would expect that, on average, approximately 47 of these questions pertain to constituency service. On the other hand, for every 100 questions submitted by a non-dynastic legislator serving in a more rural region, on average, approximately 42 of these questions are constituency service queries. We can also see that for every 100 questions submitted by a dynastic legislator in Dublin, approximately 30 of these questions are constituency service queries, while a non-dynastic legislator in Dublin has an expected count of 33 service questions.

Figure 3.6: Expected counts of constituency service questions



These results suggest that dynastic legislators in less urbanized regions invest greater effort in service activities than their non-dynastic counterparts. Somewhat surprisingly, it also appears that dynastic legislators in Dublin engage in similar levels of constituency service to non-dynastic legislators in Dublin. This indicates that while constituency service may play an integral role for the reputation of dynasties in more rural areas, this is not the case for dynasties in highly urbanized areas. It therefore maybe the case that the brand name recognition of political dynasties in

highly urbanized areas may be developed and maintained through different mechanisms than in more rural areas. One possibility is that instead of focusing on service responsiveness, dynastic politicians in highly urbanized areas focus on certain policy issues and/or provide substantial allocations to constituents.

Turning to the other predictors included in the estimated models, we can see that the coefficient for the gender indicator is negative, indicating that female legislators engage in less service responsiveness than their male counterparts. However, the coefficient is not consistently statistically different from 0 across different model specifications. Consequently, gender doesn't seem to be a particularly important predictor of service responsiveness. Contrastingly, the indicator representing a legislators' affiliation to the Fianna Fáil party is positively associated with service responsiveness, indicating Fianna Fáil legislators have a higher propensity to engage in service activities. Similarly, the positive coefficient for the government membership variable, suggests that legislators who are members of the government engage in more service activities than opposition legislators.¹⁶ This may be due to government legislators having little incentive to question the policy decisions of their colleagues. In sum, these findings indicate that government membership and party affiliation are also relatively important predictors of legislators' constituency service.

3.5 Discussion

Although dynastic politicians are prevalent in many democracies countries in the world, researches have paid little attention to how dynastic politicians represent their constituents. In this work I broaden our understanding of political dynasties by examining the relationship between membership in a political dynasty and service

¹⁶This result is consistent with previous work on the constituency service activities of legislators in Ireland (Martin, 2011*a*). It should be remembered that this indicates that government legislators employ questions for service more than oppositions legislators. This doesn't indicate government legislators ask more questions in general. Opposition legislators submit far more questions than government affiliated representatives.

responsiveness. Building on existing work on political dynasties and political representation, I posited that the unique political socialization in which their previously elected relatives emphasize the importance of the cultivation of a personal vote and the social networks associated with dynasties lead to dynastic legislators providing more service than non-dynastic legislators. Using an original dataset consisting of the dynastic ties of legislators and all parliamentary questions submitted in Ireland between 1973 and 2007, I find that dynastic legislators in rural areas appear to engage in more constituency service than non-dynastic legislators. My results also suggest that dynastic legislators in highly urbanized regions provide similar service responsiveness to non-dynastic legislators in these regions. This evidence of variance in the representational styles of dynastic legislators raises the possibility that different mechanisms drive the formation and persistence of dynasties in different geographical regions.

The finding that a substantial proportion of dynastic legislators engage in significant constituency service can be used to both support and contradict the argument that dynasties negatively impact democratic politics. If one considers service responsiveness to be a key component of political representation, the provision of more constituency service clearly does not hinder democratic politics. However, if one considers service responsiveness as a trivial time consuming exercise which distracts from policy issues and slows national development, then some dynasties clearly impede democratic processes. In any case, it is clear that more research on the emergence and impacts of political dynasties is required. While existing research shows a politician's duration in office is positively related to the likelihood of a dynasty forming, we know little about how the behavior of politicians and other factors influence the probability of a subsequent formation of a dynasty. More attention needs to be paid to the members of dynasties that were the first in their family to obtain political office.

Existing research also finds that dynasties improve the descriptive representation

of females. However, it is perhaps more important to focus on the link between dynastic female legislators and the substantive representation of females. Qualitative evidence indicates that widows and daughters of deceased politicians are often strongly encouraged by senior party members to contest the upcoming bye-election to maintain the seat for the party. There is also some speculation that in countries with gender quotas, female legislators elected through these quotas often act as proxies for their male relatives and simply tow the party line (Nanivadekar, 2006). Studying this issue further would shed further light on the influence of dynasties on the substantive representation of women and the impact of dynasties, more broadly.

Chapter 4

Ethnic Quotas and Substantive Representation of Ethnic Minorities

“We demand that there should be a law which enables there should be reservation for backward classes, SCs and STs in the private sector jobs too. We see that even after so many years of independence the condition of Dalits in the country is still very bad.” - Ram Vilas Paswan (2016).¹

4.1 Introduction

In 1949, the Constituent Assembly of India, the first parliament of newly independent India, introduced the Constitution of India. The constitution included a provision that a proportion of the seats in the national parliament would be reserved for members of the Scheduled Castes (Dalits) and Scheduled Tribes (Tribes) ethnic groups. It was initially decreed that this reservation system would be maintained for 20 years.

¹Quote from Chen (2016).

However, in the period since the constitution's introduction, a number of constitutional amendments to continue the reserved seat system have been passed. Today, 131 of the 543 members of the Lok Sabha, the lower house of the bicameral Indian Parliament, are elected through reserved seats for members of Scheduled Castes and Scheduled Tribes and there is little to suggest that the system will be discontinued in the near future. India is, of course, just one country with electoral quotas for underrepresented sections of society. In recent years, a large number of countries have introduced special institutional arrangements to enhance the political representation of females and/or minority groups. More than 100 countries have introduced electoral quotas for women while over 25 countries have introduced formal electoral arrangements to promote the election of political representatives from ethnic minority groups (Htun, 2004; Krook and O'Brien, 2010; Bird, 2014).

The recent global proliferation of electoral quotas has prompted increased attention from political scientists on the effects of such provisions. Research indicates that gender quotas have increased the descriptive representation of women and have produced greater policy responsiveness to women's rights and interests (Xydias, 2007; Franceschet and Piscopo, 2008; Bhavnani, 2009; Hughes, 2011). Gender quotas also appear to increase the provision of public goods that are more relevant to women and reduce bias against female leaders (Chattopadhyay and Duflo, 2004; Duflo, 2005; Beaman et al., 2009). Turning to ethnic quotas, research primarily focuses on the influence of quotas on the allocation responsiveness of legislators. This work indicates that such institutional arrangements can promote the targeting of goods to minority groups (Chandra and Umaira, 2011), although partisan incentives can reduce the distributive effects of quotas (Dunning and Nilekani, 2013). Meanwhile, other research suggests that while ethnic quotas in India have helped to successfully integrate individuals from minorities into the political system and reduce discrimination against minorities in districts with reserved seats, quotas appear to have no effect on the welfare of minorities living in districts where quotas are used (Jensenius, 2013).

I contribute to this debate by examining the influence of ethnic quotas on the substantive representation of ethnic minority constituents by exploring whether there are differences in the responsiveness to the interests of minority groups between legislators elected via an ethnic quota and legislators not elected through a quota. I test the empirical association between ethnic quotas and policy responsiveness to minority ethnic groups using an original dataset, that I constructed, consisting of detailed legislator information and written questions submitted by legislators to ministers in India from 1999 and 2009. India is a good case to investigate the relationship between election via an ethnic quota and the nature of legislators' activities because it is one of the few consolidated democracies in the world to have longstanding quotas for minority ethnic groups. The results of my analyses indicate that ethnic quotas increase legislators' responsiveness to the interests of the relevant minority ethnic groups. This suggests that electoral quotas help to improve the substantive representation of marginalized groups. This result provides further evidence for the introduction of quotas as a means of reducing inequality in the political representation of historically disadvantaged groups in society. This finding also highlights the importance of considering both legislative processes and legislative outcomes when assessing the impact of quotas on the substantive representation of marginalized minority groups. The policymaking process is extremely complicated and solely focusing on outcomes can therefore neglect the influence of quotas on the efforts of legislators in the legislative process.

4.2 Electoral Quotas and Substantive Representation

Although there is a consensus in the electoral quotas literature that quotas have a positive impact on the descriptive representation of marginalized sections of society,

there is a still debate regarding the effect of quotas on the substantive representation of these groups. On one hand, researchers argue that the introduction of electoral quotas prompts an increase in both the descriptive and substantive representation of marginalized groups (e.g. Xydias, 2007; Franceschet and Piscopo, 2008). On the other hand, others contend that the impact of other institutional factors and the office seeking incentives of legislators nullify any potential effects of quotas on substantive representation (e.g. Dunning and Nilekani, 2013; Jensenius, 2013).

In the gender and race politics literature, it has been argued members of marginalized minority groups can better represent these groups than non-members because they have similar experiences to other members of these groups and have different preferences to individuals that do not belong to the group (Phillips, 1995; Mansbridge, 1999; Besley, Pande and Rao, 2005). Building on this, it has been posited that quotas reinforce tendencies to identify members of the marginalized groups as more similar to other members in these groups and less similar to individuals from other groups (Franceschet and Piscopo, 2008). Ethnic quotas could therefore potentially magnify group differences and contribute to the perception that legislators from minorities are required precisely for their perspectives related to these groups. Legislators elected through quotas will also be cognizant that these institutional arrangements were introduced with the goal of promoting the substantive representation of the respective marginalized groups and, consequently, should be more responsive to the interests of constituents from these groups than legislators not elected through a quota.

Recent research on ethnic quotas has countered this logic by contending that legislators' partisan and office seeking incentives can negate the effect of quotas on the substantive representation of marginalized ethnic groups (Dunning and Nilekani, 2013; Jensenius, 2013). This research argues that legislators are very reliant on support provided by party leaders to secure the party nomination and obtain resources such as campaign finance in their attempt to retain office. In order to secure and maintain this support, legislators' allocation responsiveness will therefore be more

reflective of the goals of party leaders than the interests of marginalized groups. That is, partisan ties can undermine the incentives produced by quotas to target allocations to underrepresented historically disadvantaged groups.

In a similar vein, it is also argued that office seeking incentives of legislators can impede the positive influence of quotas on the substantive representation of marginalized ethnic groups (Jensenius, 2015). In certain contexts, it is the case that legislators' constituents are both members of marginalized minority groups and individuals that do not belong to such a group. Building on the representation literature which assumes that politicians are primarily motivated by re-election (Mayhew, 1974; Fenno, 1978; Jensenius, 2013), it is argued that legislators therefore will need to adequately represent both groups of constituents to remain in office. Consequently, legislators who are elected through quotas that disproportionately focus on the interests of a minority group may not be able to secure the nomination of their party and, even if they do, will be unlikely to secure a plurality of votes from constituents at election time. On this basis, Jensenius (2015) argues that ethnic quotas in India have not produced outcomes that have improved the socio-economic well-being of historically disadvantaged Scheduled Castes.

While it is very plausible that promoting the substantive representation of ethnic minorities will not be the primary preference of representatives elected through ethnic quotas, when assessing the influence of quotas we are interested in relative differences in the substantive representation provided by legislators elected through quotas and legislators not elected through quotas. Moreover, research arguing that ethnic quotas do not improve substantive representation primarily focuses on what Franceschet and Piscopo (2008) refer to as 'outcomes', which are more closely related to the concept of allocation responsiveness. It is also important to examine representation 'processes', which are more closely related to the concepts of policy and service responsiveness. Although there may be no observable increase in the distribution of funds or socio-economic outcomes due to quotas, it is important to remember

that substantive representation is a multifaceted concept, and hence, this is not sufficient evidence to conclude that quotas do not influence substantive representation. To provide new insight into the influence of ethnic quotas on legislative processes, I therefore test the empirical association between ethnic quotas and legislators' policy responsiveness to minority ethnic groups in this paper.

4.3 Research Design

In this section I discuss why I focus on the responsiveness of legislators to the interests of ethnic minorities in India, describe my data, specify my empirical model, and discuss my results.

4.3.1 India

India is a good country to investigate the influence of ethnic quotas on legislators' responsiveness to the interests of minority ethnic groups. Since India's first general election in 1952, a substantial share of seats in the Lok Sabha have been reserved for individuals that are members of Scheduled Castes and Scheduled Tribes. Currently, 131, or 24 per cent, of legislators that serve in the Lok Sabha are elected through reserved seats and provide ethnic minorities with substantial descriptive representation in the legislature. Electoral quotas are assigned to districts primarily on the basis of the relative population of ethnic minority group members in a district, with districts possessing the highest relative ethnic minority populations generally being assigned reserved seats.² Similar to a number of other countries that have introduced formal institutional arrangements to improve the political representation of females and/or minority groups, India employs a single member district plurality (SMD) elec-

²For more detailed information on the assignment process of ethnic quotas in India, see the section dedicated to this chapter in the Appendix.

toral system, where the candidate that wins the most votes in a district is elected.³ By focusing on India we can thereby obtain new insight into the influence of ethnic quotas on the policy responsiveness of legislators to ethnic groups' preferences in parliamentary democracies employing SMD electoral rules.

4.3.2 Data

I have constructed an original dataset comprising detailed information on legislators that served in the Lok Sabha and all questions submitted by legislators to ministers between 1999 and 2009. There are two Lok Sabhas in this period. The maximum term length of a Lok Sabha is five years.

As in many other countries, parliamentary questions are an important feature of the legislative process in the Lok Sabha. During the first hour of every sitting of the house, legislators in the Lok Sabha typically have the opportunity to question the prime minister and government ministers (Parliament of India, 2016).⁴ Legislators that submit questions are required to give between ten days and twenty days notice, although a limited number of questions that are deemed to be of significant public importance by a minister will be answered with shorter notice (Parliament of India, 2016). Legislators in the Lok Sabha submit both oral and written questions. Oral questions receive an oral answer by a minister during a legislative session, while written questions receive a written reply from a minister.⁵ As in other countries, ministers answer questions on a rota basis, enabling legislators to submit queries related to a wide variety of issues and policies.

³It is generally the case that single member district plurality electoral rules produce two party systems and single-party governments. However, the number of parties in the Lok Sabha has generally exceeded two (Chhibber and Kollman, 1998).

⁴This is at the speaker's discretion. The speaker has the authority to cancel question hour on a given day.

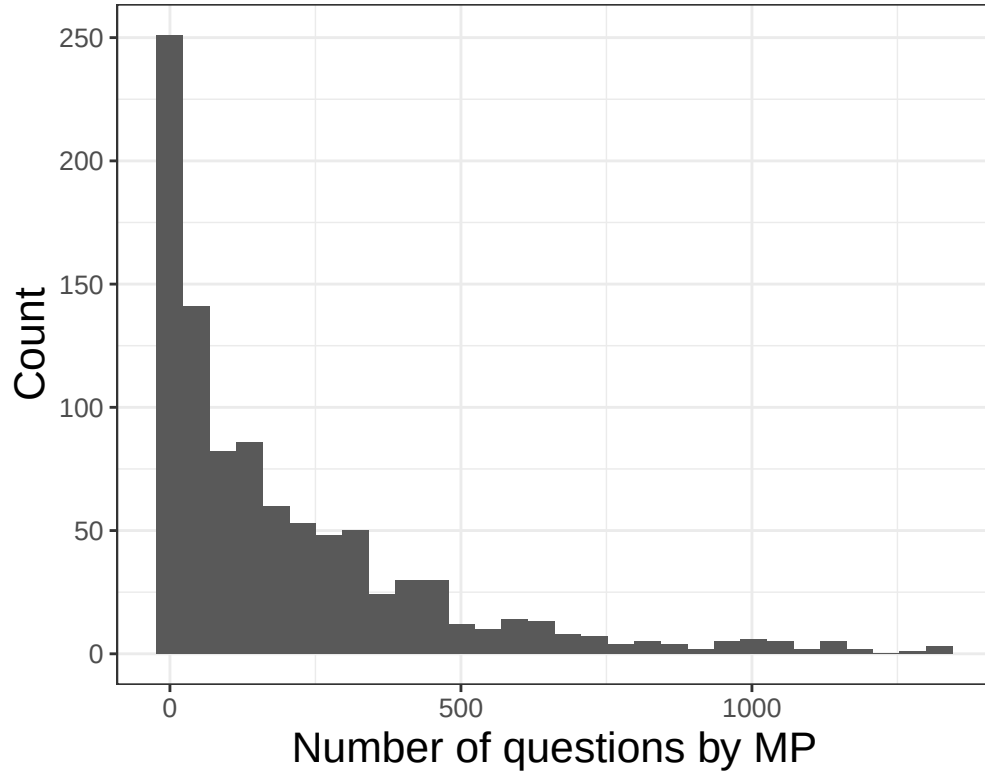
⁵Oral questions are referred to as starred questions in the Lok Sabha, while written questions are referred to as unstarred questions.

Table 4.1: Total number of questions asked by legislators in India between 1999 and 2009

Lok Sabha	Number of legislators that asked a question	Total questions asked
13	480	100,024
14	483	96,652

While there are some limitations on the content of queries submitted by legislators in the Lok Sabha, questions are a particularly rich source of information for measuring legislative responsiveness given the limited availability and complexity of speech data from the Lok Sabha and the lack of available data on legislative voting (Yadav, 2015). As outlined in Table 4.1, 480, or 84 per cent, of the legislators that served in the thirteenth Lok Sabha asked a question, with a total of 100,024 questions submitted by legislators in this period. Similarly, 483, or 82 per cent, of the legislators that served in the fourteenth Lok Sabha submitted a question to the government, with 96,652 questions being asked in this period. This demonstrates that questions provide information on the activities of a large number of legislators that serve in the Lok Sabha. In Figure 4.1, I present the distribution of the total number of questions asked by legislators in these Lok Sabhas. From this plot, we can see that there is considerable variation in how many questions legislators choose to submit. The mean number of questions submitted by a legislator in a Lok Sabha is 204, with the median legislator asking 119 questions.

