Title
The Multiple Nontransferable Vote in Theory and Practice: Dynamic Political Behavior in New Hampshire and Vancouver

Author
Gordon, Oakley Benedict

Publication Date
2022

Peer reviewed|Thesis/dissertation

The Multiple Nontransferable Vote in Theory and Practice: Dynamic Political Behavior in New Hampshire and Vancouver

By

OAKLEY BENEDICT GORDON

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Political Science

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Matthew S. Shugart, Chair

James F. Adams

Josephine Andrews

Committee in Charge

2022

For my son, Max, and his baby brother or sister who is on the way

Contents

List of Figures

List of Tables

Abstract

The Multiple Nontransferable Vote in Theory and Practice: Dynamic Political Behavior in New Hampshire and Vancouver

A great deal of academic attention has been paid to both the majoritarian elections of single winners and the proportional election of multiple winners. Research has included the study of the connection of these systems to issues varying from voter behavior, party formation and stability, equity, social justice, ethnic tensions, and polarization. Nonetheless, the intersection of these systems—the majoritarian election of multiple winners—has escaped any considerable academic attention, at both the theoretical and empirical levels. This dissertation makes a major step forward into this research void by examining the most simple and common form of majoritarian multiple-winner election: the multiple nontransferable vote (MNTV), in which M candidates are elected, each voter can vote for up to M candidates, and the M candidates with the most votes win.

Through three papers, compiled herein as three substantive chapters along with an introduction and a conclusion, this dissertation builds and then tests three theories related to the ability of political actors to react and behave strategically in political systems that employ MNTV. Aside from the focus on the same electoral system, the common thread throughout each chapter are the questions: Does MNTV present unique strategic considerations, distinct from those of single-winner majoritarian and proportional systems? Do candidates, parties, and voters understand these incentives well enough to act strategically? Finally, do these individual strategic considerations result in any unique emergent properties in the political systems that employ MNTV?

To answer these broad questions, each substantive chapter poses a more specific question regarding strategic behavior in MNTV systems. Chapter 2 finds that voters

will vote more strategically when their preferred-party's control of the legislature is threatened in New Hampshire's districted MNTV system. Chapter 3 finds that candidates and parties are more likely to enter Vancouver's MNTV system when the balance of power between the parties presents a strategic opportunity for them to do so. Chapter 4 explores the limits of actors' ability to strategically act in MNTV systems by discovering a period of system shock defined by erratic behavior following a modification to Vancouver's electoral system in 1968. Together, these chapters present a set of strategic incentives that are unique to systems that use MNTV to elect their leaders and then show that actors have a surprising (albeit limited) capability to respond to these incentives. These findings advance the study of electoral systems in a new direction and have normative implications regarding the reform and equity of MNTV systems.

Acknowledgments

I am incredibly grateful for and to the friends, colleagues, and family members who have supported me on the long road to the completion of this dissertation. First and foremost, I am thankful for my amazing wife, Jess, who supported and loved me while writing her own dissertation. I am excited to see what is in store next for the Doctors Gordon. To my son, Max, thank you for the joy and purpose you bring to my life everyday. And, for course, I am thankful for my family, parents, and parents in-law; I could not have completed this PhD during the pandemic without your help.

To my PhD cohort—Isaac, Lowell, Sam, Dan, Nahrain, Tzu-Ping, Maria, Ross, Brian, Christy, and Kendra—I could not have asked for a better group of friends with whom to start my graduate career at UC Davis. I miss our our classes together, our shared offices, our game nights, and especially our weekly dinners. When we were all together, I felt at home in a way I had rarely felt before. Some of you, I will see on our continuing adventures together. To the rest, though our paths may never cross again, I will always cherish our time together.

I am deeply grateful for my exceptionally kind and patient dissertation committee: My chair and advisor, Matthew Shugart, whose guidance, kindness, and insights have gone above and beyond that of any mentor; Jim Adams, who is perhaps the nicest person I've ever met; Jo Andrews, who has been a delight to work for as both her TA and research assistant; and Norm Matloff, who very generously joined the committee in an honorary capacity on short notice to provide critique and feedback regarding the use of machine learning in my fourth chapter. To the greater UC Davis community, from the bright students whom I taught to the planners who made our campus a beautiful bike haven, thank you for giving me an institution to be proud of; GO AGGIES! Rounding out my academic gratitude, I am grateful to Professors John Francis, Baodong Liu, and James Curry for helping me to find my academic footing and direction while I was in Utah.

The friends who have supported and cared for me throughout the years are more numerous than I deserve and far more numerous than I can list here comprehensively. In Davis, I am especially thankful for my friend Evan Rothstein, who is always a delight, and for Katy Jarvis, who was always there for me and my family. From Salt Lake City, I am thankful for my best man and hiking partner Jon Hansen, along with Mike Robant, Ian Hill, James Carroll, and everyone who trudged up a mountain with me over the years. To you and to all of my other friends, past and present, thank you for your friendship and support.

Chapter 1

Introduction

Few electoral systems raise the ire of political scientists and reformers as much as the multiple nontransferable vote (MNTV). In an MNTV election of M candidates, where $M > 1$, each voter can cast M votes but not more than one vote per candidate; the M candidates with the most votes are elected. In both theory and practice, this generally results in an at-large plurality winning most or all of the seats, with no mechanism for producing representation for smaller voting blocs, which could be represented by a proportional electoral system (or by single member districts, if the bloc is geographically concentrated). Whether adopted out of an inability to find a better system or as a deliberate ploy to suppress ethnic and partisan minorities, MNTV produces elections that are generally uncompetitively elected and results that are generally unrepresentative. Regardless of its normative shortcomings and simple rules, however, the behavior of actors within the system is far from simple. Predicting and understanding how politicians, parties, and voters behave and adapt in MNTV systems is important, not only because it informs efforts to reform the system where it is used, but also because it helps build a more generalized understanding of behavior in all majoritarian systems.

Despite the wide condemnation of MNTV electoral systems and the importance of understanding them for the purpose of reform, the effects of the system—beyond its production of non-proportional results—are understudied yet surprisingly complex and dynamic. This can, perhaps, be attributed to the fact that MNTV is generally employed at subnational and local levels, where it has escaped the scrutiny of researchers focused on national elections and where it less likely to produce data that is both plentiful enough and accessible enough for analysis. This dissertation begins to remedy this lack of academic examination on the complex effects of MNTV elections on voters, parties, and candidates.

The three substantive chapters of this dissertation explore three different topics related to the behavior of actors within MNTV systems. Chapter 2 examines how voters limit their split-ticket voting (i.e. voting for candidates from multiple parties in the same MNTV election) in the elections of New Hampshire's lower state house. Chapters 3 and 4 both use data from municipal election in Vancouver, BC. Chapter 3 creates and test a theory of strategic entry for both candidates and parties, finding that political actors are able to determine and exploit strategically ideal moments to enter the political system. Chapter 4 employs a novel artificial neural network to identify a period of uncertainty produced by a change of magnitude in the city's electoral system.

1.1 Dissertation Outline

1.1.1 Split-Ticket Voting in New Hampshire

Chapter 2 begins this dissertation's empirical examination of behavior in MNTV systems with a look at voter behavior in New Hampshire. New Hampshire uses MNTV to elect the 400-member lower house of its state legislature. The state also provides an opportunity to study the effects of MNTV when the party system of the

MNTV election is heavily embedded in the party system of contemporaneous higher offices (i.e. the MNTV election of the New Hampshire State House of Representatives is always down-ballot of single-winner elections ranging from the state senate to the US presidency). In short, MNTV affects the strategies of parties, politicians, and voters, but does not have a great effect upon the number or size of parties, thereby allowing said strategic effects to be isolated and studied without complication from a shifting party system.