Figure 4.1: Distribution of number of questions by MPs between 1999 and 2009

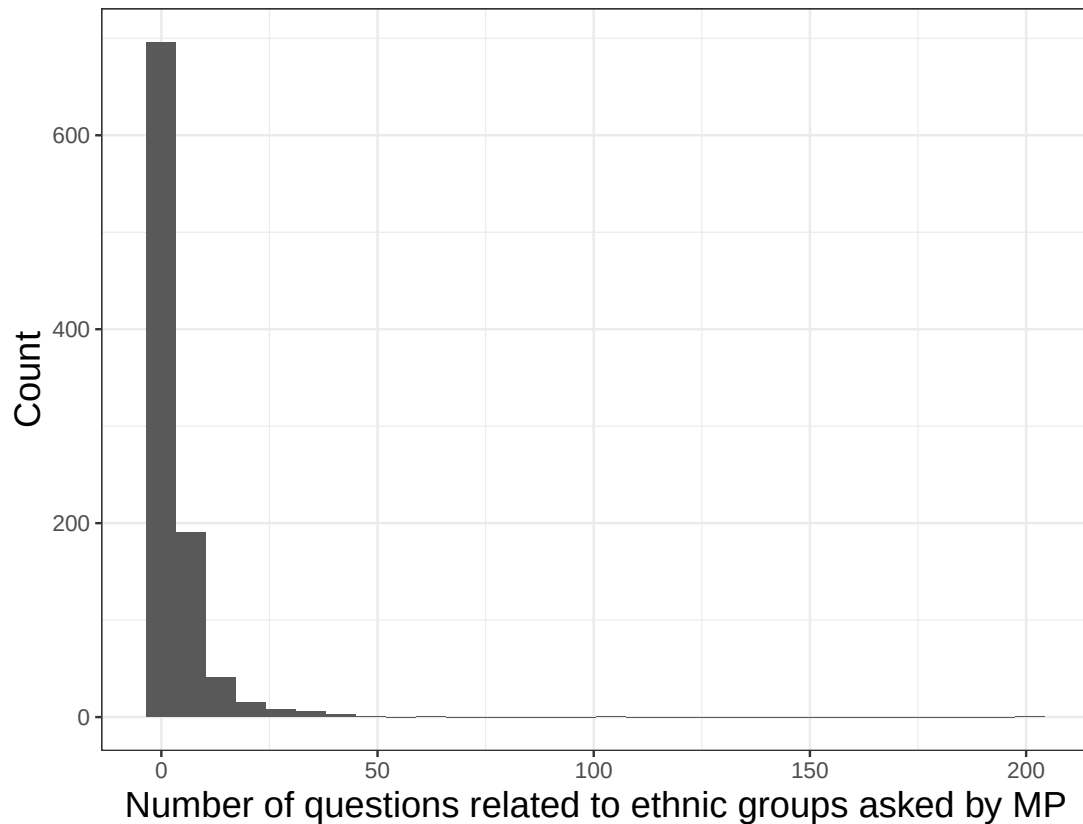


To operationalize legislators responsiveness to the interests of ethnic groups, I employ the counts of questions pertaining to ethnic groups' interests. I classify a question as related to the interests of ethnic minorities if the question contains at least one of the following terms: 'caste', 'tribe', 'dalit', 'ST', and 'SC'. Legislators typically use these terms when discussing the marginalized groups that are the subject of the ethnic quota in India. This rule for classifying documents is akin to the dictionary approach discussed by Grimmer and Stewart (2013). To do this, I first obtained the text of questions and related metadata by scraping webpages of questions from the website of the Lok Sabha.⁶ I then utilized regular expressions to determine whether

⁶More detailed information on this process is provided in the technical details section of the Appendix.

any of the relevant terms were contained in the text of a submitted question. As a final step, I read a random sample of questions that were classified as related to ethnic minority groups to ensure the validity of this strategy.⁷ In the thirteenth Lok Sabha, 2,030 questions related to ethnic groups were asked by MPs, with legislators asking 1,507 such questions in the fourteenth Lok Sabha.

Figure 4.2: Distribution of number of questions related to marginalized ethnic groups by MPs between 1999 and 2009



The primary predictor of interest in this study is whether a legislator was elected via an ethnic quota or not. Details on whether a legislator represents a district with an ethnic quota were obtained from detailed legislator biographies on the website

⁷I provide examples of these questions in the Chapter 4 section of the Appendix.

of the Lok Sabha. In Table 4.2 I present the number of legislators elected through ethnic quotas in both the thirteenth and fourteenth Lok Sabhas. We can see that 124, or 22 per cent, of the legislators that served in the thirteenth Lok Sabha obtained office through a reserved seat. Similarly, 125, or 21 per cent, of the legislators that served in the fourteenth Lok Sabha were elected via an ethnic quota. In testing the empirical association between ethnic quotas and legislators' policy responsiveness to ethnic minorities, I also account for a set of other factors that are likely predictors of legislators' representation of minorities' interests. As questions are a tool primarily employed by opposition legislators, I include an indicator variable denoting whether a legislator is in government or a member of the opposition. To account for potential differences in male and female legislators' attention to minority ethnic groups, I also include an indicator variable representing the gender of a legislator.

Table 4.2: Total number of legislators elected through ethnic quotas in India 1999-2009

Lok Sabha	Legislators elected via an ethnic quota
13	124
14	125

I measure all of the variables in this study at the legislator level for both the thirteenth and fourteenth Lok Sabhas. For the outcome variable, this means that I count the number of questions related to ethnic minority groups submitted by all legislators that asked a question in each Lok Sabha.

Model specification

To understand the influence of ethnic quotas on representation of the minority ethnic groups' interests, I estimate negative binomial models of the total number of questions related to the interests of Scheduled Castes and Scheduled Tribes submitted by

legislators in the Lok Sabha. In these models, I include an offset variable to account for variation in the total number of questions submitted by different legislators. For instance, consider the case where a legislator elected via an ethnic quota and a legislator not elected through a quota both submitted 13 questions related to the interests of Scheduled Castes and Scheduled Tribes in the period of a given Lok Sabha. These counts should only be considered similar if both legislators have similar exposure, i.e. the legislators submit a comparable total number of questions. As demonstrated earlier, there is considerable variation in the number of questions submitted, and I therefore include the log of the total number of questions submitted by each legislator in a Lok Sabha session as an offset variable. For some of the model specifications, I use a multilevel regression framework to estimate the relationship between election through a reserved seat and a legislator’s attention to the interests of ethnic minorities. In these specifications, to allow for additional heterogeneity and to help account for variation across Lok Sabhas, variation across geographical regions, and variation across parties, I model the intercept as a function of parliament, state, and party indicators, thereby creating parliament, state, and party random effects.

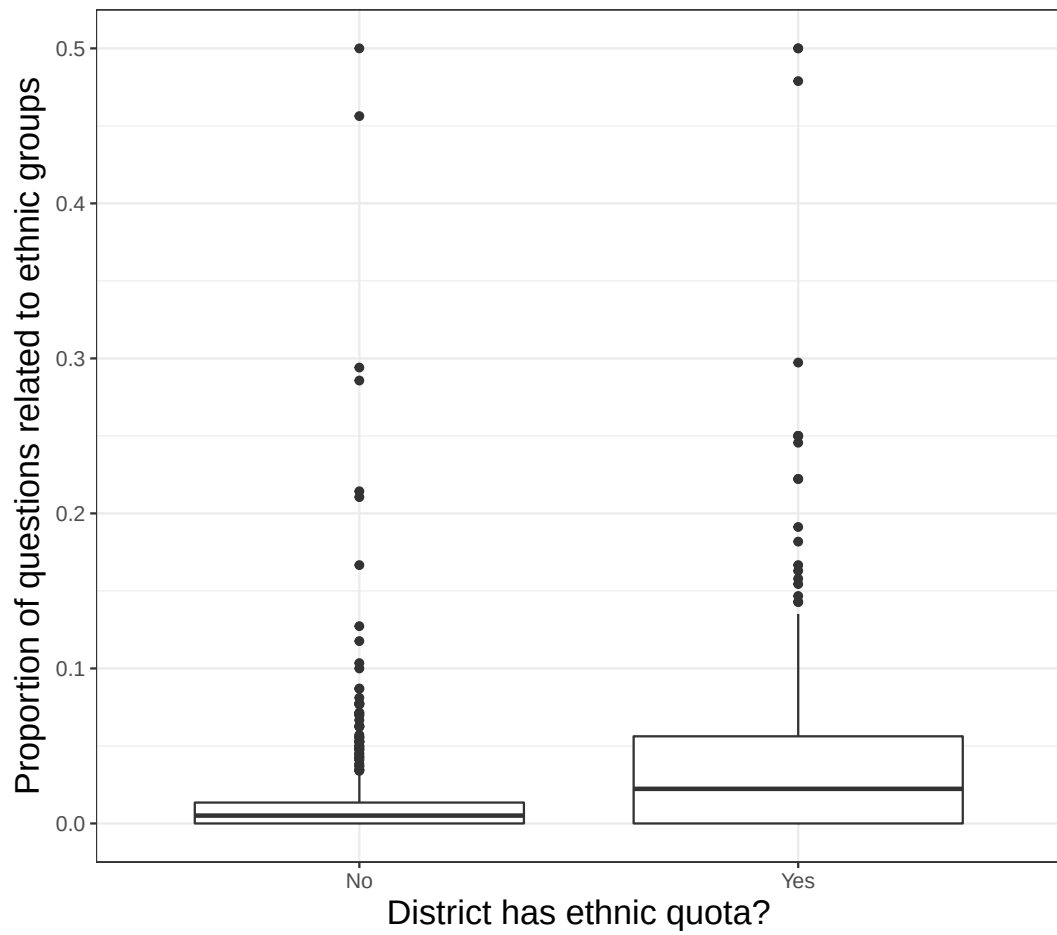
4.4 Results

Prior to discussing the results of the various models outlined above, I first examine the bivariate relationship between election via an ethnic quota and the proportion of legislators’ questions related to Scheduled Castes and Scheduled Tribes. In Figure 4.3, I present the proportion of legislators’ questions related to Scheduled Castes and Scheduled Tribes for both legislators who are elected through ethnic quotas and legislators not elected through ethnic quotas.⁸ From this plot, we can see that the proportion of submitted questions pertaining to ethnic groups is substantially higher

⁸For greater interpretability, I exclude observations where the proportion of submitted questions related to ethnic groups exceeds 0.5 from this plot.

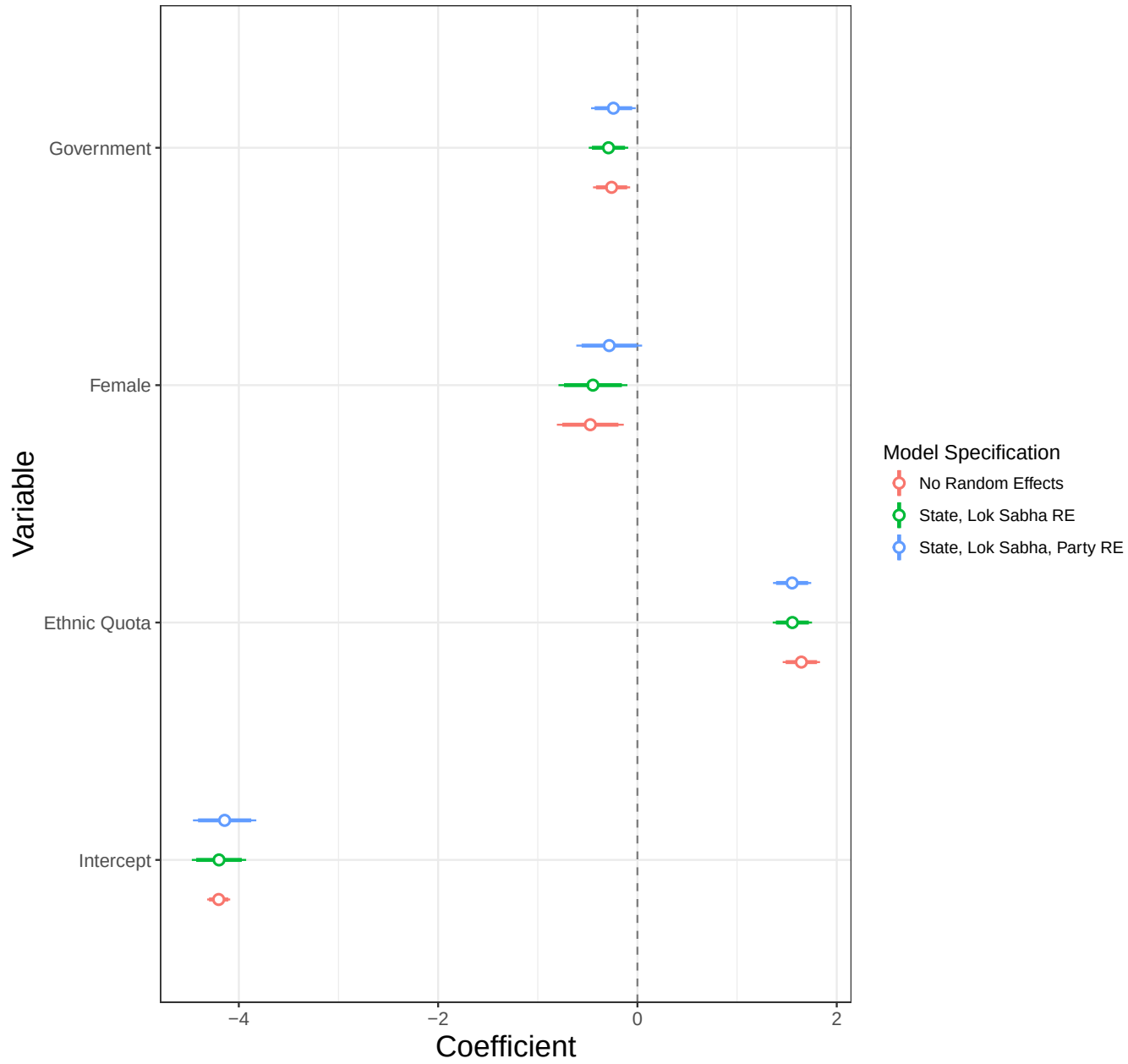
for legislators elected in districts with ethnic quotas.⁹

Figure 4.3: Bivariate relationship between ethnic quotas and questions related to ethnic groups



⁹A nonparametric Wilcoxon sum rank test indicates that this difference is statistically significant.

Figure 4.4: Ethnic quotas and policy responsiveness in India 1999-2009



I next present the results of the various estimated negative binomial count models in Figure 4.4.¹⁰ In line with the previous examination of the bivariate relationship

¹⁰A table of coefficients and other model statistics can be found in the section dedicated to this

between ethnic quotas and questions related to ethnic groups, we can see that the coefficient indicating a legislator was elected via an ethnic quota is positive and statistically significant in each model. This indicates that election via an ethnic quota is a strong and robust predictor of a legislator's responsiveness to the interests of ethnic minority groups. It is also clear that the magnitude of the coefficient is much larger than the magnitude of the coefficients for the government and gender indicator variables. To better understand the substantive influence of ethnic quotas on legislators' representation of the interests of ethnic minority groups, I calculate the expected counts of submitted questions pertaining to ethnic minority groups for every 100 questions asked for both legislators elected through an ethnic quota and those not elected through an ethnic quota, while setting the other variables included in the model at their modal values.

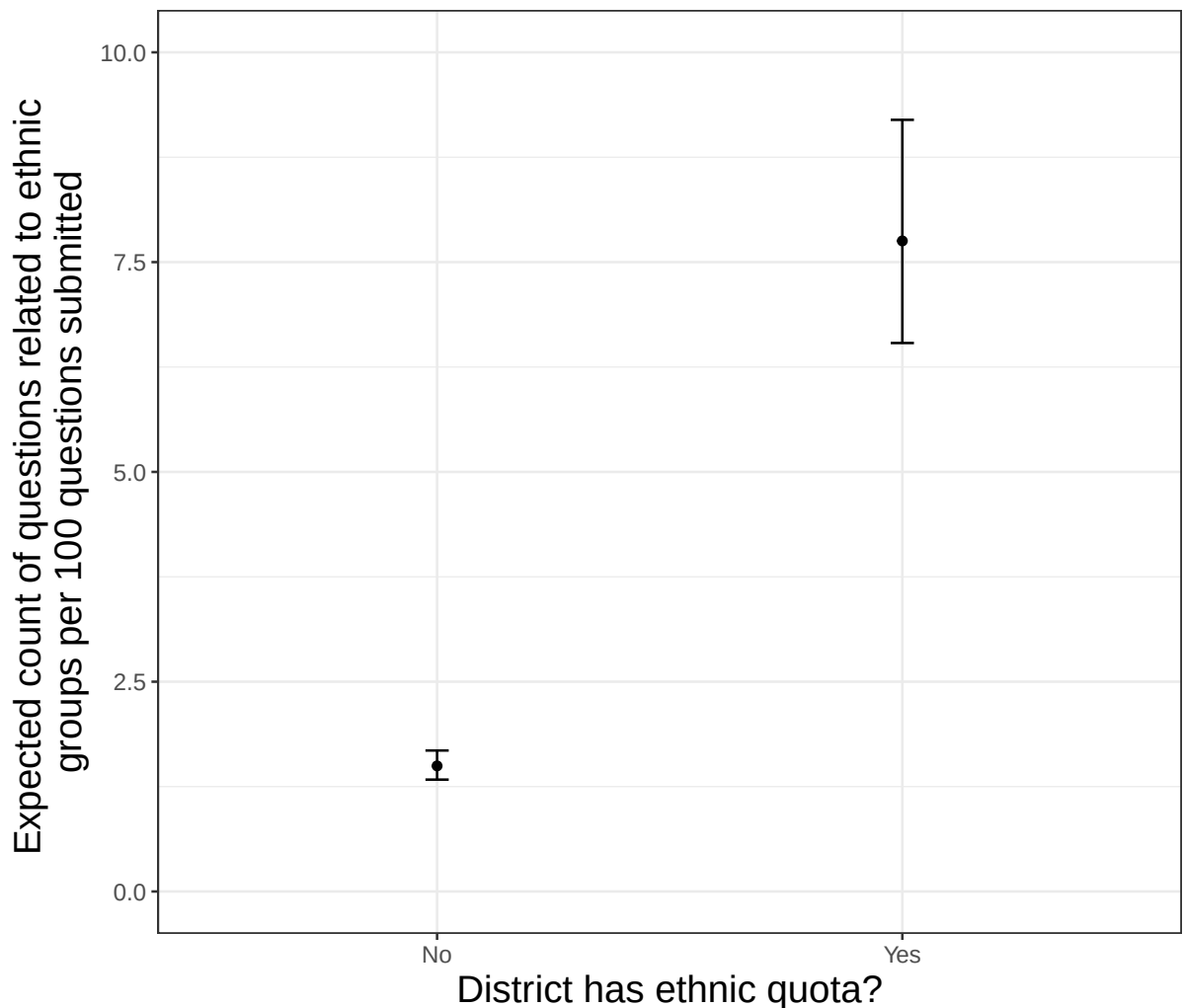
I present these expected counts in Figure 4.5.¹¹ From this plot, it is evident that the expected count of submitted questions pertaining to ethnic minority groups is considerably higher for legislators that were elected via an ethnic quota. More specifically, for every 100 questions submitted by a legislator elected in a district with an ethnic quota, we would expect that, on average, 8 of these questions would relate to Scheduled Castes and Scheduled Tribes. Contrastingly, for every 100 questions submitted by a legislator elected in a district that does not have an ethnic quota, we would expect that, on average, only 2 of these questions will pertain to Scheduled Castes and Scheduled Tribes. That is, we would expect for every 100 questions submitted, a legislator that obtained office through an ethnic quota will ask 6 more questions pertaining to Scheduled Castes and Scheduled Tribes than a legislator elected in a district without an ethnic quota, all else equal. As the confidence intervals of these expected counts do not overlap, we can be confident that this difference is substantively

chapter in the Appendix.

¹¹The expected counts were generated using the estimated parameters from the negative binomial model specification that did include any random effects.

meaningful. These results provide strong evidence of a robust relationship between ethnic quotas and legislators' responsiveness to the preferences and interests of ethnic minority groups in India. These results also lend further supporting evidence to the literature positing that quotas for underrepresented sections of society enhance the substantive representation of these groups.

Figure 4.5: Expected counts of questions related to ethnic minority groups



Turning to the other predictors included in the estimated models, we can see

that the magnitude of the coefficients for both the gender and government indicators are both quite small in magnitude relative to the ethnic quota variable. For the model specifications that do include party random effects, the coefficient for the gender indicator is statistically significant. When party random effects are included in the model, the gender indicator is neither a substantively or statistically important predictor of legislators' responsiveness to ethnic groups. Likewise, although the estimated parameters for the government indicator can be considered different from zero at conventional levels of statistical significance for all specifications, the magnitudes of the coefficients for this variable are very small relative to the coefficients for the ethnic quota indicator. In sum, these results suggest that neither of these variables is a particularly important predictor of legislators' attention to the interests of ethnic minorities in India.

4.5 Discussion

In recent years a large number of countries have introduced electoral quotas in an effort to enhance the political representation of marginalized and underrepresented demographic groups. This has produced a sizable political science literature investigating the effects of these institutions. While the majority of this literature argues that electoral quotas enhance the quality of representation of marginalized groups, recent research has presented evidence that office-seeking and partisan incentives can mitigate and nullify the positive impact of quotas. Using an original dataset consisting of all parliamentary questions submitted by legislators in the Lok Sabha between 1999 and 2009, I contribute to this debate by examining the empirical association between election in a district with an ethnic quota and legislators' responsiveness to the interests of marginalized ethnic groups. Using a number of different model specifications, I find that there is a strong, positive, and robust empirical relationship between longstanding electoral quotas for members of Scheduled Castes and Sched-

uled Tribes and legislators' policy responsiveness to the interests of these historically disadvantaged groups. This finding provides further support to research which argues that quotas positively influence the policy representation of marginalized sections of society and indicates that the introduction of quotas may help alleviate inequity in the political representation of minority demographic groups.