The general question posed by chapter 2 is how well voters in an MNTV are able to strategically limit their split-ticket voting in response to partisan considerations arising from the magnitude and competitiveness of their districts. The chapter builds a theory that split-ticket voting is especially detrimental to a voter's preferred party when the magnitude and competitiveness of their district are both high. The underlying logic of this theory, as demonstrated mathematically in the chapter, is that even a small number of split-ticket votes can "flip" the entirety of one district's delegation from one party to another when a district is especially competitive. Further, such a flip at the district level is more likely to affect the partisan majority of the chamber as a whole when the district magnitude is high, relative to other districts. Taken together, this means that voters with any preference regarding which party should control the chamber will be more likely to ignore their cross-party candidate preferences and instead vote straight-ticket if they are voting in a competitive high-magnitude district.

In addition to the large number of members of the chamber, the election of New Hampshire House of Representatives is somewhat unique in that it uses variable-magnitude districts, rather than uniform-magnitude districts or an at-large election. This provides a unique variable that is required to leverage the chapter's main substantive question. While New Hampshire provides a rare opportunity to study

MNTV with both variable magnitude and variable competitiveness between districts, the chapter still faces the hurdle of measuring the level of split-ticket voting within each district. This is difficult at face value, because split-ticket voting cannot be directly inferred from aggregate election results and individual ballots are difficult to obtain. To solve this problem, chapter 2 proposes a new method for approximating the percentage of ballots that were cast as split-tickets based upon election totals, then it uses this metric as its main dependent variable.

The results of the analysis show that voters are indeed more likely to vote straight-ticket in competitive high-magnitude districts, despite the facts that one might both expect (correctly or not) voters in competitive districts to be less partisan and, therefore, more inclined to vote for candidates of multiple parties and expect voters in higher-magnitude districts, who have more votes, to be more likely to spare a vote or two for a candidate outside of their preferred party. Nonetheless, these countervailing expectations are not completely unfounded. The analysis only upholds the chapter's main theory in an interactive model, in which the effects of magnitude and competitiveness are measured through a multiplicative term. This suggests that competitiveness and magnitude both, on their own, increase split-ticket voting and only decrease split-ticket voting when the two variables increase together.

1.1.2 Strategic Party Formation and Candidate Entry in Vancouver, BC

Chapters three and four round out the dissertation with an examination of party and candidate behavior in city politics Vancouver, British Columbia, drawing upon a dataset compiled for this chapter containing every candidacy from 1936 to 2018. Despite both systems using MNTV elections, the political systems of New Hampshire's state legislative elections and Vancouver's municipal elections are extremely different. New Hampshire's party system is a local extension of the national party system; Vancouver's parties are endemic to the city. New Hampshire's party system is stable; Vancouver's parties rise, merge, split, and fall chaotically from election to election. Thus, while New Hampshire provides an opportunity to study MNTV elections that are wholly embedded in and stabilized by a larger overarching political system, Vancouver provides an opportunity to study MNTV elections that exist in a relative vacuum, off-cycle from provincial and national elections, and stabilized solely by the city's mayoral race. Before delving into the specific research questions and methods regarding Vancouver's elections, the first and most obvious takeaway from a comparison of New Hampshire and Vancouver is that MNTV produces a fundamentally unstable political system if left without stabilization by external forces.

Chapter 3 examines strategic entry of parties and candidates within the chaotic party landscape produced by the city's at-large MNTV elections of its city council, school board, and park board, along with the plurality election of its mayor. In a more stable majoritarian system, the calculus of entry is rather simple and static. In the ever-changing realm of Vancouver's party system, however, in which major parties can both rise and collapse relatively suddenly (and in which intrepid independents may triumph unexpectedly), the calculus of when to enter a race as a new candidate (or as an experienced candidate in a new party) can mean the difference between success and failure.

Chapter 3 suggests a relatively simple, yet mathematically derived, theory that suggests that certain conditions in a given election provide the greatest chance of success for new small parties seeking to win office by drawing support away from the city's largest party. The question that follows is whether this theory of opportunity can explain observed candidate and party entry in different elections. If so, the results would suggest a higher degree of strategic sophistication among potential candidates and party formateurs.

To answer this question, chapter 3 employs a model to explain the absolute and relative number of new candidates and members of new parties. The chapter also introduces a new variety of the "effective number" metric, which it uses as an independent variable to help measure party fragmentation. Effective number metrics are used to provide a count of the number of parties within and, by extension, the fragmentation of a party system, because the number of registered parties will generally overstate the number of relative parties for all practical purposes. The effective number is generally calculated by using either the parties' shares of the total vote or their shares of the elected seats. Chapter 3 introduces the use of the parties' shares of the nominated candidates to produce a third metric of fragmentation. This metric is of unique use in Vancouver because, unlike many places, parties are much more likely to nominate partial slates rather than full slates The effective number of parties based on nominees is of practical use for chapter 3 because the theory built by the chapter assumes a high degree of variability in the number of candidates nominated by each party.

Chapter 3 ultimately finds that candidate entry and party formation can indeed be predicted by the strategic opportunity presented by the party system after any given election. As the largest party grows larger in terms of vote share, new parties and new candidates will enter the system in an attempt to exploit the fact that they need to "poach" diminishing shares of largest party's voter base. Together with chapter 2, this suggests that both voters and candidates are able to navigate the unique and dynamic strategic incentives presented to them by MNTV.

1.1.3 Electoral Reform and System Shock in Vancouver, BC

While chapters two and three explore how strategically candidates and voters behave in MNTV election, chapter 4 asks how well political actors respond to changes within an MNTV system. If actors are generally strategic in a stable system, then how long does it take for them to adapt to change? Specifically, chapter 3 examines the effects of a change in magnitude and the end of staggered terms in Vancouver's municipal councils on the behavior of the city's politicians, with the goal of determining the nature and duration of said effects.

The nature of chapter 4's question requires an ability to find especially nuanced and subtle patterns in the aggregate and individual behavior of candidates throughout Vancouver's history. To do so, chapter 4 employs a novel artificial deep neural network, which is trained to predict the behavior of any given candidate based upon their own political history, the history of their party, and the political history of the city. Once trained, this neural network reveals patterns between candidates, between parties, and between years

The patterns revealed by this network show a definitive period of systemic instability arising following electoral reform in 1968 and ending 11 elections later in 1990. This period of instability was defined by erratic candidate behavior and was seemingly driven by greater uncertainty by all actors within the system. Further, the findings suggest that the period of instability was neither uniform nor instantaneous. Instead, the system transitioned or "spiraled," so to speak, into instability over the eight years following the reform and remained in a state of peak instability for fourteen years thereafter before stabilizing into its current equilibrium.

1.1.4 Key Conclusions

Together, these chapters show that politicians and voters alike can respond strategically to the surprisingly complex incentives that are generated by the simple rules of MNTV elections, with caveat that political actors seem to have trouble adapting to changes in MNTV systems. This dissertation presents itself as one of a very few attempts of academia to understand the dynamic political systems under the MNTV. It is, however, hopefully among the first steps of many in the field to examining and better modeling behavior in MNTV and other less studied systems used around the world.