While this study provides supporting evidence for a positive relationship between ethnic quotas and responsiveness to marginalized ethnic groups, the introduction and utilization of electoral quotas nevertheless pose significant challenges to legislators and policymakers. Firstly, there are multiple different forms of electoral quotas. This study focused on the implementation of an electoral quota in the form of a reserved seat system in which only members of marginalized ethnic groups can contest the election. Other forms of electoral quotas include national electoral quotas regulating the proportion of candidates from different demographic groups that must be candidates in an election, political party quotas where the party mandates that a certain proportion of its candidates must be from specified minority groups, and reserved seats for parties representing these groups. Secondly, in addition to gender and ethnicity, electoral quotas have been introduced in countries to reduce under-representation on the basis of religion, age, nationality, language, disability and profession (Hughes, 2011). Although most countries have electoral quotas based on one of these categories, a small proportion have introduced quotas that try to enhance representation for intersectional combinations of these groups (Krook and O'Brien, 2010). A final challenge is that there is mixed evidence regarding whether introducing quotas for one marginalized section of society helps improve representation for other minority groups (Folke, Freidenvall and Rickne, 2015). Understanding how a quota targeted at one minority group influences the representation of other marginalized sections of society is critical for understanding the broad representational impacts of quotas. These challenges give rise to the difficult questions of how many marginalized minority groups in a country should be guaranteed representation through quotas and

how such a system can be implemented effectively. Both policymakers and political scientists will need to carefully consider how these different complexities can affect the future quality of different components of political representation.

Chapter 5

Politweeting: Exploring Legislators' Use of Twitter

“Welcome to my Twitter page, if I can be of any assistance to you please leave me a message or send me an email at robert.troy@oireachtas.ie” - First tweet by Deputy Robert Troy (2012).

5.1 Introduction

Since its foundation in 2006, the social media platform Twitter has quickly become an important forum for political discussion. Large numbers of Twitter users discuss political events on Twitter everyday. Representatives, other political actors, and citizens all employ Twitter for communication. Currently, Twitter is perhaps the primary communication platform utilized by the incumbent President of the United States, Donald Trump. Moreover, as noted by Barberá (2015), a large proportion of legislators in advanced democracies have Twitter accounts. Given this, it is unsurprising that political scientists have started to identify Twitter as a relevant medium for understanding politics and a useful source of data for examining a variety of different

political questions.

The existing literature examining legislators' usage of Twitter indicates that legislators discuss a number of different issues in tweets and that legislators employ Twitter for impression management and self promotion, and, to a lesser extent, constituency service (Golbeck, Grimes and Rogers, 2010; Lassen and Brown, 2011; Barberá et al., 2014). Barberá (2015) also shows that Twitter network data can be used to accurately replicate conventional measures of legislators' ideologies and measure the ideology of other users, enabling researchers to learn about the political preferences of a very large number of individuals. In a similar vein, evidence indicates that partisan differences in the language of Congress members' tweets closely resemble partisanship patterns observed when employing other data sources (Radford and Sinclair, 2016). Recent work also suggests that legislators strategically utilize hashtags when discussing particularly divisive issues (Hemphill, Culotta and Heston, 2013).¹ Twitter has also been used to study the campaign activities of candidates in the months leading up to an election, with existing evidence suggesting that the platform is primarily employed for communicating candidates' plans for government (Evans, Cordova and Sipole, 2014).

In this chapter I contribute to the growing Twitter literature by exploring legislators' utilization of Twitter using a new dataset consisting of all tweets by Irish legislators that served in the 31st Dáil Éireann between 2011 and 2016. Using this data, I examine the relationship between different political factors and both legislators' use of Twitter and the content of legislators' tweets. In these analyses, I find that the age of a legislator, whether a legislator holds a position in the cabinet, and whether a legislator serves in a highly urbanized area are the strongest predictors of Twitter usage. I also show that legislators discuss both political and non-political issues in tweets and that a number of the topics discussed are closely related to important components of political representation. I further demonstrate that the empirical

¹Hashtags are terms or phrases preceded with the # sign that are used to indicate the tweet relates to a particular topic.

relationship between different political factors and these topics are in line with prior theoretical expectations. This indicates that valid measures of political representation can be generated from the text of tweets and provides evidence that Twitter is a useful resource for gleaning new insights and testing theories related to legislators' activities and representation of constituents.

5.2 Who tweets?

I first explore the relationship between different political factors and legislators' characteristics and whether or not a legislator has a Twitter account. In Table 5.1, I present the total number of TDs that served in the 31st Dáil Éireann who have a Twitter account. From this table we see can that 153 out of 173, or 88 per cent, of TDs who were elected to the 31st Dáil have a Twitter account.

Table 5.1: TDs who served in the 31st Dáil Éireann Twitter status

Account	No Account
153	20

To understand whether there are systematic differences between legislators with a Twitter account and legislators that do not have a Twitter account, I model whether a legislator is a Twitter user as a function of a number of potentially relevant demographic, political, and constituency covariates. To capture the possibility that the electoral performance of legislators influences the decision to use Twitter, I employ the percentage of first preference votes obtained by a TD in their election to Dáil Éireann as a measure of their electoral strength. To capture the possibility that high ranking politicians are more likely to use Twitter, I include an indicator variable capturing whether a legislator serves in the government cabinet. Similarly, to capture whether there is a systematic difference in the propensity of government and oppo-

sition legislators to adopt Twitter, I include a binary variable capturing whether a legislator is a member of a governing party or not. To capture potential differences in Twitter usage across demographic groups, I also include variables measuring the age of legislators in 2011 and a binary indicator representing the gender of a legislator. Finally, to account for possibility that social media usage patterns differs across more urban and less urban areas, I include an indicator capturing whether a legislator serves in a Dublin constituency.² Data for these predictors was obtained from the Elections Ireland website and Wikipedia.

I model the relationship between having a Twitter account and the various factors discussed above using logistic regression and LASSO logistic regression models. The LASSO logistic regression model is a modified version of logistic regression which penalizes coefficients to improve the predictive performance of the model (Hastie, Tibshirani and Friedman, 2009). The LASSO penalty can result in the coefficients of some variables having a value of exactly zero and is therefore useful for identifying the most important variables in a model (Fariss and Jones, 2017).³

In Figure 5.1, I plot the coefficients from the models discussed above. From the plot it is immediately clear that the majority of the factors are not strong predictors of legislators' adoption of Twitter. The only variable that is statistically different from zero in the logit model is the age of a legislator. The coefficient of age is negative indicating that legislators are less likely to adopt Twitter as they get older. Similarly, age is the only variable that has a non-zero coefficient in the LASSO specification. Together, this indicates that age appears to be the most useful predictor of whether a legislator has a Twitter account. To better understand legislators usage of Twitter, I therefore focus on variation in how much legislators with Twitter accounts tweet.

²As noted earlier in this work, the population density of the Dublin region is far higher than that of any other city or large town in Ireland.

³The LASSO model was also used in Chapter 3 for classifying questions as constituency service and more details on this model can be found in the technical details section of the Appendix.

Figure 5.1: The adoption of Twitter by TDs in the 31st Dáil Éireann

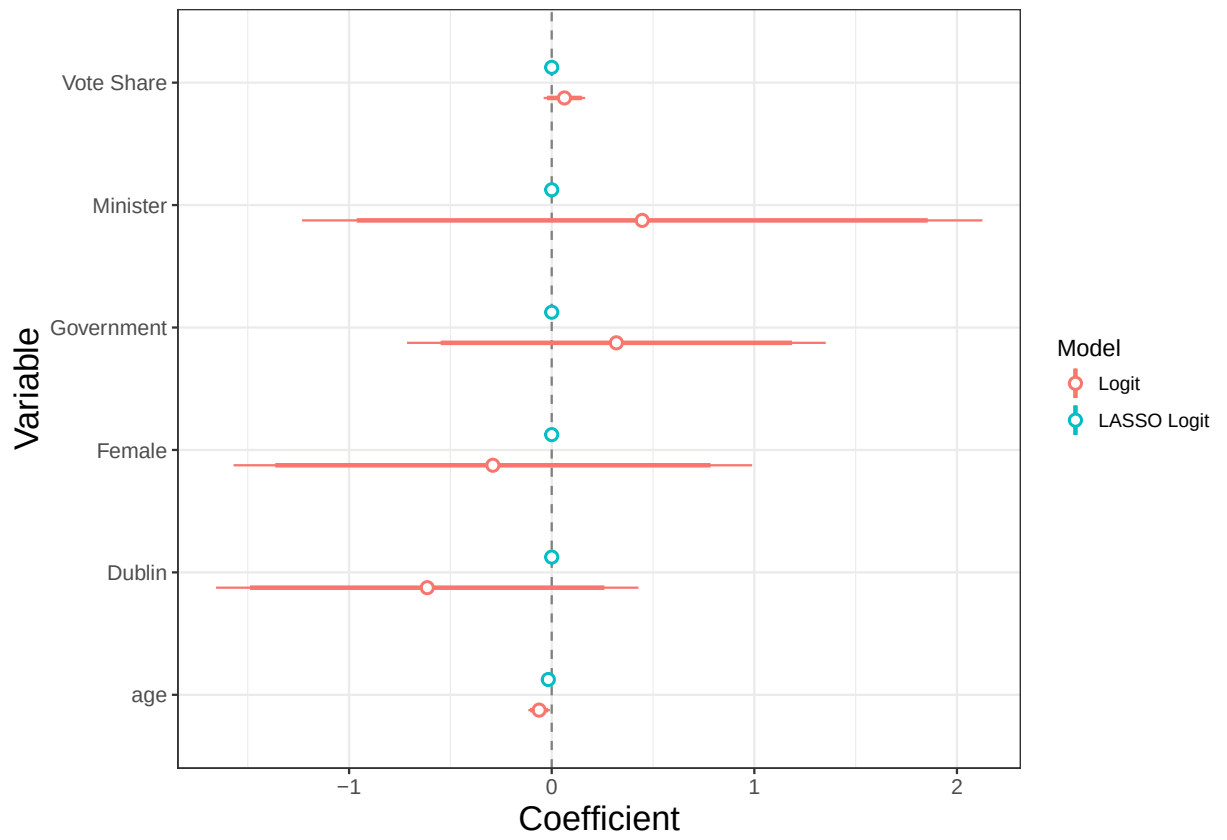
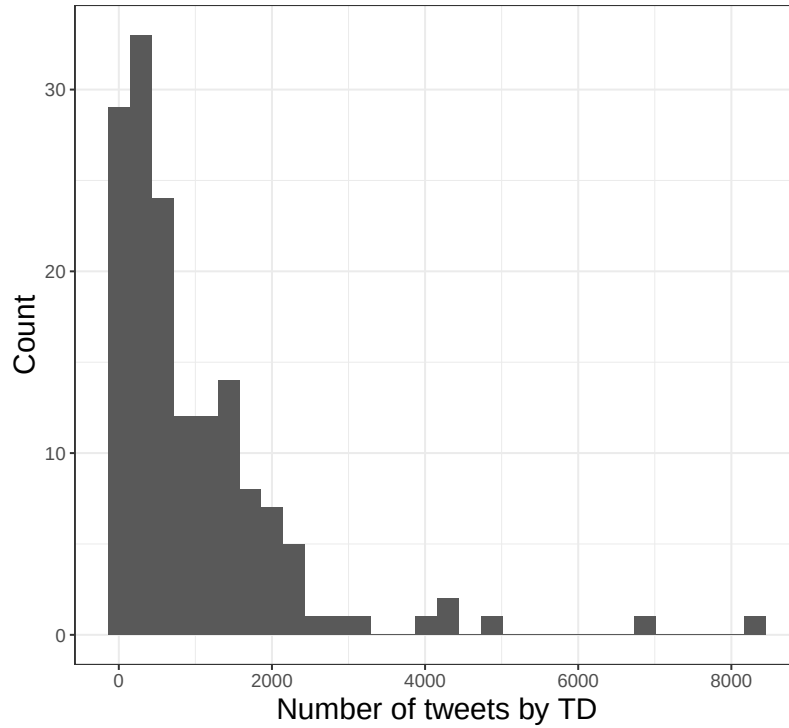


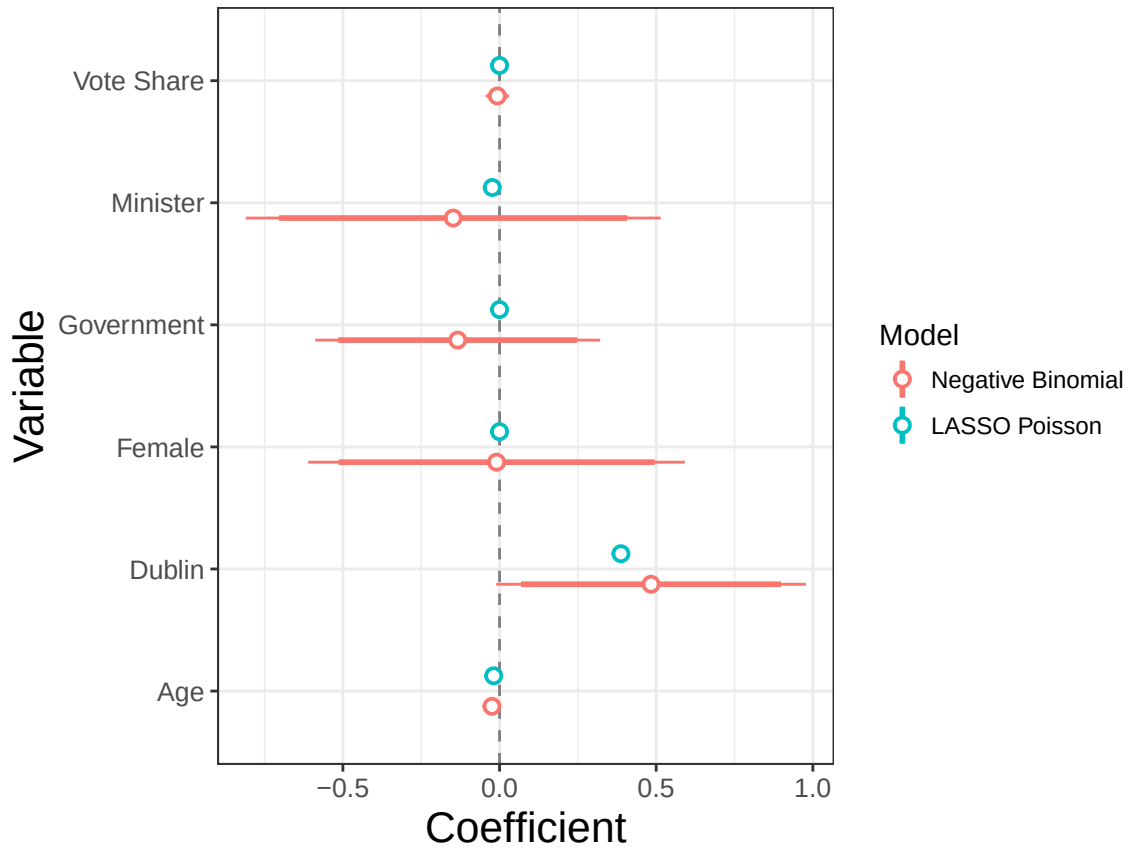
Figure 5.2: Variation in Twitter usage by TDs in the 31st Dáil Éireann



In Figure 5.2, I present the distribution of tweets posted by legislators during the 31st Dáil Éireann. We can see that this distribution is very skewed with a long right tail, indicating a small number of legislators tweet far more than the typical legislator. To better understand this variation, I model the number of tweets posted by a legislator with a Twitter account as a function of the same variables included in the logistic regression models. I estimate the relationship between the number of tweets posted and these factors using negative binomial and LASSO Poisson models.⁴

⁴The data displays overdispersion which causes the standard errors of coefficients from a Poisson regression model to be incorrectly estimated. Overdispersion therefore can also lead to incorrect inferences. As the LASSO Poisson model does not produce standard errors, the problems associated with overdispersion are not particularly relevant for this model. The LASSO Poisson model works in a similar fashion to LASSO logit model, shrinking variables to zero to improve the predictive accuracy of a model.

Figure 5.3: The usage of Twitter by TDs in the 31st Dáil Éireann



In Figure 5.3, I plot the coefficients from the negative binomial and LASSO Poisson regression models discussed above. As in the logit model presented earlier, the only variable that is statistically different from zero in the negative binomial model is the age of a legislator. The coefficient of age is negative suggesting younger legislators post more tweets than older legislators. Similarly, the age variable has a non-zero coefficient in the LASSO model for assessing variable importance. This again indicates that age is a useful predictor of Twitter use by legislators. It is also noticeable that the coefficient for the indicator capturing whether a legislator serves in a Dublin constituency is much larger than the coefficients for all other variables. This suggests

that legislators in highly urbanized areas tweet more frequently than legislators in less urbanized regions. The only other factor with a non-zero coefficient in the LASSO model is the indicator capturing whether a legislator served as a minister in the government cabinet. This coefficient is negative, suggesting that ministers tweet less than those who do not serve in the cabinet. In sum, these results suggest that the age of a legislator, whether a legislator holds a position in the cabinet, and the geographic region a legislator serves in are all useful predictors of legislators' Twitter usage. However, while the other factors in the model do not seem to influence the volume of tweets posted by legislators, this does not mean that they do not affect what legislators tweet about in a substantively meaningful way.