But how does this dissertation's empirical examination of MNTV elections inform our normative understanding of an electoral system that is already well-understood to be unrepresentative and uncompetitive? For one, it shows that there are no unknown redeeming qualities hidden beneath the surface of the system. Indeed, if there is any normative implications of this dissertations, it is that the system may be even worse than was previously known. Chapter 2 shows that the only touted benefit of MNTV—that it grants voters greater choice by allowing them to vote for multiple candidates—can backfire as small shares of split-ticket voting can produce governing majorities representing electoral minorities. Further, the system rewards and generates strategic voting, thereby punishing those who exercise their ability to vote sincerely for candidates of multiple parties. Chapter 3 shows that MNTV can create extreme instability and uncertainty in a party system, thereby depriving voters of the ability to cultivate familiarity with political parties and hold those parties accountable for their successes and failures. Finally, chapter 4 shows that even relatively minor changes to an MNTV system can create years of even greater instability and uncertainty, which may benefit politicians seeking to exploit such a chaotic environment, but does not yield a political system that is stable, predictable, understandable, and representative for the average voter.

Chapter 2

Split-Ticket Voting in New Hampshire

2.1 New Hampshire's Multiple Nontransferable Vote: An Unusual Version of an Unusual System

In every congressional election, most state legislative elections, and many local elections, representatives are elected in single member districts in which every voter can cast one vote for one candidate and the candidate with the most votes wins. Some state and local legislators, however, are elected in multiple-seat districts that use a generalized form of this system that can elect more than one candidate in the same race. This generalized electoral system is called "multiple nontransferable voting" (or just "MNTV"), "unlimited voting" or"plurality at-large voting."

When MNTV is used to elect M seats: each party usually nominates M candidates; each voter may cast up to M votes (but may choose to cast fewer); a voter may not cast more than one vote for the same candidate; and the M candidates who receive the most votes are elected. While it is not uncommon for states to employ MNTV to elect representatives in districts of a magnitude (M) of two, New

Hampshire's state house elections are unusual both because magnitude varies between districts and because its larger district magnitudes are very high compared to any other state. Its smallest districts only have one representative apiece,[1] but its largest district has 11 representatives.

The MNTV system is not merely unusual, however. It presents voters with a more complex set of options, but also presents them with a much more complex set of systemic incentives that affect when rational voters should vote sincerely and when they should vote strategically. Studying its effects on voter behavior will show how well voters can understand and strategically adapt to complex and uncommon systems, even when the system is only used to elect one house of their state legislature while other simpler systems are used to elect the rest of the offices on their ballot.

New Hampshire presents the ideal case to study voters' ability to react rationally to the unusual and complex incentives of the MNTV system. The strategic incentives of the MNTV are determined, in part, by district magnitude. Variation in magnitude yields variation in strategic incentives, which can then be compared to variation in voter behavior to test how well voters understand and respond to the structural incentives of their electoral system. In this chapter, I explore the relationship between district magnitude and strategic incentives and derive testable hypotheses about the effects of magnitude on voter behavior, specifically the effects on voters casting split-ticket ballots.

[1] When MNTV is used to elect one candidate, the system is identical to a simple plurality vote. Each voter casts one vote for one candidate, then the candidate with the most votes wins. This makes MNTV one of several systems that can be described as a generalized form of the simple plurality vote.

Table 2.1: A Stylized Sample Ballot in an MNTV System.
This voter has cast a split-ticket, because they voted for candidates from more than one party.

Vote for three candidates:	
Republican Candidate #1	✓
Republican Candidate #2	✓
Republican Candidate #3	
Democratic Candidate #1	
Democratic Candidate #2	✓
Democratic Candidate #3	

Table 2.2: Hypothetical Results from an Election of Three Seats using MNTV.

Candidate	Vote Count	
Republican Candidate #1	10,011	**elected**
Republican Candidate #2	10,032	**elected**
Republican Candidate #3	10,075	**elected**
Democratic Candidate #1	9,992	
Democratic Candidate #2	9,922	
Democratic Candidate #3	9,983	

2.2 The Understudied MNTV System

Existing research says very little about the MNTV system, either in America or abroad. This is, most likely, due to four reasons. Firstly, it is a rare electoral system, at least at the national level. Secondly, within the United States, it is almost only used in select state legislative and local elections and not in elections which tend to capture the interest of Americanist and comparativist political scientists.

Thirdly, many of the theoretical works around more common systems do not directly apply to MNTV. It is not a proportional system, so findings regarding the effects of proportional systems should not and do not apply to MNTV systems. One could ask how will the effects of MNTV follow the effects of other majoritarian systems, such as the tendency of such electoral systems to foster a two-party system. However, even a cursory examination of states with MNTV confirms that they have as much of a two party system as any other state, which is a finding that is neither theoretically interesting nor surprising. Finally, there is very little observable variation in the relevant institutional variables of MNTV systems. This is especially true with district magnitude. Of the few states that use MNTV, almost all have districts of uniform magnitude. That uniform magnitude is, in turn, almost always two members per district. New Hampshire's wide variety of district magnitudes, including districts with as many as 11 members, presents variance with which to test theories that tie varying magnitude to vary strategic incentives in the MNTV system, such as the theory which I detail below.

Neither MNTV nor New Hampshire's uniqueness have completely eluded scholars' attention, however. Almost all of the work on MNTV focuses on the system's effects on descriptive representation. None of the work focuses on the strategic considerations that MNTV presents to voters, much less how these considerations vary with changes in magnitude or competitiveness. Rule (1990); Welch (1990); Matland and Brown (1992); Leal, Martinez-Ebers and Meier (2004); Trounstine and Valdini (2008) all examine the effects of MNTV systems on minority and women's representation in various local and state contexts. They all find, with various caveats and different levels of confidence, that MNTV lowers the number of minorities in office but slightly increases the representation of women. The system's suppression of minority representation can generally be attributed to the fact that MNTV has fewer and larger districts than SMD, given the same total number of representatives. Because minority areas are more likely to comprise the majority of a district if the districts are smaller and more plentiful, minority areas should control more seats in SMD districts. The success of women under MNTV can perhaps be attributed to the ability of parties to nominate more than one candidate for the same race; this allows parties to nominate men *and* women in the same race, where they may have only nominated men if forced to choose just one in an SMD election.

One common element of most of these works is that MNTV is treated as a binary variable: either the election has more than one seat or it does not. This fails to capture the potential effects of varying district magnitude. A better measurement for these studies would have been to treat MNTV as a generalized form of SMD, then measure district magnitude. This may have proved especially useful in Leal, Martinez-Ebers and Meier (2004)'s study, which examined school board elections, which in turn have district magnitudes that vary as greatly as New Hampshire's house districts.

Finally, one piece by Squire (1992) finds that New Hampshire has the least professional legislature in the country. New Hampshire, along with Vermont, has the highest representation of homemakers, students, and retirees in the legislature. This is interesting, although tangential, to this chapter because Vermont's state senate is the only other state house in the country with large-magnitude districts. Squire's paper does not, unfortunately, explore whether or not the professional attributes of the Vermont and New Hampshire state legislatures are caused by their electoral system.

In sum, MNTV has caught the attention of a small number of scholars. These scholars, however, have limited their research to who is elected and have largely ignored any features of the electoral system itself, such as district magnitude. There has been no attempt to study the unique strategic incentives that MNTV presents voters, to explain how these incentives vary *between* different MNTV systems, nor least of all how well voters respond to these different and varying incentives.

2.3 Voter Preferences, Split-Tickets, and Strategic Incentives under MNTV

In order to understand how well voters respond to the varying strategic incentives of MNTV, these varying incentives must be modeled. Specifically, I model the strategic incentives to cast straight-ticket votes, given a sincere preference to cast a split-ticket vote. Once these incentives have been modeled, the following sections will compare the variations in the incentives between districts to the number of split-ticket votes that were cast in each district. If the number of split-tickets significantly decreases when the incentives to cast straight-tickets increases, then I can conclude that voters understand and are responding rationally to their system.