5.3 Understanding the content of legislators' tweets

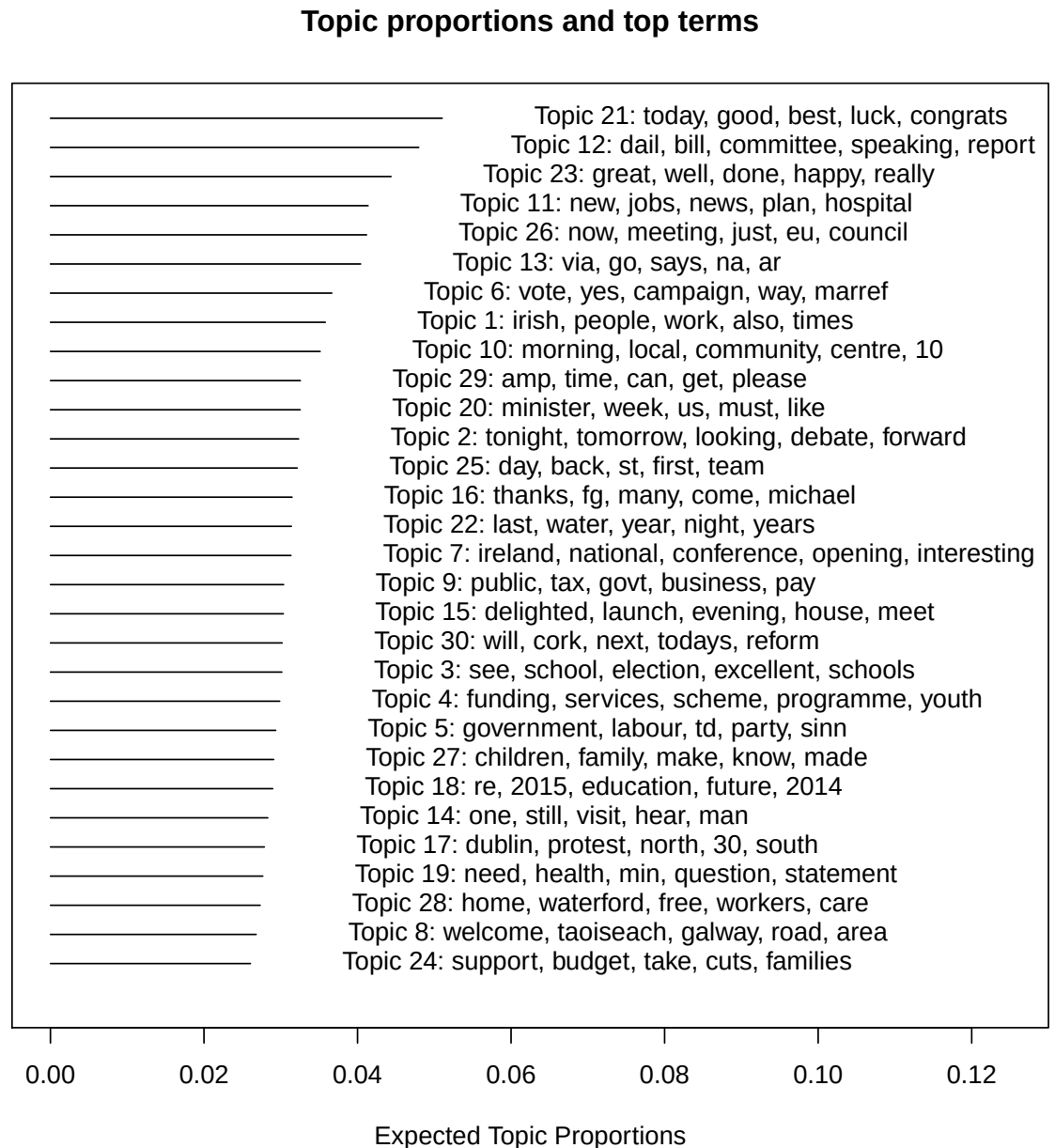
I next examine the content of legislators' tweets and how different political factors are related to what legislators tweet about. To understand the content of legislators tweets, I use a topic modelling approach. Topic models are statistical algorithms which provide information about the underlying thematic structure of a collection of documents (Blei, 2012). The primary quantity of interest produced by a topic model is the group of topics discussed in the collection of documents. Topic models estimate these topics by using information about the frequency of word co-occurrence in different documents (Grimmer and Stewart, 2013). In the context of tweets by legislators, topics mostly represent politically relevant concepts. For instance one topic could relate to government economic policy while another topic could pertain to the issue of education. Topic models therefore help us to understand what issues legislators focus on in tweets and can provide further information on the priorities of legislators.

In recent years, topic models have become an extremely popular analytical tool in political science research and political scientists have developed a number of different

topic models for analyzing politically relevant text (Quinn et al., 2010; Grimmer, 2010; Roberts et al., 2014). I use the correlated topic model which is the most basic implementation of the structural topic model introduced by Roberts et al. (2014). This builds on the widely used Latent Dirichlet Allocation model by allowing topics to exhibit correlation (Blei, Ng and Jordan, 2003; Blei and Lafferty, 2007).⁵ The structural topic model has been previously utilized to study a wide variety of text corpora including tweets related to Edward Snowden, news reports pertaining to China, and open ended survey responses from the American National Election Study (Roberts et al., 2014; Lucas et al., 2015; Roberts, Stewart and Airolidi, 2016). As with other topic models, this model assumes that each document/tweet is a mixture of different topics and the researcher must specify the number of topics to be estimated, K . I estimated a number of different models with different numbers of topics and, following the guidelines of Quinn et al. (2010), chose $K = 30$ which produced both relatively interpretable and meaningful topics.

⁵A detailed comparison of different topic modelling approaches can be found in the Supplementary Appendix of Roberts et al. (2014).

Figure 5.4: Topic proportions and top terms associated with each topic from correlated topic model



In Figure 5.4, I present the expected proportion of the documents in the corpus that are assigned to a topic for all 30 estimated topics. This plot also presents the top five words associated with each topic. These words can help label topics, since topic models do not inform the user what a topic measures.⁶ From the plot we can see that legislators tweet about a variety of different matters. It is apparent that legislators do not exclusively tweet about politically relevant issues. For instance, the topic with the largest proportion of documents assigned to it is a topic pertaining to well wishes messages posted by legislators. In these messages, legislators wish various individuals and teams *good luck* or the *best of luck* in their exams or match. Other non-political topics include a topic pertaining to St Patrick’s Day festivities and a topic representing messages of congratulations sent by legislators. We can also see that legislators post about their legislative activities such as committee meetings, constituency clinics, and launch events, different policy domains such as economic policy and health policy, various referenda, water charges and anti-government demonstrations protesting these charges, and allocations to constituencies.⁷ This indicates that legislators use Twitter for advertising, position taking, and credit claiming, which are, according to Mayhew (1974), the three primary activities of legislators. This provides evidence that legislators’ tweets may be a useful resource for testing theories related to legislative behaviour and legislators’ representation of constituents.

In the following sections, I provide a series of validations to assess the effectiveness of using different estimated topics as measures of legislators’ policy responsiveness, allocation responsiveness, service responsiveness, and advertising activity. I pay particular attention to construct validity, which refers to how well a measure captures the theoretical construct of interest, and its subcategory concurrent validity, which refers to how well a measure distinguishes between groups we theoretically expect it

⁶To label the topics I therefore also read a number of documents that are most strongly associated with a given topic. I provide a table of topic labels in the Chapter 5 section of the Appendix.

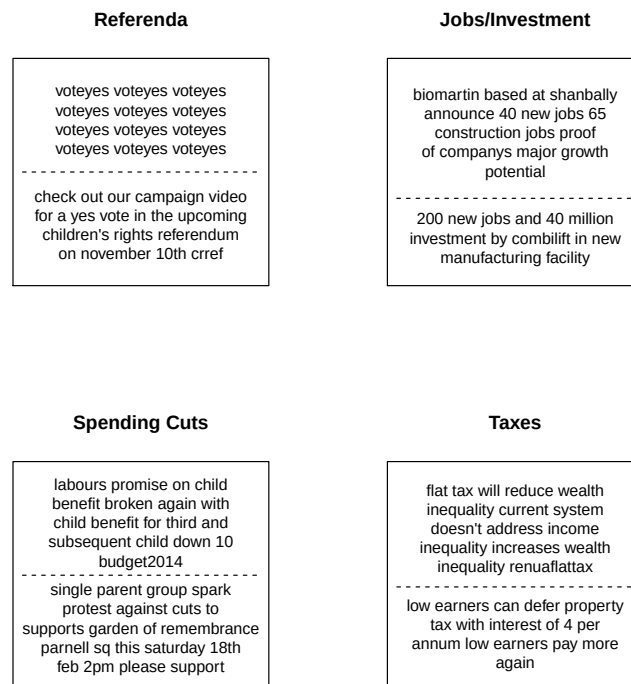
⁷Clinics are meetings which occur in the constituency of a legislator in which constituents can meet with a legislator and can request assistance with an issue etc.

to distinguish between (Trochim and Donnelly, 2001).

5.3.1 Policy responsiveness on Twitter

I first examine the relationship between affiliation with a government party and the attention that legislators pay to different policy areas on Twitter to validate that the observed relationships meet theoretical expectations. For this exercise, I focus on the economic policy topics related to job creation and investment, taxes, and government spending cuts. Additionally, I also examine the topic pertaining to government proposed referenda. To provide more insight into what these topics capture, I present sample representative tweets from these topics in Figure 5.5.

Figure 5.5: Example tweets from different policy topics



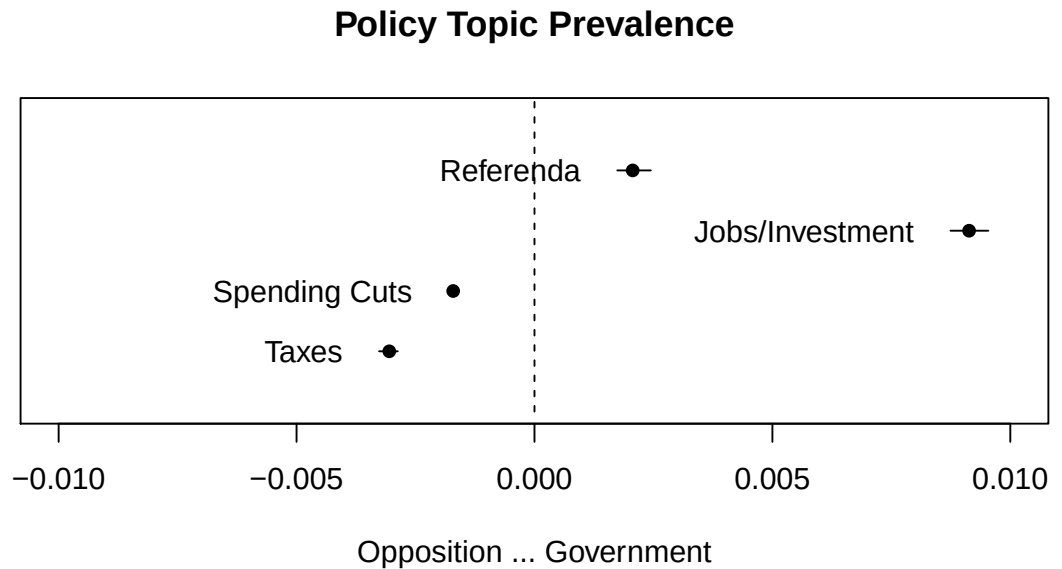
I primarily focus on economic policy issues as it is generally argued that the economy is an important factor that influences electoral outcomes (Powell Jr and Whitten, 1993; Alesina and Rosenthal, 1995; Lewis-Beck and Stegmaier, 2000). I expect that government legislators will discuss the issues of job creation and investment more than the opposition given the evidence that job creation is positively related to the electoral performance of government legislators (Lewis-Beck and Stegmaier, 2000). On the other hand, I expect that the opposition will emphasize high taxes more than government legislators given that research shows increases in taxes reduce support for governing parties (Niemi, Stanley and Vogel, 1995). Similarly, I expect the opposition to discuss cuts to government spending more than government legislators as governments prefer to highlight spending and transfer increases prior to elections (Rogoff, 1987). Lastly, with regard to referenda, I anticipate that government legislators are more likely to tweet more about this topic given the relevant constitutional amendments are originally proposed by the government.

To understand the link between government affiliation and legislators attention to these policy issues on Twitter, I follow Roberts et al. (2014) and estimate the relationship between the different political factors and topical prevalence, which captures how much a topic is discussed.⁸ In Figure 5.6, I present the difference in how much government and opposition legislators discuss these policy issues in their tweets. We can see that government legislators are more likely to discuss referenda than opposition legislators, potentially in an effort to convince voters to pass the proposed amendments. We can also see that government legislators, as expected, post tweets highlighting job creation and investment more than opposition legislators. In sharp contrast to this, opposition legislators appear to instead prefer to emphasize the spending cuts and taxes introduced by the government, which presumably negatively impact their con-

⁸Specifically, the *stm* package estimates the relationship between covariates and topical prevalence using a linear regression model where the outcome is the expected proportion of a tweet that is assigned to the given topic (Roberts, Stewart and Tingley, 2014).

stituents, when discussing the economy on Twitter. These results provide support to the idea that these topics boast construct validity for studying policy responsiveness.

Figure 5.6: Difference in government and opposition legislators' attention to policy issue



5.3.2 Allocation responsiveness on Twitter

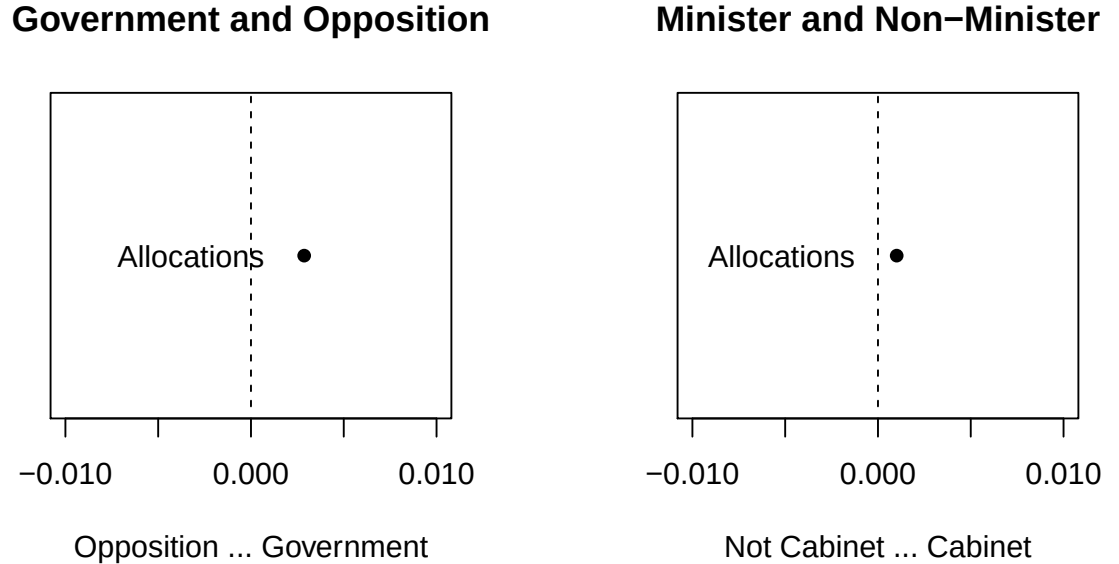
As noted in previous chapters of this work, another important component of legislators' representation of constituents is allocation responsiveness, which refers to the allocation of pork barrel goods that benefit specific constituencies or small groups of individuals in constituencies (Eulau and Karps, 1977). As briefly mentioned earlier in this section, one of estimated topics produced by the topic model of legislators' tweets pertains to legislators announcements about allocations for different local projects. I present sample representative tweets from this allocation topic in Figure 5.7.

Figure 5.7: Example tweets from allocations topic

Allocations
offaly lottery funding welcomed by the allocation of national lottery funding from the department of
delighted that sphere 17 regional youth services have received 50000 in funding for youth cafe today
good news announced regarding the skibbereen flood relief scheme confirmation received from dept of public
125000 for donegal festivals under 2016 festivals programme

To examine whether tweets can provide insight into allocation responsiveness, I follow a similar process as above and examine the relationship between different political factors and legislators’ discussion of allocations on Twitter. I explore how government affiliation and holding a ministerial position influence a legislator’s focus on allocations in tweets. I focus on these factors as the Minister of Finance in Ireland must approve of all government spending decisions, thereby increasing the likelihood that the government legislators will provide more allocations, and previous research suggests that legislators serving in the cabinet increase allocations to their district (Golden and Picci, 2008; Suiter and O’Malley, 2014).

Figure 5.8: Difference in legislators' attention to allocations



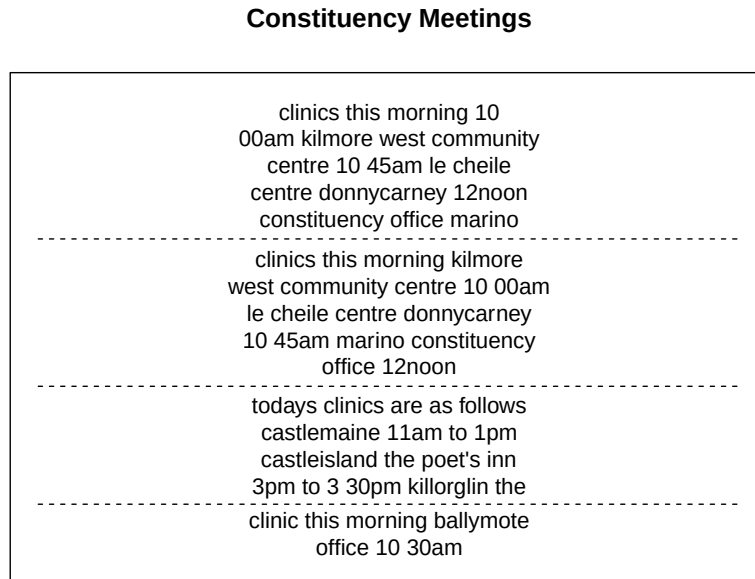
In Figure 5.8, I present the respective differences in how much government and opposition legislators and ministers and non-ministers discuss allocations in their tweets. As expected, we can see that government legislators appear to tweet more about local grants and funding than opposition legislators. Similarly, it also appears that cabinet ministers pay more attention to allocations in tweets than legislators not in the cabinet. The positive relationship between these factors and the topical prevalence of this measure provides further evidence that tweets can substantively contribute to our understanding of legislative behavior and political representation.

5.3.3 Service responsiveness on Twitter

I next focus on service responsiveness, the service that legislators provide to constituents or groups in their constituency, which is another important component of

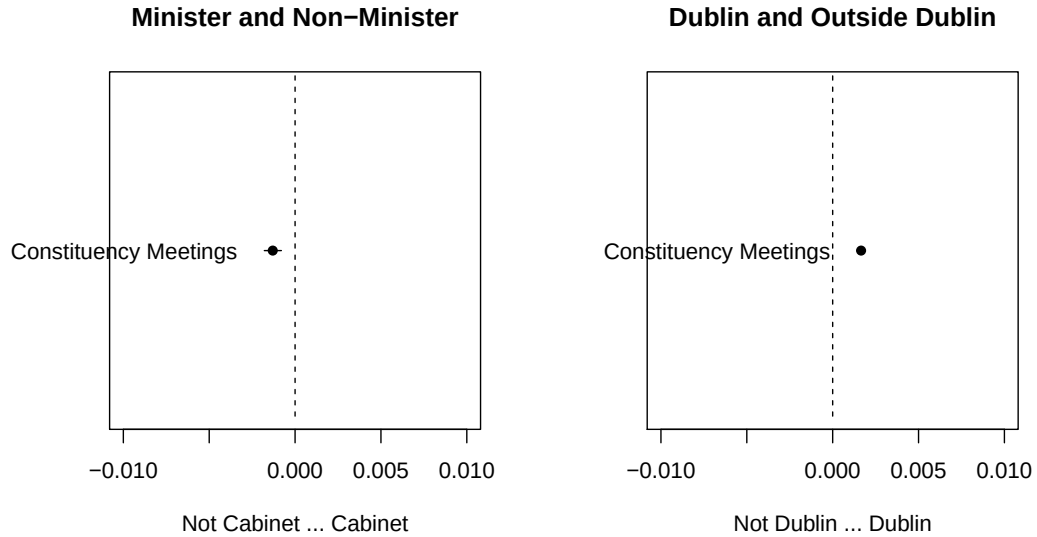
political representation (Eulau and Karps, 1977; Cain, Ferejohn and Fiorina, 1987). To do this, I investigate the relationship between holding ministerial office and serving in a Dublin office and legislators' promotion of constituency meetings in tweets.

Figure 5.9: Example tweets from constituency clinics topic



I focus on the estimated constituency clinics topic which captures legislators' announcements about local clinics and present example representative tweets from this topic in Figure 5.9. I expect that ministers will have less time to hold clinics due to the greater policy workload of a cabinet position and, in line with research on the differences between urban and rural politics (Clarke, 1978; Komito, 1992), that legislators serving in the more urbanized Dublin region will engage in less constituency work.

Figure 5.10: Differences in legislators' to constituency activity



In Figure 5.10, I present the respective differences in how much ministers and non-ministers and legislators serving in Dublin and outside of the Dublin region tweet about local constituency clinics. We can see that cabinet members appear to tweet less about constituency meetings than legislators that do not hold a ministerial position. On the other hand, and in contrast to prior expectations, it seems that legislators serving in Dublin constituencies tweet more about their constituency clinics than legislators serving in more rural geographical regions. One possible explanation for this is that urban legislators tweet more about constituency meetings in an attempt to highlight the service they provide while it is not necessary for rural legislators to do this. Despite this result, the constituency meetings topic provides information on an aspect of legislative activity that is not often found in other data sources.⁹

⁹In Ireland, legislators typically notify constituents about clinics using adverts in local publications, making it difficult to construct a comprehensive dataset. Unlike other contexts, a large proportion of legislators in Ireland do not have websites. It does appear that some legislators also use Facebook to post information about clinics. However, Facebook only provides very limited data for certain politicians' pages.

5.3.4 Legislator advertising on Twitter

As a final exercise for assessing the utility of tweets for studying representation, I focus on topics related to the advertising behavior of legislators. Mayhew (1974) argues that advertising, the effort to build a positive image using non-political content and engaging in non-political activities, is one of the primary activities of representatives. As discussed earlier, legislators appear to engage in substantial advertising behavior when tweeting. Here I focus on two topics that appear to represent advertising behavior, the topic capturing well wishes messages posted by legislators and the topic representing Irish language tweets. In Figure 5.11, I present example tweets from these topics.

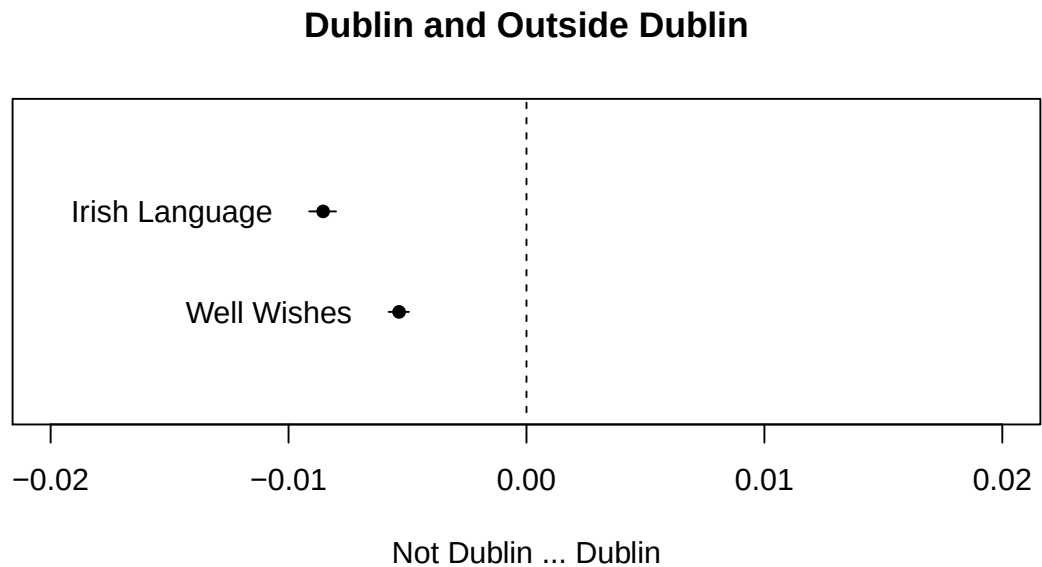
Figure 5.11: Example tweets from advertising topics

Irish Language	Well Wishes
agus anios ta mo leaba ag scairt orm oiche mhaith daoibhse ta suil agam go mbeidh bua ag aontroim amarach codladh samh zzzzzz and so 2 bed la fada a bhi ann s mise scrioste ach sasta mar rinne muid obair maith inniu oiche mhaith s codladh samh pancakes amarach	wishing the tipperary footballers all the best today in the munster senior football championship semi final in 2 best of luck to the ballyduff hurlers playing in the munster intermediate final in thurles today gaa

I expect that legislators serving in areas outside of the Dublin region will tweet more good luck messages since Gaelic games, the subject of the majority of these well wishes tweets, are very popular in more rural communities (Gallagher and Marsh,

2004). Similarly, I also expect more Irish language tweets from legislators in less urbanized areas as all Gaeltacht regions, areas where Irish is recognized as the primary language, lie outside of the Dublin geographic area.

Figure 5.12: Difference in urban and rural legislators' advertising behavior



In Figure 5.12, I present the differences in the prevalence of Irish language messages and tweets wishing individuals and teams well in upcoming events between legislators serving in Dublin and outside of the Dublin region. Once again, these results indicate the estimated topics from Twitter data have construct validity and can be utilized to better understand the types of advertising behavior legislators engage in.

The series of analyses presented above suggests that legislators employ tweets to discuss both political and non-political issues and, for the most part, legislators' use of Twitter matches our prior theoretical expectations. These validations also provide evidence that tweets are a useful resource for understanding and learning about legislators' behavior and representation of constituents.

5.4 Discussion

Using a new dataset comprised of all tweets by legislators that served in the 31st Dáil Éireann, I build on the growing literature pertaining to political activity on Twitter by providing a detailed examination of legislators’ utilization of Twitter. I show that the age of a legislator is the most important predictor for whether a legislator has a Twitter account and the age of a legislator, whether a legislator holds a ministerial role, and the geographic region of a legislator all influence the frequency of legislators’ Twitter usage. Using a statistical topic model to understand the subjects that legislators focus on in tweets, I find that legislators pay substantial attention to both political and non-political issues. I also demonstrate that a number of topics which legislators’ tweet about can be used to measure important components of political representation and legislative behavior. In a series validation exercises, I show that these topics generally covary with different political factors in line with theoretical expectations. This suggests that the content of legislators’ tweets is a useful resource for learning about legislators’ representation of constituents and generating new insights into legislative behavior.

In this work, I solely focus on the usage of Twitter by Irish legislators. However, Twitter is also utilized by a substantial proportion of legislators in a number of advanced democracies (Barberá, 2015). Therefore a natural next step in the study of Twitter usage by legislators is to compile a cross-national dataset of legislators’ tweets from a variety of countries. Given that tweets can be easily accessed by researchers, the construction of such a dataset is relatively straightforward compared to compiling a cross-national dataset of legislative speeches or questions. Such data would provide detailed information on the activities of a large number of political figures and would enable researchers to test theories of legislative behavior related to electoral systems and other institutional factors (Carey and Shugart, 1995).¹⁰

¹⁰Of course, analyzing legislators tweets from countries employing mixed electoral systems could

When legislators tweet they often post a picture in addition to written text. Given the proliferation of images on social media platforms in recent years, political scientists have unsurprisingly begun to explore how to utilize image data to better understand political processes (Anastasopoulos et al., 2016; Casas and Webb Williams, 2016). Legislators on Twitter post images of their political activities such as canvassing in addition to images of non-political activities such as the sporting events they attend. The classification of these images into categories closely tied to different constructs from the political representation literature would identify substantively meaningful variation in the different types of images that legislators post, thereby gaining a deeper insight into the priorities, activities, and role orientation of legislators.

also be employed to study such questions.

Chapter 6

Conclusion

In this work, I build on and contribute to the vast literature pertaining to legislators' representation of constituents. This literature is one of the most important in the field of political science as politicians' representation of constituents' interests plays a key role in democratic politics. In many democratic countries, how a legislator represents their constituents can play a decisive role in whether constituents vote to keep the legislator in office. It is therefore very important to understand how different political factors influence how a politician represents constituents. This work makes two primary contributions to this understanding. Firstly, I analyze how different data sources can be utilized to measure different elements of political representation, thereby contributing to the literature focused on the conceptualization and measurement of representation. Secondly, I examine how two important political factors, membership in a political dynasty and ethnic quotas, influence legislators' attention to different components of political representation. In the following paragraphs I provide more details about these contributions and discuss potential extensions and research projects related to this dissertation.

As noted in one of the first sections of this work, political representation can be seen as a multidimensional concept. However, this conceptualization also creates the

challenge of how researchers should measure and operationalize each of the components of political representation. In my first empirical chapter, I help to address this challenge by comparing legislators' speeches, questions, and tweets, three data sources which can typically be obtained for parliamentary democracies. In this analysis, I demonstrate that the majority of legislators communicate using all three platforms and each platform seems to provide good coverage of different legislative activities. I also find that legislators' speeches appear to be especially useful for understanding the policy responsiveness of legislators, that legislative questions are useful for measuring both the service and policy responsiveness of legislators, and that legislators' tweets can provide a unique insight into the advertising activity of legislators. The results of these content analysis models suggest that legislators utilize these communication platforms to emphasize different aspects of their legislative activities.

Building on this comparison and the literature studying legislators' activity using Twitter data, I also examined how different political factors relate to the frequency and content of legislators' tweets. In this exploratory analysis, I find that the age of a legislator, whether a legislator holds ministerial office, and whether a legislator serves in an urbanized geographic region all appear to influence the frequency of legislators' Twitter usage. Using topic models to understand the content of legislators' tweets, I also demonstrate that legislators discuss political issues as well as non-political issues in their tweets. In this analysis I also show that a number of the topics legislators tweet about can be utilized to study important concepts in political representation such as allocation responsiveness and legislators' advertising behavior. In a series of validation exercises, I demonstrate that the covariance between these topics and political factors generally meet theoretical expectations. This chapter contributes to the extant representation literature by highlighting legislators' tweets as a useful resource for studying political representation.

While this study highlights the utility of tweets for understanding the behavior of legislators, further analysis of the text of legislators' tweets should be completed.

A somewhat surprising finding from my analyses of legislators' tweets is that legislators in urban areas appear to tweet more about their constituency clinics than representatives serving in rural areas. One potential explanation for this result is that legislators in rural areas may simply use more traditional mediums to advertise their clinics. An alternative explanation is that constituency service and clinics play such a central role in the activities of legislators in more rural regions that there is no need to tweet about clinics. This indicates that further systematic investigation is required to understand legislators use of Twitter for constituency service.

Turning to the influence of different political factors on different aspects of political representation, I also examined the empirical association between membership in a political dynasty and the service responsiveness of legislators. I posited that the unique political socialization in which their previously elected relatives emphasize the importance of the cultivation of a personal vote and the social networks associated with dynasties lead to dynastic legislators providing more service than non-dynastic legislators. To investigate the relationship between membership in a political dynasty and service responsiveness I use a new dataset of parliamentary questions and demonstrate how supervised machine learning techniques can be utilized to generate a measure of service responsiveness from the text of questions. The results of a statistical analysis employing this measure indicate that the level of service responsiveness of dynastic legislators in rural areas is higher than that of non-dynastic legislators serving in rural areas. Contrastingly, these results also indicate that there is a negligible difference in the service responsiveness of dynastic and non-dynastic legislators in highly urbanized areas. This suggests that there may be systematic differences in the operations of dynasties in different regions. This chapter therefore sheds new light on the activities of dynastic legislators and contributes to the literature pertaining to political dynasties in addition to the representation literature.

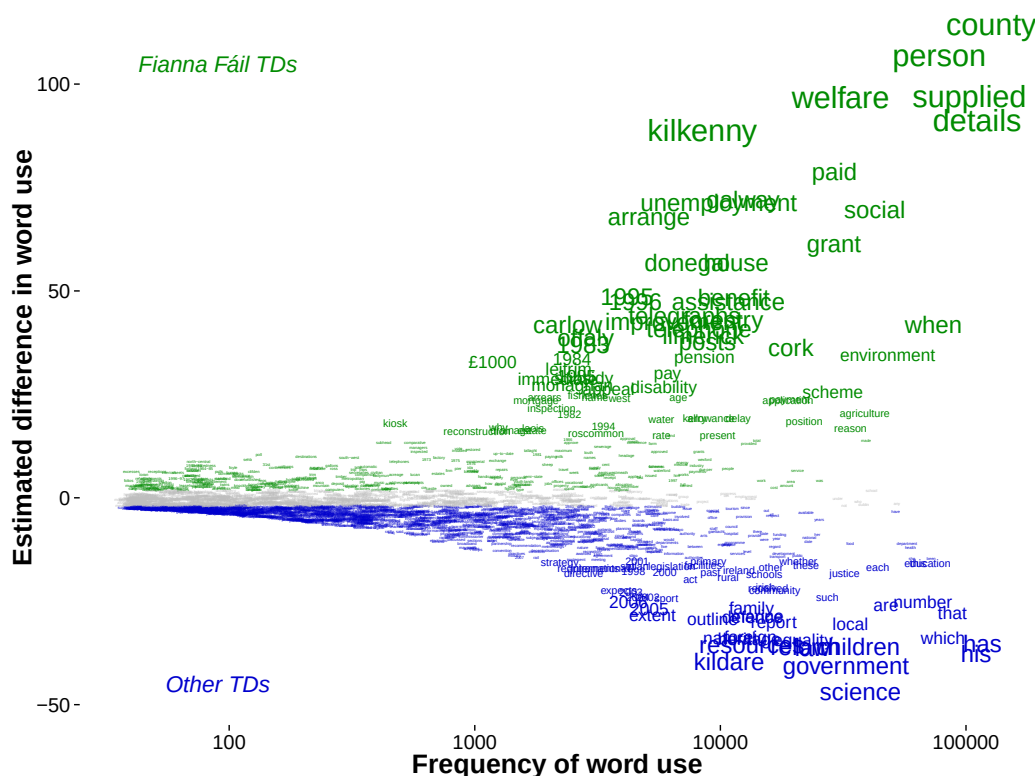
Similarly, I also investigated the relationship between ethnic quotas and the responsiveness of legislators to the interests of minority ethnic groups. This is an

important topic as a large number of countries have introduced special institutional arrangements to enhance the political representation of historically marginalized and disadvantaged groups. Despite this, there is a still debate regarding the effect of quotas on the substantive representation of such groups. I contribute to this debate by finding a strong positive relationship between election through an ethnic quota and the number of questions pertaining to ethnic minorities submitted by a legislator. This indicates that although the influence of quotas may not be visible when studying outcomes such as a constituency's level of development, quotas may nonetheless positively influence the efforts of legislators in the legislative process. However, there are still a number of important complexities related to quotas that need to be resolved by researchers and policymakers alike.

Beyond this analysis of the link between ethnic quotas and the responsiveness of legislators to the interests of minority ethnic groups, I do not extensively examine the role that institutions play in the representation process in this work. One institution which can influence a legislator's representation is the political party they are affiliated with. When modelling the empirical association between dynastic membership and service responsiveness, I also identified a positive relationship between membership in Fianna Fáil, the dominant party in Irish politics in the period of study, and legislators' focus on service. In an additional content analysis of legislative questions, I explore this relationship further and provide more evidence that Fianna Fáil legislators are more focused on constituency service than non-Fianna Fáil legislators.¹ In Figure 6.1 I present differences in language usage in parliamentary questions between Fianna Fáil and non-Fianna Fáil legislators.

¹I again utilize the informative Dirichlet prior model of Monroe, Colaresi and Quinn (2009) for this analysis. I also use the same question data that I employed in Chapter 4.

Figure 6.1: Differences in words usage between Fianna Fáil and non-Fianna Fáil legislators



In this plot the font size of words is proportional to the estimated difference in word usage. We can see that the vast majority of words most strongly associated with Fianna Fáil are representative of constituency service. Fianna Fáil legislators often submit questions pertaining to a *social welfare* payment or *grant* for a *person* in their *county*. This result is of particular interest as researchers have consistently found it difficult to differentiate Fianna Fáil and Fine Gael, traditionally the second largest political party in Ireland (Mair, 1987; Laver, Benoit and Garry, 2003; Gallagher and Marsh, 2004). One possible explanation for the difference presented here is the party structure of Fianna Fáil, which, until recently, differed significantly to that of other parties and was centred around small local party branches known as *cumann*

(Carty, 1976).² It estimated that at one time there were over 3000 of these grassroots organizations across Ireland (McGee, 2011). Each *cumann* was represented by a small number of delegates at constituency candidate selection meetings.³ This structure could therefore create incentives for Fianna Fáil legislators to engage in constituency service to help secure the party nomination from grassroots members. This brief discussion also demonstrates the potential of further investigation into the influence of institutions on legislators’ representation of constituents.

There are also a number of potential methodological research projects related to this work that could be completed in the future. In this project, I employ large datasets of legislative text and a variety of different text analytics tools to generate new insights on the topic of political representation. In recent years a number of new deep learning models and tools for analyzing text data have been developed by computer scientists (LeCun, Bengio and Hinton, 2015).⁴ A particularly important development in this domain is word vector representations, also known as word embeddings (Mikolov et al., 2013; Pennington, Socher and Manning, 2014). A word embedding represents a word as a k dimensional vector, where 100 and 300 are popular values of k . Compared to the bag of words representation utilized throughout this work, word embeddings more accurately capture the semantic meaning and similarity of words. Research has shown that word embeddings combined with modern neural network architectures such as convolutional and recurrent neural networks can improve the predictive performance of text classification and other natural language processing tasks (Kim, 2014; Chen and Manning, 2014). This approach is an obvious natural extension of the procedure I presented for classifying legislative questions as constituency service. More broadly, these new methods may also help political

²As noted by Gallagher and Marsh (2004), grassroots members are not particularly influential in Fine Gael.

³In recent years, Fianna Fáil’s party leadership has played a greater role in the selection of the party’s candidates.

⁴Deep learning is a subfield of machine learning pertaining to ‘deep’ neural network models, i.e. neural network models with multiple hidden layers.

scientists using text data to better understand political phenomena.

This project highlights the usefulness of utilizing computational social science methodologies to study representation. As outlined, this approach should also be useful for providing new insight into legislators' activities and improving our collective understanding of politicians' representation of constituents. I hope the research presented in this project provides a foundation for further work in this area.

Bibliography

- Achen, Christopher H. 1977. "Measuring representation: Perils of the correlation coefficient." *American Journal of Political Science* pp. 805–815.
- Achen, Christopher H. 1978. "Measuring representation." *American Journal of Political Science* pp. 475–510.
- Aldrich, John H and David W Rohde. 2001. The logic of conditional party government: Revisiting the electoral connection. In *Congress Reconsidered, Seventh Edition*, ed. Dodd, Lawrence C and Oppenheimer. Bruce I. CQ Press p. 269292.
- Alesina, Alberto and Howard Rosenthal. 1995. *Partisan politics, divided government, and the economy*. Cambridge University Press.
- Anastasopoulos, L Jason, Dhruvil Badani, Crystal Lee, Shiry Ginosar and Jake Williams. 2016. "Photographic home styles in Congress: a computer vision approach." *arXiv preprint arXiv:1611.09942* .
- Ansolabehere, Stephen, James M Snyder Jr and Charles Stewart III. 2000. "Old voters, new voters, and the personal vote: Using redistricting to measure the incumbency advantage." *American Journal of Political Science* pp. 17–34.
- Arulampalam, Wiji, Sugato Dasgupta, Amrita Dhillon and Bhaskar Dutta. 2009. "Electoral goals and center-state transfers: A theoretical model and empirical evidence from India." *Journal of Development Economics* 88(1):103–119.
- Asako, Yasushi, Takeshi Iida, Tetsuya Matsubayashi and Michiko Ueda. 2015. "Dynastic politicians: Theory and evidence from Japan." *Japanese Journal of Political Science* 16(01):5–32.
- Banerjee, Abhijit and Rohini Somanathan. 2007. "The political economy of public goods: Some evidence from India." *Journal of development Economics* 82(2):287–314.
- Barberá, Pablo. 2015. "Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data." *Political Analysis* 23(1).

- Barberá, Pablo, Richard Bonneau, Patrick Egan, John T Jost, Jonathan Nagler and Joshua Tucker. 2014. Leaders or followers? Measuring political responsiveness in the US Congress using social media data. In *110th American Political Science Association Annual Meeting*.
- Barnes, Samuel Henry. 1977. *Representation in Italy: institutionalized tradition and electoral choice*. University of Chicago Press.
- Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande and Petia Topalova. 2009. “Powerful women: does exposure reduce bias?” *The Quarterly journal of economics* 124(4):1497–1540.
- Beger, Andreas, Cassy L Dorff and Michael D Ward. 2016. “Irregular leadership changes in 2014: Forecasts using ensemble, split-population duration models.” *International Journal of Forecasting* 32(1):98–111.
- Berry, Christopher R and Anthony Fowler. 2015. “Cardinals or Clerics? Congressional Committees and the Distribution of Pork.” *American Journal of Political Science* .
- Besley, Timothy and Marta Reynal-Querol. 2015. “The Logic of Hereditary Rule: Theory and Evidence.” .
- Besley, Timothy, Rohini Pande and Vijayendra Rao. 2005. “Participatory democracy in action: Survey evidence from South India.” *Journal of the European Economic Association* 3(2-3):648–657.
- Bhavnani, Rikhil R. 2009. “Do electoral quotas work after they are withdrawn? Evidence from a natural experiment in India.” *American Political Science Review* 103(01):23–35.
- Bird, Karen. 2014. “Ethnic quotas and ethnic representation worldwide.” *International Political Science Review* 35(1):12–26.
- Blei, David M. 2012. “Probabilistic topic models.” *Communications of the ACM* 55(4):77–84.
- Blei, David M, Andrew Y Ng and Michael I Jordan. 2003. “Latent dirichlet allocation.” *Journal of machine Learning research* 3(Jan):993–1022.
- Blei, David M and John D Lafferty. 2007. “A correlated topic model of science.” *The Annals of Applied Statistics* pp. 17–35.
- Boylan, Christopher, Ryan McMahon and Burt L Monroe. 2016. “Racially Differentiated Language in NFL Scouting Reports.” *ACM SIGKDD 16 Workshop on Large Scale Sports Analytics* .

- Braganca, Arthur, Claudio Ferraz and Juan Rios. 2015. "Political dynasties and the quality of government."
- Butler, Daniel M, Christopher F Karpowitz and Jeremy C Pope. 2012. "A field experiment on legislators home styles: Service versus policy." *The Journal of Politics* 74(02):474–486.
- Butler, Daniel M and David E Broockman. 2011. "Do politicians racially discriminate against constituents? A field experiment on state legislators." *American Journal of Political Science* 55(3):463–477.
- Cain, Bruce, John Ferejohn and Morris Fiorina. 1987. *The personal vote: Constituency service and electoral independence*. Harvard University Press.
- Carey, John M and Matthew Soberg Shugart. 1995. "Incentives to cultivate a personal vote: A rank ordering of electoral formulas." *Electoral studies* 14(4):417–439.
- Carson, Jamie L, Erik J Engstrom and Jason M Roberts. 2007. "Candidate quality, the personal vote, and the incumbency advantage in congress." *American Political Science Review* 101(02):289–301.
- Carson, Jamie L and Jeffery A Jenkins. 2011. "Examining the electoral connection across time." *Annual Review of Political Science* 14:25–46.
- Carty, RK. 1976. "Social cleavages and party systems: a reconsideration of the Irish case." *European Journal of Political Research* 4(2):195–203.
- Casas, Andreu and Nora Webb Williams. 2016. "Images that Matter: Online Protests and the Mobilizing Role of Pictures."
- Chandra, Kanchan and Wamiq Umaira. 2011. "Indias democratic dynasties." *New York University* .
- Chattopadhyay, Raghavendra and Esther Duflo. 2004. "Women as policy makers: Evidence from a randomized policy experiment in India." *Econometrica* 72(5):1409–1443.
- Chen, Danqi and Christopher D Manning. 2014. A Fast and Accurate Dependency Parser using Neural Networks. In *EMNLP*. pp. 740–750.
- Chen, Liu Cheun. 2016. "Ram Vilas Paswan demands reservation for SCs, STs in private sector." *India Today* .
- Chhibber, Pradeep and Ken Kollman. 1998. "Party aggregation and the number of parties in India and the United States." *American political Science review* 92(02):329–342.

- Clarke, Harold D. 1978. "Determinants of provincial constituency service behaviour: A multivariate analysis." *Legislative Studies Quarterly* pp. 601–628.
- Clarke, Harold D, Richard G Price and Robert Krause. 1975. "Constituency service among Canadian Provincial Legislators: Basic findings and a test of three hypotheses." *Canadian Journal of Political Science* 8(04):520–542.
- Clinton, Joshua D. 2006. "Representation in Congress: constituents and roll calls in the 106th House." *Journal of Politics* 68(2):397–409.
- Collins, Stephen. 1993. *Spring and the Labour story*. O'Brien Press.
- Conde, Carlos H. 2007. "Family dynasties bind politics in Philippines." *Reuters* .
- Converse, Philip E. and Roy Pierce. 1986. *Political Representation in France*. Harvard University Press.
- Cox, Gary W and Mathew D McCubbins. 2005. *Setting the agenda: Responsible party government in the US House of Representatives*. Cambridge University Press.
- Cranmer, Skyler J and Bruce A Desmarais. 2017. "What can we Learn from Predictive Modeling?" *Political Analysis* .
- Cruz, Cesi, Julien Labonne and Pablo Querubin. 2017. "Politician family networks and electoral outcomes: Evidence from the Philippines." *American Economic Review* .
- Dal Bó, Ernesto, Pedro Dal Bó and Jason Snyder. 2009. "Political dynasties." *The Review of Economic Studies* 76(1):115–142.
- D'Orazio, Vito, Steven T Landis, Glenn Palmer and Philip Schrodtt. 2014. "Separating the Wheat from the Chaff: Applications of Automated Document Classification Using Support Vector Machines." *Political analysis* 22(2).
- Downs, Anthony. 1957. "An economic theory of political action in a democracy." *The journal of political economy* pp. 135–150.
- Dufo, Esther. 2005. "Why political reservations?" *Journal of the European Economic Association* 3(2-3):668–678.
- Dunning, Thad and Janhavi Nilekani. 2013. "Ethnic quotas and political mobilization: caste, parties, and distribution in Indian village councils." *American Political Science Review* 107(01):35–56.
- Ellickson, Mark C and Donald E Whistler. 2001. "Explaining state legislators' casework and public resource allocations." *Political Research Quarterly* 54(3):553–569.

- Erikson, Robert S. 1978. "Constituency opinion and congressional behavior: A reexamination of the Miller-Stokes representation data." *American Journal of Political Science* pp. 511–535.
- Eulau, Heinz and Paul D Karps. 1977. "The puzzle of representation: Specifying components of responsiveness." *Legislative Studies Quarterly* pp. 233–254.
- Evans, Heather K, Victoria Cordova and Savannah Sipole. 2014. "Twitter style: An analysis of how house candidates used twitter in their 2012 campaigns." *PS: Political Science & Politics* 47(02):454–462.
- Fallon, Johnny. 2011. *Dynasties: Irish Political Families*. New Island.
- Fariss, Christopher J, Fridolin J Linder, Zachary M Jones, Charles D Crabtree, Megan A Biek, Ana-Sophia M Ross, Taranamol Kaur and Michael Tsai. 2015. "Human rights texts: converting human rights primary source documents into data." *PloS one* 10(9):e0138935.
- Fariss, Christopher J and Zachary M Jones. 2017. "Enhancing Validity in Observational Settings When Replication is Not Possible." *Political Science Research and Methods* pp. 1–16.
- Farrell, David M and Roger Scully. 2007. *Representing Europe's citizens?: electoral institutions and the failure of parliamentary representation*. Oxford University Press.
- Feinstein, Brian D. 2010. "The dynasty advantage: Family ties in congressional elections." *Legislative Studies Quarterly* 35(4):571–598.
- Fenno, Richard F. 1977. "US House members in their constituencies: An exploration." *American Political Science Review* 71(03):883–917.
- Fenno, Richard F. 1978. *Home style: House members in their districts*. Harper Collins.
- Ferejohn, John. 1986. "Incumbent performance and electoral control." *Public choice* 50(1):5–25.
- Fiorina, Morris P. 1977. "The case of the vanishing marginals: The bureaucracy did it." *American Political Science Review* 71(01):177–181.
- Fiorina, Morris P. 1981. *Retrospective voting in American national elections*. Yale University Press.
- Fiorina, Morris P. 1989. *Congress, keystone of the Washington establishment*. Yale University Press.

- Folke, Olle, Johanna Rickne and Daniel M Smith. 2016. "Gender and Dynastic Political Recruitment: Theory and Evidence."
- Folke, Olle, Lenita Freidenvall and Johanna Rickne. 2015. "Gender quotas and ethnic minority representation: Swedish evidence from a longitudinal mixed methods study." *Politics & Gender* 11(02):345–381.
- Fowler, Anthony and Andrew B Hall. 2015. "Congressional seniority and pork: A pig fat myth?" *European Journal of Political Economy* 40:42–56.
- Franceschet, Susan and Jennifer M Piscopo. 2008. "Gender quotas and women's substantive representation: Lessons from Argentina." *Politics & Gender* 4(03):393–425.
- Franklin, Mark N and Philip Norton. 1993. *Parliamentary Questions: For the Study of Parliament Group*. Oxford University Press, USA.
- Freeman, Patricia K and Lilliard E Richardson Jr. 1996. "Explaining variation in casework among state legislators." *Legislative Studies Quarterly* pp. 41–56.
- Gaines, Brian J. 1998. "The impersonal vote? Constituency service and incumbency advantage in British elections, 1950-92." *Legislative Studies Quarterly* pp. 167–195.
- Gallagher, Michael and Lee Komito. 2006. The constituency role of Dáil deputies. In *Politics in the Republic of Ireland*, ed. John Coakley and Michael Gallagher. Psychology Press.
- Gallagher, Michael and Michael Marsh. 2004. "Party membership in Ireland: the members of Fine Gael." *Party Politics* 10(4):407–425.
- Geys, Benny and Daniel M Smith. 2016. "Political Dynasties in Democracies: Causes, Consequences, and Remaining Puzzles." *Economic Journal*.
- Golbeck, Jennifer, Justin M Grimes and Anthony Rogers. 2010. "Twitter use by the US Congress." *Journal of the American Society for Information Science and Technology* 61(8):1612–1621.
- Golden, Miriam A and Lucio Picci. 2008. "Pork-barrel politics in postwar Italy, 1953–94." *American Journal of Political Science* 52(2):268–289.
- Golden, Miriam and Brian Min. 2013. "Distributive politics around the world." *Annual Review of Political Science* 16:73–99.
- Golder, Matt and Jacek Stramski. 2010. "Ideological congruence and electoral institutions." *American Journal of Political Science* 54(1):90–106.

- Grimmer, Justin. 2010. "A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases." *Political Analysis* pp. 1–35.
- Grimmer, Justin. 2013a. "Appropriators not position takers: The distorting effects of electoral incentives on congressional representation." *American Journal of Political Science* 57(3):624–642.
- Grimmer, Justin. 2013b. *Representational Style in Congress: What Legislators Say and Why It Matters*. Cambridge University Press.
- Grimmer, Justin and Brandon M Stewart. 2013. "Text as data: The promise and pitfalls of automatic content analysis methods for political texts." *Political analysis* pp. 267–297.
- Grimmer, Justin, Solomon Messing and Sean J Westwood. 2012. "How words and money cultivate a personal vote: The effect of legislator credit claiming on constituent credit allocation." *American Political Science Review* 106(04):703–719.
- Grofman, Bernard. 1999. *Elections in Japan, Korea, and Taiwan under the single non-transferable vote: The comparative study of an embedded institution*. University of Michigan Press.
- Grose, Christian R, Neil Malhotra and Robert Parks Van Houweling. 2015. "Explaining explanations: How legislators explain their policy positions and how citizens react." *American Journal of Political Science* 59(3):724–743.
- Hall, Richard L. 1998. *Participation in congress*. Yale Univ Pr.
- Harden, Jeffrey J. 2013. "Multidimensional Responsiveness: The Determinants of Legislators' Representational Priorities." *Legislative Studies Quarterly* 38(2):155–184.
- Harden, Jeffrey J. 2016. *Multidimensional democracy: A supply and demand theory of representation in American legislatures*. Cambridge University Press.
- Hastie, Trevor, Robert Tibshirani and Jerome Friedman. 2009. *The elements of statistical learning*. Springer.
- Hemphill, Libby, Aron Culotta and Matthew Heston. 2013. "Framing in social media: How the us congress uses twitter hashtags to frame political issues.".
- Hill, Daniel W and Zachary M Jones. 2014. "An empirical evaluation of explanations for state repression." *American Political Science Review* 108(03):661–687.

- Höllinger, Franz and Max Haller. 1990. "Kinship and social networks in modern societies: a cross-cultural comparison among seven nations." *European Sociological Review* 6(2):103–124.
- Election Commission of India. 2017. "Delimitation of Constituencies."
- Election Commission of India. 2016. "FAQs - Delimitation of Constituencies."
- Parliament of India. 2016. "Rules of Procedure and Conduct of Business in Lok Sabha."
- Central Statistics Office Ireland. 2011. "Profile 1 Town and Country."
- Houses of the Oireachtas. 2016. "A Brief Guide To How Your Parliament Works."
- Parliament of India. 2016. "Question hour in Lok Sabha."
- Htun, Mala. 2004. "Is gender like ethnicity? The political representation of identity groups." *Perspectives on Politics* 2(03):439–458.
- Huber, John D and G Bingham Powell. 1994. "Congruence between citizens and policymakers in two visions of liberal democracy." *World Politics* 46(03):291–326.
- Hughes, Melanie M. 2011. "Intersectionality, quotas, and minority women's political representation worldwide." *American Political Science Review* pp. 604–620.
- Jackson, John E and David C King. 1989. "Public Goods, Private Interests, and Representation." *American Political Science Review* 83(04):1143–1164.
- James, Gareth, Daniela Witten, Trevor Hastie and Robert Tibshirani. 2013. *An introduction to statistical learning*. Vol. 6 Springer.
- Jensenius, Francesca Refsum. 2013. "Power, performance and bias: Evaluating the electoral quotas for scheduled castes in India."
- Jensenius, Francesca Refsum. 2015. "Development from representation? A study of quotas for the scheduled castes in India." *American Economic Journal: Applied Economics* 7(3):196–220.
- Jurafsky, Dan and James H. Martin. 2017. *Speech & Language Processing (Third Edition)*.
- Jurafsky, Dan, Victor Chahuneau, Bryan R Routledge and Noah A Smith. 2014. "Narrative framing of consumer sentiment in online restaurant reviews." *First Monday* 19(4).

- Kalla, Joshua L and David E Broockman. 2015. "Campaign contributions facilitate access to congressional officials: A randomized field experiment." *American Journal of Political Science* 60(3):545–558.
- Kennedy, Geraldine. 2002. *The Irish Times Nealon's Guide to the 29th Dáil and Seanad*. Gill & Macmillan.
- Kim, Yoon. 2014. "Convolutional neural networks for sentence classification." *arXiv preprint arXiv:1408.5882*.
- Komito, Lee. 1989. Dublin politics: symbolic dimensions of clientelism. In *Ireland from Below: Social Change and Local Communities*, ed. Chris Curtin and Thomas Wilson. pp. 240–259.
- Komito, Lee. 1992. "Brokerage or friendship? Politics and networks in Ireland." *Economic and Social Review* 23(2):129.
- Krook, Mona Lena and Diana Z O'Brien. 2010. "The politics of group representation: Quotas for women and minorities worldwide." *Comparative Politics* 42(3):253–272.
- Kurtz, Donn M. 1989. "The political family: A contemporary view." *Sociological Perspectives* 32(3):331–352.
- Laband, David N and Bernard F Lentz. 1985. "Favorite sons: Intergenerational wealth transfers among politicians." *Economic Inquiry* 23(3):395–414.
- Labonne, Julien, Sahar Parsa and Pablo Querubin. 2015. "Political Dynasties, Term Limits and Female Political Empowerment: Evidence from the Philippines."
- Lassen, David S and Adam R Brown. 2011. "Twitter: The electoral connection?" *Social Science Computer Review* 29(4):419–436.
- Laver, Michael, Kenneth Benoit and John Garry. 2003. "Extracting policy positions from political texts using words as data." *American Political Science Review* 97(02):311–331.
- LeCun, Yann, Yoshua Bengio and Geoffrey Hinton. 2015. "Deep learning." *Nature* 521(7553):436–444.
- Lee, David S, Enrico Moretti and Matthew J Butler. 2004. "Do voters affect or elect policies? Evidence from the US House." *The Quarterly Journal of Economics* 119(3):807–859.
- Levitt, Steven D and Catherine D Wolfram. 1997. "Decomposing the sources of incumbency advantage in the US House." *Legislative Studies Quarterly* pp. 45–60.

- Lewis-Beck, Michael S and Mary Stegmaier. 2000. "Economic determinants of electoral outcomes." *Annual Review of Political Science* 3(1):183–219.
- Lizzeri, Alessandro and Nicola Persico. 2001. "The provision of public goods under alternative electoral incentives." *American Economic Review* pp. 225–239.
- Lopez, Linette. 2011. "The Political Dynasties That Destroyed Greece, And The Real Reason The Country Is Such A Mess." *Reuters* .
- Lucas, Christopher, Richard A Nielsen, Margaret E Roberts, Brandon M Stewart, Alex Storer and Dustin Tingley. 2015. "Computer-assisted text analysis for comparative politics." *Political Analysis* 23(2):254–277.
- Lucey, Anne. 2016. "Danny Healy-Rae to run in Kerry constituency." *Irish Times* .
- Mair, Peter. 1987. *The Changing Irish Party System: Organisation, Ideology and Electoral Competition*. Burns & Oates.
- Mansbridge, Jane. 1999. "Should blacks represent blacks and women represent women? A contingent" yes"." *The Journal of politics* 61(3):628–657.
- Marsh, Michael. 2007. "Candidates or parties? Objects of electoral choice in Ireland." *Party Politics* 13(4):500–527.
- Martin, Shane. 2011*a*. "Parliamentary questions, the behaviour of legislators, and the function of legislatures: An introduction." *The Journal of Legislative Studies* 17(3):259–270.
- Martin, Shane. 2011*b*. "Using parliamentary questions to measure constituency focus: An application to the Irish case." *Political Studies* 59(2):472–488.
- Mayhew, David R. 1974. *Congress: The electoral connection*. Yale University Press.
- McClendon, Gwyneth H. 2016. "Race and responsiveness: An experiment with South African politicians." *Journal of Experimental Political Science* 3(01):60–74.
- McGee, Harry. 2011. "Fianna Fáil has lost the local knowledge. The grassroots are not being listened to." *Irish Times* .
- McGee, Harry. 2016. "Michael Healy Rae: Flat of cap, slight of figure, big of personality." *Irish Times* .
- Mendoza, Ronald U, Edsel L Beja Jr, Victor S Venida and David B Yap. 2012. "Inequality in democracy: Insights from an empirical analysis of political dynasties in the 15th Philippine Congress." *Philippine Political Science Journal* 33(2):132–145.

- Mezey, Michael L. 1994. "New perspectives on parliamentary systems: a review article." *Legislative Studies Quarterly* .
- Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*. pp. 3111–3119.
- Milesi-Ferretti, Gian Maria, Roberto Perotti and Massimo Rostagno. 2002. "Electoral systems and public spending." *Quarterly Journal of Economics* pp. 609–657.
- Miller, Wakken E and Donald E Stokes. 1963. "Constituency influence in Congress." *American political science review* 57(01):45–56.
- Minhas, Shahryar, Jay Ulfelder and Michael D Ward. 2015. "Mining texts to efficiently generate global data on political regime types." *Research & Politics* 2(3):2053168015589217.
- Monroe, Burt L and Amanda G Rose. 2002. "Electoral systems and unimagined consequences: Partisan effects of districted proportional representation." *American Journal of Political Science* pp. 67–89.
- Monroe, Burt L and Ko Maeda. 2004. Talks cheap: Text-based estimation of rhetorical ideal-points. In *annual meeting of the Society for Political Methodology*. pp. 29–31.
- Monroe, Burt L, Michael P Colaresi and Kevin M Quinn. 2009. "Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict." *Political Analysis* p. mpn018.
- Montgomery, Douglas C. 1991. *Design and analysis of experiments*.
- Nanivadekar, Medha. 2006. "Are quotas a good idea? The Indian experience with reserved seats for women." *Politics & Gender* 2(01):119–128.
- Nealon, Ted. 1973. *Nealon's Guide to the 20th Dail and Seanad*. Dublin, Platform Press.
- Nealon, Ted. 1977. *Nealon's Guide to the 21st Dail and Seanad*. Dublin, Platform Press.
- Nealon, Ted. 1981. *Nealon's Guide to the 22nd Dail and Seanad*. Dublin, Platform Press.
- Nealon, Ted. 1982. *Nealon's Guide to the 23rd Dail and Seanad*. Dublin, Platform Press.

- Nealon, Ted. 1983. *Nealon's Guide to the 24th Dail and Seanad*. Dublin, Platform Press.
- Nealon, Ted. 1987. *Nealon's Guide to the 25th Dail and Seanad*. Dublin, Platform Press.
- Nealon, Ted. 1993. *Nealon's Guide to the 27th Dail and Seanad*. Dublin, Platform Press.
- Nealon, Ted. 1997. *Nealon's Guide to the 28th Dail and Seanad*. Dublin, Platform Press.
- Niemi, Richard G, Harold W Stanley and Ronald J Vogel. 1995. "State economies and state taxes: Do voters hold governors accountable?" *American Journal of Political Science* pp. 936–957.
- Norton, Philip and David Michael Wood. 1993. *Back from Westminster: British Members of Parliament and their constituents*. Cambridge Univ Press.
- O' Shea, Owen. 2011. *Heirs to the Kingdom: Kerry's Political Dynasties*. O'Brien Press.
- O'Connor, Niall. 2016. "Oireachtas printing controversy: TD insists his 72,500 calendars 'are not for Christmas'." *Irish Independent* .
- Page, Benjamin I and Robert Y Shapiro. 1983. "Effects of public opinion on policy." *American political science review* 77(01):175–190.
- Pennington, Jeffrey, Richard Socher and Christopher D Manning. 2014. Glove: Global Vectors for Word Representation. In *EMNLP*. Vol. 14 pp. 1532–1543.
- Persson, Torsten and Guido Enrico Tabellini. 2002. *Political economics: explaining economic policy*. MIT press.
- Phillips, Anne. 1995. *The politics of presence*. Clarendon Press.
- Pitkin, Hanna Fenichel. 1967. *The concept of representation*. Univ of California Press.
- Poole, Keith T and Howard Rosenthal. 1985. "A spatial model for legislative roll call analysis." *American Journal of Political Science* pp. 357–384.
- Powell, G Bingham. 2000. *Elections as instruments of democracy: Majoritarian and proportional visions*. Yale University Press.
- Powell, G Bingham. 2004. "Political representation in comparative politics." *Annu. Rev. Polit. Sci.* 7:273–296.

- Powell, G Bingham and Georg S Vanberg. 2000. "Election laws, disproportionality and median correspondence: Implications for two visions of democracy." *British Journal of Political Science* 30(03):383–411.
- Powell Jr, G Bingham and Guy D Whitten. 1993. "A cross-national analysis of economic voting: taking account of the political context." *American Journal of Political Science* pp. 391–414.
- Proksch, Sven-Oliver and Jonathan B Slapin. 2012. "Institutional foundations of legislative speech." *American Journal of Political Science* 56(3):520–537.
- Querubin, Pablo. 2016. "Family and politics: Dynastic persistence in the Philippines." *Quarterly Journal of Political Science* 11:151–181.
- Quinn, Kevin M, Burt L Monroe, Michael Colaresi, Michael H Crespin and Dragomir R Radev. 2010. "How to analyze political attention with minimal assumptions and costs." *American Journal of Political Science* 54(1):209–228.
- Radford, Jason and Betsy Sinclair. 2016. "Electronic Homestyle: Using Twitter to Estimate Partisan Ideal Points for Members of Congress."
- Rahman, Ashikur. 2013. Essays on political dynasties: evidence from empirical investigations PhD thesis London School of Economics and Political Science.
- Roberts, Margaret E, Brandon M Stewart and Dustin Tingley. 2014. "stm: R package for structural topic models."
- Roberts, Margaret E, Brandon M Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson and David G Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58(4):1064–1082.
- Roberts, Margaret E, Brandon M Stewart and Edoardo M Airolidi. 2016. "A model of text for experimentation in the social sciences." *Journal of the American Statistical Association* 111(515):988–1003.
- Rogoff, Kenneth S. 1987. "Equilibrium political budget cycles."
- Rossi, Martín. 2014. "Family Business: Causes and Consequences of Political Dynasties."
- Saalfeld, Thomas. 2011. "Parliamentary questions as instruments of substantive representation: Visible minorities in the UK House of Commons, 2005–10." *The Journal of Legislative Studies* 17(3):271–289.

- Scholl, Edward L. 1986. "The electoral system and constituency-oriented activity in the European Parliament." *International Studies Quarterly* 30(3):315–332.
- Schwarz, Daniel, Denise Traber and Kenneth Benoit. 2015. "Estimating intra-party preferences: comparing speeches to votes." *Political Science Research and Methods* pp. 1–18.
- Shepsle, Kenneth A and Barry R Weingast. 1981. "Political preferences for the pork barrel: A generalization." *American journal of political science* pp. 96–111.
- Shugart, Matthew Søberg, Melody Ellis Valdini and Kati Suominen. 2005. "Looking for locals: voter information demands and personal vote-earning attributes of legislators under proportional representation." *American Journal of Political Science* 49(2):437–449.
- Slapin, Jonathan B and Sven-Oliver Proksch. 2010. "Look whos talking: Parliamentary debate in the European Union." *European Union Politics* 11(3):333–357.
- Smith, Daniel Markham. 2012. *Succeeding in politics: dynasties in democracies* PhD thesis University of California, San Diego.
- Smith, Daniel Markham and Shane Martin. 2016. "Political Dynasties and the selection of cabinet ministers." *Legislative Studies Quarterly* .
- Spirling, Arthur and Iain McLean. 2007. "UK OC OK? Interpreting optimal classification scores for the UK House of Commons." *Political Analysis* pp. 85–96.
- Spirling, Arthur and Kevin Quinn. 2010. "Identifying intraparty voting blocs in the UK House of Commons." *Journal of the American Statistical Association* 105(490):447–457.
- Stimson, James A, Michael B MacKuen and Robert S Erikson. 1995. "Dynamic representation." *American Political Science Review* 89(03):543–565.
- Stratmann, Thomas and Martin Baur. 2002. "Plurality Rule, Proportional Representation, and the German Bundestag: How Incentives to Pork-Barrel Differ across Electoral Systems." *American Journal of Political Science* pp. 506–514.
- Suiter, Jane and Eoin O'Malley. 2014. "Yes, minister: the impact of decision-making rules on geographically targeted particularistic spending." *Parliamentary Affairs* 67(4):935–954.
- Thachil, Tariq and Emmanuel Teitelbaum. 2015. "Ethnic Parties and Public Spending New Theory and Evidence From the Indian States." *Comparative Political Studies* .

- Trochim, William MK and James P Donnelly. 2001. *Research methods knowledge base*. Atomic Dog Pub.
- Turner, Lorraine and Carmel Crimmins. 2011. "Irish want political overhaul; end to cronyism." *Reuters* .
- Urbinati, Nadia and Mark E Warren. 2008. "The concept of representation in contemporary democratic theory." *Annual Review of Political Science* 11:387–412.
- Vaishnav, Milan and Neelanjan Sircar. 2011. "Core or swing? The role of electoral context in shaping pork barrel."
- Ward, Hugh and Peter John. 1999. "Targeting benefits for electoral gain: Constituency marginality and the distribution of grants to English local authorities." *Political Studies* 47(1):32–52.
- Ward, Michael D, Brian D Greenhill and Kristin M Bakke. 2010. "The perils of policy by p-value: Predicting civil conflicts." *Journal of Peace Research* 47(4):363–375.
- Weingast, Barry R and William J Marshall. 1988. "The industrial organization of Congress; or, why legislatures, like firms, are not organized as markets." *The Journal of Political Economy* pp. 132–163.
- Xydias, Christina V. 2007. "Inviting More Women to the Party: Gender Quotas and Women's Substantive Representation in Germany." *International Journal of Sociology* 37(4):52–66.
- Yadav, Vineeta. 2015. "Studying legislative party politics in data scarce environments A new empirical approach." *Party Politics* .

Appendix

A.1 Technical Details

In the following sections I outline various technical aspects from different sections in this work.

A.1.1 Text as data

In this section I outline the steps for creating and processing the various text datasets used in this project.

Ireland legislative text data

The majority of the legislative text data used in this project was originally collected and cleaned for the Legislative Speech Projects, directed by Burt Monroe. Specifically, the data covering the period from 1918 to 2007 was converted into a structured XML format by Cheryl Monroe. I extend the temporal scope of this data by including all legislative speeches and questions from 2007 to 2016. To obtain this data I wrote custom functions in Python which identified and scraped the relevant webpages from the website of Dáil Éireann. In total, I scraped over 190,000 of these webpages. The format of these webpages is HTML and differs from the HTML originally obtained in the Legislative Speech Projects.

Figure A.1: Example of a legislative text HTML page



Houses of the
Oireachtas
Tithe an Oireachtais

[Current Debates](#) > [Dáil Debates](#) > [2015](#)

- ▶ **Written Answers Nos. 1 - 14**
- ▶ **Pyrite Remediation Programme Implementation**
- ▶ **Local Authority Housing Provision**
- ▶ **Water Charges Yield**
- ▶ **Local Authority Housing Provision**
- ▶ **Housing Estates**

Dáil Éireann Debate
Vol. 870 No. 2

⏪ ⏩

Page 69 of 99

⏪ ⏩

Wednesday, 4 March 2015

Written Answers Nos. 1 - 14

The following are questions tabled by Members for written response and the ministerial replies as received on the day from the Departments [unrevised].

Questions Nos. 1 to 9, inclusive, answered orally.

Pyrite Remediation Programme Implementation

10. Deputy Clare Daly asked the **Minister for the Environment, Community and Local Government** his views regarding the impact of extending the number of areas from which the Pyrite Resolution Board can receive applications; if extra funding and staff have been made available as a result of this decision; and if he is satisfied that sufficient funding is in place to remediate the number of affected dwellings. [8864/15]

Minister of State at the Department of the Environment, Community and Local Government (Deputy Paudie Coffey): I am satisfied that the Pyrite Resolution Board has the resources, both human and financial, to meet the needs of the pyrite remediation scheme.

Additionally, in 2012 the format of the Dáil website was significantly altered. I therefore wrote a further series of original Python functions to extract the legislative text and relevant corresponding metadata and export this information to JSON format. The functions for scraping and extracting the data could be reused to maintain an up-to-date public database of legislative speeches and questions from Dáil Éireann. For the analyses of these speeches and questions presented in Chapter 2 and Chapter 3, I removed links, punctuation, and digits from the text of speeches and questions. I also normalized accented Irish characters and converted every word into lower case

format.

Ireland Twitter data

To get the tweets of Irish TDs that served between 2011 and 2016, I first obtained the Twitter usernames of TDs with Twitter accounts from biographical sketches on the website of Dáil Éireann. I then wrote a custom function in Python which utilized the Tweepy Python library to access the Twitter API and download the Twitter feeds of TDs. The format of these downloads is JSON. I wrote a further series of functions to extract and convert the text of tweets and relevant metadata from this form. I preprocessed this data by excluding retweets as well as tweets that are replies to other Twitter users. For the analysis of tweets in Chapter 2, I removed hashtags, links, Twitter handles, punctuation, and digits from the text of Tweets. I again normalized accented Irish characters and converted every word into lower case format. For the topic model presented in Chapter 5, I again removed links, Twitter handles, punctuation, and digits from the text of Tweets. For this chapter, I also removed stopwords and words that did not occur in at least 50 of the tweets in the corpus while retaining hashtags. In Figure A.2, I present an example tweet in the format returned by the Twitter API.

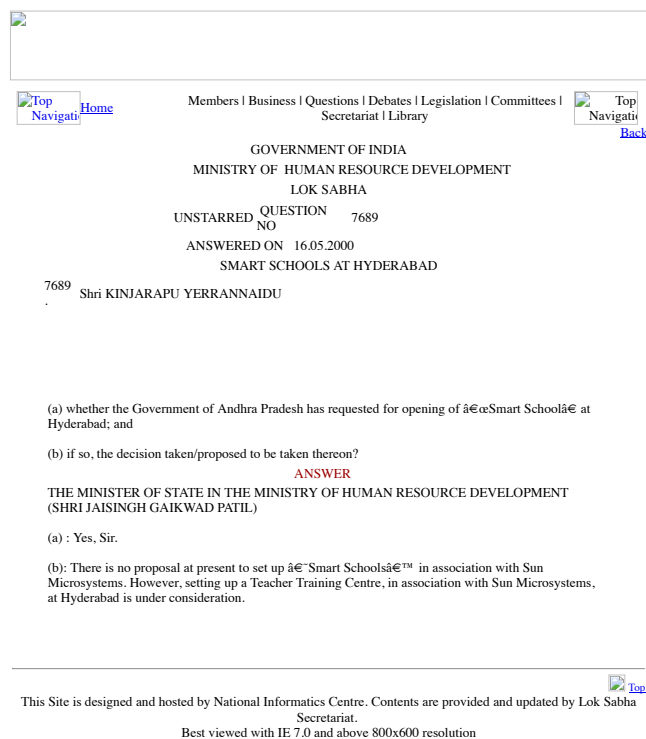
Figure A.2: Example of a tweet returned by API

```
{
  "contributors": null,
  "truncated": true,
  "text": "I've gotten my flu vaccine, have you? If you are elderly or in a high risk category it is vital that you do.\u2026 https://t.co/WFuAQzqWue",
  "is_quote_status": false,
  "in_reply_to_status_id": null,
  "id": 819900279617372160,
  "favorite_count": 39,
  "source": "<a href='\"http://twitter.com/download/iphone\"' rel='\"nofollow\"'>Twitter for iPhone</a>",
  "retweeted": false,
  "coordinates": null,
  "entities": {
    "symbols": [],
    "user_mentions": [],
    "hashtags": [],
    "urls": [
      {
        "url": "https://t.co/WFuAQzqWue",
        "indices": [108, 131],
        "expanded_url": "https://twitter.com/i/web/status/819900279617372160",
        "display_url": "twitter.com/i/web/status/8\u2026"}
    ],
    "in_reply_to_screen_name": null,
    "in_reply_to_user_id": null,
    "retweet_count": 19,
    "id_str": "819900279617372160",
    "favorited": false,
    "user": {
      "follow_request_sent": false,
      "has_extended_profile": false,
      "profile_use_background_image": true,
      "default_profile_image": false,
      "id": 38000023,
      "profile_background_image_url_https": "https://pbs.twimg.com/profile_background_images/837258032/f7d6261e31e090db0d00ddc29645bc0c.png",
      "verified": false,
      "translator_type": "none",
      "profile_text_color": "3C3940",
      "profile_image_url_https": "https://pbs.twimg.com/profile_images/578322413814824960/TH-ElloJ_normal.png",
      "profile_sidebar_fill_color": "95E8EC",
      "entities": {
        "description": {
          "urls": []
        }
      },
      "followers_count": 5898,
      "profile_sidebar_border_color": "FFFFFF",
      "id_str": "38000023",
      "profile_background_color": "0099B9",
      "listed_count": 93,
      "is_translation_enabled": false,
      "utc_offset": null,
      "statuses_count": 1632,
      "description": "",
      "friends_count": 1272,
      "location": "",
      "profile_link_color": "0099B9",
      "profile_image_url": "http://pbs.twimg.com/profile_images/578322413814824960/TH-ElloJ_normal.png",
      "following": false,
      "geo_enabled": true,
      "profile_banner_url": "https://pbs.twimg.com/profile_banners/38000023/1367422408",
      "profile_background_image_url": "http://pbs.twimg.com/profile_background_images/837258032/f7d6261e31e090db0d00ddc29645bc0c.png",
      "screen_name": "HMcEntee",
      "lang": "en",
      "profile_background_tile": true,
      "favourites_count": 1337,
      "name": "Helen McEntee TD",
      "notifications": false,
      "url": null,
      "created_at": "Tue May 05 19:38:37 +0000 2009",
      "contributors_enabled": false,
      "time_zone": null,
      "protected": false,
      "default_profile": false,
      "is_translator": false
    },
    "geo": null,
    "in_reply_to_user_id_str": null,
    "possibly_sensitive": false,
    "lang": "en",
    "created_at": "Fri Jan 13 13:33:54 +0000 2017",
    "in_reply_to_status_id_str": null,
    "place": {
      "full_name": "Dublin City, Ireland",
      "url": "https://api.twitter.com/1.1/geo/id/7dde0febc9ef245b.json",
      "country": "Ireland",
      "place_type": "city",
      "bounding_box": {
        "type": "Polygon",
        "coordinates": [
          [
            [-6.387438, 53.2987449],
            [-6.1078047, 53.2987449],
            [-6.1078047, 53.4110598],
            [-6.387438, 53.4110598]
          ]
        ]
      },
      "contained_within": [],
      "country_code": "IE",
      "attributes": {}
    },
    "id": "7dde0febc9ef245b",
    "name": "Dublin City"
  }
}
```

India parliamentary question data

To obtain all the parliamentary questions submitted by Indian members of parliament between 1999 and 2009, I wrote a custom function to scrape the webpages containing the questions from the website of the Lok Sabha. Like the Ireland legislative text data discussed above, the format of these webpages is HTML. I therefore wrote a further series of Python functions to extract the text and metadata of the questions from this format.

Figure A.3: Example of parliamentary question HTML page



Bag of words representation

For the statistical analysis conducted in Chapter 2, Chapter 3, Chapter 4, and Chapter 5, I employed a bag of words representation of documents from the relevant text sources. A bag of words representation ignores the order of words and assumes words are independent by simply considering the frequency of a word in a document (Jurafsky and Martin, 2017). For each corpora, I construct a document term matrix where each document in the corpus is represented by a row and each word in the corpus is represented by a column.

A.1.2 Details of Fightin’ Words model

In Chapter 2, I employed the Fightin’ Words model to compare the content of legislators’ tweets, speeches, and questions. In this section, I formally outline the details of this model.

The log odds ratio informative Dirichlet prior model developed by Monroe, Colaresi and Quinn (2009), informally known as the Fightin’ Words algorithm, is a Bayesian statistical model for identifying words that distinguish a category of texts from another. This model has been utilized to study partisan differences in Congressional speeches, to understand the content of human rights reports, to compare the language of restaurant menus, and to identify differences in NFL scouts’ reports of white and minority quarterback prospects (Monroe, Colaresi and Quinn, 2009; Jurafsky et al., 2014; Fariss et al., 2015; Boylan, McMahon and Monroe, 2016). This model requires that differences in the usage of any particular word across different categories of text must be very strong before we can be confident that we are identifying a systematic difference in language usage.

Formally, starting without considering the different categories of text, the model assumes the vector of frequencies, $\mathbf{y}$, for the W unique words in the combined corpus,

follows a multinomial distribution:

$$\mathbf{y} \sim MN(n, \boldsymbol{\pi}) \quad (\text{A.1})$$

where n represents the total number of words in the corpus, and $\boldsymbol{\pi}$ represents a vector capturing the probabilities for the W distinctive words in the corpus. The model next assumes, that $\boldsymbol{\pi}$ follows a Dirichlet distribution:

$$\boldsymbol{\pi} \sim \text{Dirichlet}(\boldsymbol{\alpha}) \quad (\text{A.2})$$

where $\boldsymbol{\alpha}$ is a prior vector containing information about the relative usage frequency of the unique words W with all $\alpha_w > 0$. As the Dirichlet is the conjugate prior of the multinomial distribution, there is an analytic solution for $\boldsymbol{\pi}$:

$$\hat{\boldsymbol{\pi}} = \frac{\mathbf{y} + \boldsymbol{\alpha}}{n + a_0} \quad (\text{A.3})$$

To calculate the odds of the word w relative to all other $W - 1$ words as:

$$\Omega_w = \frac{\pi_w}{1 - \pi_w} \quad (\text{A.4})$$

Moving to the case the when we want to compare two categories of text i and j , we can utilize the log odds ratio, δ_w^{i-j} to understand differences in the usage of word w across the categories:

$$\delta_w^{i-j} = \log \left(\frac{y_w^i + \alpha_w}{n^i + \alpha_0 - y_w^i - \alpha_w} \right) - \log \left(\frac{y_w^j + \alpha_w}{n^j + \alpha_0 - y_w^j - \alpha_w} \right) \quad (\text{A.5})$$

The variance of this estimate can be calculated as:

$$\sigma^2(\delta_w^{i-j}) \approx \frac{1}{y_w^i + \alpha_w} + \frac{1}{y_w^j + \alpha_w} \quad (\text{A.6})$$

To account for the fact that words employed infrequently will have greater variance

than more frequently used terms, we use the z-score, ζ_w^{i-j} , of the log odds ratio as the statistic for evaluating linguistic differences in categories of text:

$$\zeta_w^{i-j} = \frac{\delta_w^{i-j}}{\sqrt{\sigma^2(\delta_w^{i-j})}} \quad (\text{A.7})$$

As shown by Monroe, Colaresi and Quinn (2009), this model based approach for comparing groups of text has advantages over comparing term frequencies, normalized frequencies, or TF-IDF weightings, which fail to account for variance in word usage and do not utilize information about the expected distribution of words.

A.1.3 Details of LASSO model

The least absolute shrinkage and selection operator (LASSO) model is a widely used statistical model which modifies the standard linear regression and generalized linear models (Hastie, Tibshirani and Friedman, 2009). The LASSO uses regularization to enhance the predictive accuracy of these models and identify the most influential predictors of a given outcome by shrinking the coefficients of certain predictors to exactly zero.

More formally, lets assume, as in chapters 3 and 5, that we are interested in predicting a binary outcome, $Y = 0, 1$, with a series of predictors represented by X . We can model probability that $Y = 1$ given the value of the predictors as:

$$Pr(Y = 1|X = x) = \frac{e^{\beta_0 + \beta^T x}}{1 + e^{\beta_0 + \beta^T x}} \quad (\text{A.8})$$

which can be equivalently written as:

$$\log\left(\frac{Pr(Y = 1|X = x)}{Pr(Y = 0|X = x)}\right) = \beta_0 + \beta^T x \quad (\text{A.9})$$

We then use standard maximum likelihood methods to find the parameters β_0, β

that maximize the loglikelihood function:

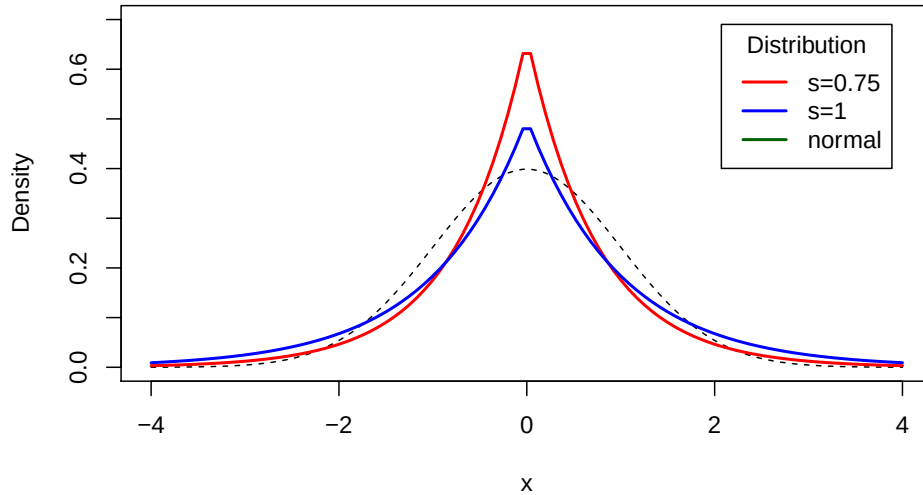
$$\max_{\beta_0, \beta} \left[\frac{1}{N} \sum y_i (\beta_0 + \beta^T x) - \log(1 + e^{\beta_0 + \beta^T x}) \right] \quad (\text{A.10})$$

The above demonstrates how standard logistic regression is estimated. The LASSO model modifies the procedure slightly by minimizing a loss function composed of the negative loglikelihood function and a ℓ_1 regularization penalty:

$$\min_{\beta_0, \beta} \left[\frac{1}{N} \sum y_i (\beta_0 + \beta^T x) - \log(1 + e^{\beta_0 + \beta^T x}) \right] + \lambda \|\beta\|_1 \quad (\text{A.11})$$

where λ controls the level of regularization. A coordinate descent algorithm is often used to minimize the above loss function (Hastie, Tibshirani and Friedman, 2009). Perhaps the most intuitive way to understand how the LASSO operates is to use a Bayesian interpretation where the priors for the coefficients have a double exponential, or Laplace, distribution with a mean of zero (James et al., 2013).

Figure A.4: Laplace Distribution



In A.4, I present the Laplace distribution with different values for the scale parameter and a standard normal distribution for comparison. From this we can see that a greater proportion of the probability mass of the Laplace distribution is centered around zero, which helps to clarify why a number of the coefficients for the LASSO model are estimated as exactly zero.

A.2 Chapter 3

Table A.1: Example questions classified as constituency service

Example	Text of question
1	asked the minister for social welfare the reason for the delay in paying disability benefit to a person details supplied in glanworth county cork and asked the minister for social welfare the reason for the delay in paying disability benefit to a person details supplied in county cork and asked the minister for social welfare the reason for the delay in paying deserted wives benefit to a person details supplied in cobh county cork and asked the minister for social welfare the reason for the delay in paying deserted wives benefit to a person details supplied in cobh county cork
2	asked the minister for foreign affairs if a person details supplied is the current holder of a valid passport if two children have been added to that passport and if not if he will have the case investigated as a matter of urgency
3	asked the minister for education if she will make a statement on the proposal to make it mandatory for vec scholarship applicants to have attained two honours in the leaving certificate asked the minister for education the percentage of students holding vec scholarships who entered athlone regional technical college county westmeath in the autumn of 1982 who would not comply with a two honours requirement
4	asked the minister for education if tenders have been received in relation to an extension to a national school details supplied in county offaly
5	asked the minister for health and children if a person details supplied in county westmeath will be admitted to hospital for a hip replacement operation and if he will make a statement on the matter

Table A.2: Example questions classified as not constituency service

Example	Text of question
1	asked the minister for social welfare the names and functions of the commissions committees and state-sponsored bodies abolished by him during his term of office to date and the names and functions of the commissions committees and state-sponsored bodies set up by him or proposed to be set up by him since taking office
2	asked the minister for defence his plans to purchase new troop and equipment-carrying aircraft the aircraft available for such use by the defence forces and his views on whether the position is satisfactory
3	asked the minister for agriculture and food if she is satisfied that sufficient steps are being taken at eu level to monitor the importation of beef from brazil if her attention has been drawn to the fact that beef is being imported into brazil from other countries with no tagging and no traceability if she is further satisfied that the industry here and especially the consumers here are fully safeguarded and if she will make a statement on the matter
4	asked the minister for transport when he transposed directive 2003/20/ec into irish law the categories of vehicles for which ireland was granted exemptions from the directive the reason an exemption was sought in each case when these derogations were applied for by ireland when he expects these categories to be included in law here if a publicity or information campaign has been organised by his department to inform passengers and transport providers of the new requirements and if he will make a statement on the matter
5	asked the minister for education and science the plans in place to assist pupils who have been determined disadvantaged after they leave primary school the linkage there is between primary and secondary schools to assist such children and if he will make a statement on the matter

I present the coefficients from the different negative binomial models examining the empirical relationship between dynastic membership and service responsiveness below:

Table A.3: Coefficients from negative binomial models

Dynastic	0.103** (0.047)	0.136** (0.065)	0.126** (0.063)
Dublin	-0.226*** (0.042)	-0.325*** (0.057)	-0.342*** (0.065)
Government	0.147*** (0.035)	0.174*** (0.021)	0.176*** (0.021)
Vote Share	0.011*** (0.003)	-0.001 (0.002)	-0.002 (0.002)
Female	-0.135** (0.061)	-0.101 (0.074)	-0.090 (0.073)
Fianna Fáil	0.236*** (0.034)	0.212*** (0.044)	0.203*** (0.043)
Dynastic×Dublin	-0.234*** (0.087)	-0.195* (0.110)	-0.167 (0.108)
Constant	-1.056*** (0.053)	-0.936*** (0.105)	-0.899*** (0.108)
TD random effects		✓	✓
Dáil Éireann random effects		✓	✓
Constituency random effects			✓
Observations	1,332	1,332	1,332
Log Likelihood	-5,484.826	-5,159.856	-5,156.420
θ	3.728*** (0.173)		
Akaike Inf. Crit.	10,985.650	10,341.710	10,336.840
Bayesian Inf. Crit.		10,398.850	10,399.170

Note:

*p<0.1; **p<0.05; ***p<0.01

A.2.1 Details of PR-STV

Under STV, voters can indicate preferences up to the number of candidates that are listed on a ballot. Seats are allocated to candidates that reach a quota. A quota can be defined as:

$$1 + \frac{N}{s + 1} \tag{A.12}$$

where N is the total number of votes cast and s is the number of seats in a constituency. If all seats have not been filled after the counting of first preference votes then another count takes place. This is done by either distributing the surplus votes of an elected candidate that has reached the quota or by eliminating the candidates with the fewest votes and distributing these votes to the remaining candidates according to the preferences on the ballot (Marsh, 2007). This process continues until all seats have been filled.

A.3 Chapter 4

Table A.4: Example questions related to ethnic minority groups

Example	Text of question
1	Whether 'backlog/carried forward vacancies' reserved for SCs, STs and OBCs are required to be treated as a 'separate and distinct Group' as provided under Article 16(4) B of the Constitution to overcome 50% ceiling limit on vacancies to be reserved in an year; if so, indicate 'backlog/carried forward vacancies' ascertained as per para 5 of the DOPT OM No. 36012/2/96-Estt.(Res.) dated 2nd July 1997 in respect of (1) SCs,(2) STs and OBCs in Group A, B, C and D categories of services as on 29th August 1997 in the Ministry of Parliamentary Affairs when Special Recruitment Drives etc. meant for filling such vacancies were stopped; indicate year-wise such carried forward vacancies filled during the last four years and those which remain unfilled; and the fresh vacancies/posts accrued to reserved classes in all categories of posts during the last four years as per post based rosters?
2	(a) whether the Government are aware that some schemes for the welfare of those belonging to Scheduled Castes and Scheduled Tribes are not being implemented effectively; (b) if so, the details thereof alongwith reasons therefor; (c) whether any complaints have been received by the National Commission for Scheduled Castes and Scheduled Tribes in this regard; and (d) the steps being taken by the Government to ensure speedy and effective implementation of the schemes?
3	(a) the number of posts of teachers lying vacant in the Kendriya Vidyalayas of Bihar and the number of posts reserved for the teachers belonging to the Scheduled Castes and Scheduled Tribes out of these; (b) the reasons for delay in filling up of these vacancies; and (c) the action proposed to be taken to fill up these vacancies?
4	(a) whether the Government pay special attention to provide loan facility to the poor and dalits; (b) if so, the details thereof; and (c) the number of the poor and dalits in Rajasthan and Maharashtra provided loans by nationalised banks during the last three years indicating the amount of loans provided?
5	(a) the total amount of financial assistance provided by the National Scheduled Castes finance & Development Corporation (NSCFDC), under its various schemes, during the last three years, scheme-wise; (b) the cumulative recovery status of NSCFDC dues during this period scheme-wise; and (c) the steps taken by the Government to help improve recoveries of NSCFDC dues for regular schemes as well as for micro-credit finance?

Table A.5: Example questions not related to ethnic minority groups

Example	Text of question
1	(a) whether the Railways have taken note of the recent reduction in airfares; (b) if so, whether the Railways apprehend diminishing patronage for its prestigious Rajdhani trains the fare on which for First Class AC works out to be costlier than airfares; (c) if so, whether the Railways propose to fall in line with the airlines in so far as the viability of continuing Rajdhani services; and (d) if so, the details thereof?
2	(a) whether there is any proposal under the consideration of the Government to provide Water Convergence System in the hilly and high rainfall areas; (b) if so, the details thereof; and (c) if not, the reasons therefor?
3	(a) whether the Government are aware of some cases of poaching by foreigners who forayed the rich marine wealth in Andaman and Nicobar Islands in the recent past; (b) if so, the details thereof; and (c) the action taken against the guilty persons?
4	(a) the amount spent by the Government on research work in agriculture sector during the last three years; (b) the details of the research work undertaken and the achievements made in this regard; and (c) the steps being taken to make further improvement in this regard?
5	(a) whether the Annual Status of Education Report (ASER) released by Planning Commission, the quality of elementary education is far behind the satisfactory level in rural India as reported in Times of India dated January 19, 2006; (b) if so, the reasons therefor alongwith the reaction of the Government thereon; (c) the steps taken by the Government to improve the quality of elementary education in rural India; (d) whether the Government has recently conducted survey of the Sarva Shiksha schools and found students facing several problems in these schools as reported in the Times of India dated January 26, 2006; (e) if so, the facts and details thereof; (f) whether the Government has decided to monitor the problems of Sarva Shiksha schools itself; and (g) if so, the details thereof?

I present the coefficients from the different negative binomial models examining the empirical relationship between ethnic quotas and legislators' responsiveness to the interests of minority ethnic groups below:

Table A.6: Coefficients from negative binomial models

Ethnic Quota	1.645*** (0.096)	1.555*** (0.101)	1.552*** (0.098)
Female	-0.473*** (0.171)	-0.447** (0.177)	-0.283* (0.168)
Government	-0.259*** (0.095)	-0.290*** (0.101)	-0.241** (0.114)
Constant	-4.202*** (0.059)	-4.199*** (0.139)	-4.142*** (0.162)
Lok Sabha random effects		✓	✓
State random effects		✓	✓
Party random effects			✓
Observations	963	963	963
Log Likelihood	-1,904.908	-1,890.856	-1,864.338
θ	0.971*** (0.069)		
Akaike Inf. Crit.	3,817.815	3,795.713	3,744.676
Bayesian Inf. Crit.		3,829.803	3,783.636
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01			

A.3.1 Details on the assignment of ethnic quotas to districts in India

In the Constitution of India it is specified that subsequent to every census that a Delimitation Commission needs to be formed to decide on the boundaries of constituencies and the assignment of ethnic quotas to constituencies (Jensenius, 2013; Election Commission of India, 2016). Since independence Delimitation Commissions have been established in 1952, 1963, 1973, and 2002 (Jensenius, 2013). States in India are assigned seats to the Lok Sabha relative to their populations. In a similar vein, the number of ethnic quotas for Scheduled Castes and Tribes in states is proportional to the relative population of these groups (Election Commission of India, 2017). For Scheduled Tribes, ethnic quotas are assigned to the constituencies in states with the highest proportion of members from this ethnic group. For Scheduled Castes, ethnic quotas are also assigned to the constituencies in states with the highest proportion of members from this ethnic group but with the added requirement that the quotas also need to be geographically dispersed across the state (Election Commission of India, 2017). This means that constituencies with higher levels of Scheduled Caste members are not necessarily assigned an ethnic quota.

A.4 Chapter 5

Table A.7: Topic labels for Twitter topic model

Topic	Topic Label	Top Words
Topic 1	Miscellaneous	irish, people, work, also, times
Topic 2	Media Appearances	tonight, tomorrow, looking, debate, forward
Topic 3	Schools	see, school, election, excellent, schools
Topic 4	Allocations	funding, services, scheme, programme, youth
Topic 5	Sinn Féin	government, labour, td, party, sinn
Topic 6	Referenda	vote, yes, campaign, way, marref
Topic 7	Business/Trade	ireland, national, conference, opening, interesting
Topic 8	Left Opposition	welcome, taoiseach, galway, road, area
Topic 9	Taxes	public, tax, govt, business, pay
Topic 10	Constituency Clinics	morning, local, community, centre, 10
Topic 11	Jobs/Investment	new, jobs, news, plan, hospital
Topic 12	Legislative Activities	dail, bill, committee, speaking, report
Topic 13	Irish Language	via, go, says, na, ar
Topic 14	Deaths/Local	one, still, visit, hear, man
Topic 15	Launches/Activities	delighted, launch, evening, house, meet
Topic 16	Political Parties	thanks, fg, many, come, michael
Topic 17	Dublin Protests	dublin, protest, north, 30, south
Topic 18	Dates/Deadlines	re, 2015, education, future, 2014
Topic 19	Health	need, health, min, question, statement
Topic 20	Photos	minister, week, us, must, like
Topic 21	Well Wishes	today, good, best, luck, congrats
Topic 22	Water Charges	last, water, year, night, years
Topic 23	Congratulations	great, well, done, happy, really
Topic 24	Spending Cuts	support, budget, take, cuts, families
Topic 25	Saint Patrick's Day	day, back, st, first, team
Topic 26	Meetings	now, meeting, just, eu, council
Topic 27	Outdoor Activities	children, family, make, know, made
Topic 28	Services/Grants	home, waterford, free, workers, care
Topic 29	Twitter Language	amp, time, can, get, please
Topic 30	Local Events	will, cork, next, todays, reform

Christopher Boylan

Pennsylvania State University
christopheriboylan@gmail.com
cboylan.com
github.com/cib117

Education	Pennsylvania State University	
	Ph.D Political Science	2017
	M.A.S. Applied Statistics	<i>Expected</i> 2018
	Trinity College Dublin, Ireland	
	M.Sc. International Politics	2011
Relevant Experience	University College Cork, Ireland	
	B.A. English and History	2010
	Data Scientist	July 2017–
	Verisk Maplecroft	
	Analytic Science Intern	Summer 2016
Peer-Reviewed Articles	Verisk Innovative Analytics	
	Graduate Research Assistant	Fall 2012 - Spring 2017
	Pennsylvania State University	
	Christopher Boylan, Ryan McMahon, and Burt L. Monroe. 2016. “Racially Differentiated Language in NFL Scouting Reports”. <i>Workshop on Large-Scale Sports Analytics at ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '16)</i> .	
Skills	Text Mining	Machine Learning
	Predictive Modeling	Design of Experiments
	Python · R · Stata · Stan · SQL · HTML · CSS	Data Visualization
		Causal Inference
Awards and Fellowships	PiCP Fellowship, Pennsylvania State University	2016
	Steighton A. Watts Graduate Scholarship, Pennsylvania State University	2016
	NSF Travel Grant for Meeting of the Society for Political Methodology	2016
	QuaSSI Predoctoral Research Fellowship, Pennsylvania State University	2015
	Big Data Social Science IGERT Affiliate, Pennsylvania State University	2015
	Conference Travel Award, Pennsylvania State University	2015
	Summer Graduate Research Award, Pennsylvania State University	2015
	Graduate Scholar Award, Pennsylvania State University	2012
Personal Information	Robert W. Graham Endowed Fellowship, Pennsylvania State University	2012
	Nationality: Irish	