I define a straight-ticket as ballot in which a voter votes only for the candidates of one party and a split-ticket as one in which a voter votes for candidates of more than one party. If a voter "undervotes" by casting fewer votes than they are allowed yet still only voted for candidates of one party, then I categorize their ballot as a straight-ticket. Ballots that do not fit the straight-ticket definition are split-tickets.

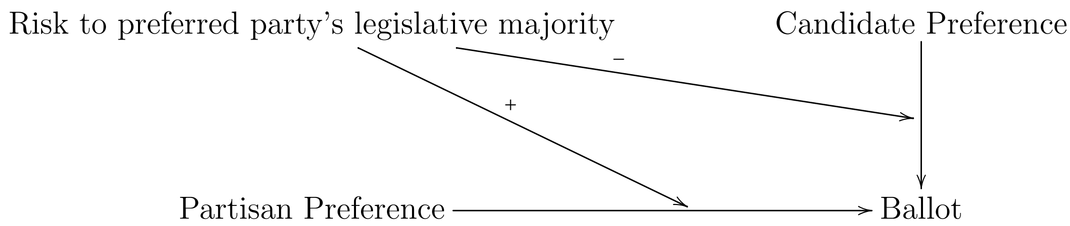
By elimination, split-tickets are all and only ballots in which a voter cast votes for candidates from more than one party. These definitions ignore any conceptual differences between voters who are loyal enough to a party to support all of its candidates and voters who are loyal enough to only vote for the party's candidates but not enough to fill out their entire ballot. This definition does, however, not only lend itself well to the data extrapolation method described later in the chapter, it also neatly divides the electorate into those who exclusively voted for one party's candidates and those who did not. As a final note, for the purposes of these definitions, each independent candidate is treated as the only member of a single-candidate party.

My theory is based upon the following assumptions. First, I assume that most voters prefer that a specific party control the state legislature; I call this party their "preferred party." Second, I assume that some voters may have preferences for specific candidates outside of their preferred party. For instance, a Republican voter may like one of the Democratic candidates running in their district; I call this candidate a "preferred opposing candidate." A voter may have more than one preferred opposing candidate.

Third, if a voter can express preference for both their preferred party and a preferred opposing candidate, then the will be inclined to do so. Expressing preferences for both is impossible in most systems, because the choice between a preferred opposing candidate and preferred party is zero-sum. In MNTV, however, a voter can express preference for both their preferred party and for preferred opposing candidate by casting a split-ticket. Finally and most importantly, voters will be less inclined to cast a split-ticket ballot, even at the cost of supporting their preferred opposing candidate, if split-ticket ballots carry an elevated risk of costing their preferred party control of the legislature.

The question, therefore, is when split-tickets carry more risk of flipping control

Figure 2.1: Basic Model of Ballot Choice as a Result of Voter Preferences and Party Risk.



of the legislature and when they carry less risk of doing so. I theorize that split-tickets carry the greatest threat to legislative majorities when district magnitude and competitiveness are *both* high. This is because there is only notable risk to the party's control of the legislature when a district's entire delegation can be flipped from one party to the other *and* when a flipped delegation is large enough to possibly affect the majority of the legislature. As I show below, split-tickets can easily flip competitive district's delegations and large-magnitude districts can more easily affect the legislative majority.

Competitive districts can be flipped between parties if even a relatively small share of the district's majority party's voters cast split-tickets. This is because each split-ticket both lowers the vote count of one or more majority party's candidates, while simultaneously increasing the vote count of one or more of the candidates from the district's minority party. Therefore, even if a majority of the voters cast a majority of each of their votes to one party, the other party can still sweep every seat in the district. These variable risks of split-tickets in safe and competitive districts is demonstrated in a series of hypothetical examples in table 2.3.

Risk to a party's legislative majority requires high magnitude, very simply, because a district with fewer seats has a smaller chance of affecting which party holds

Table 2.3: The Effect of 12 Split-tickets in a Hypothetical Election of Three Seats with 100 Voters.

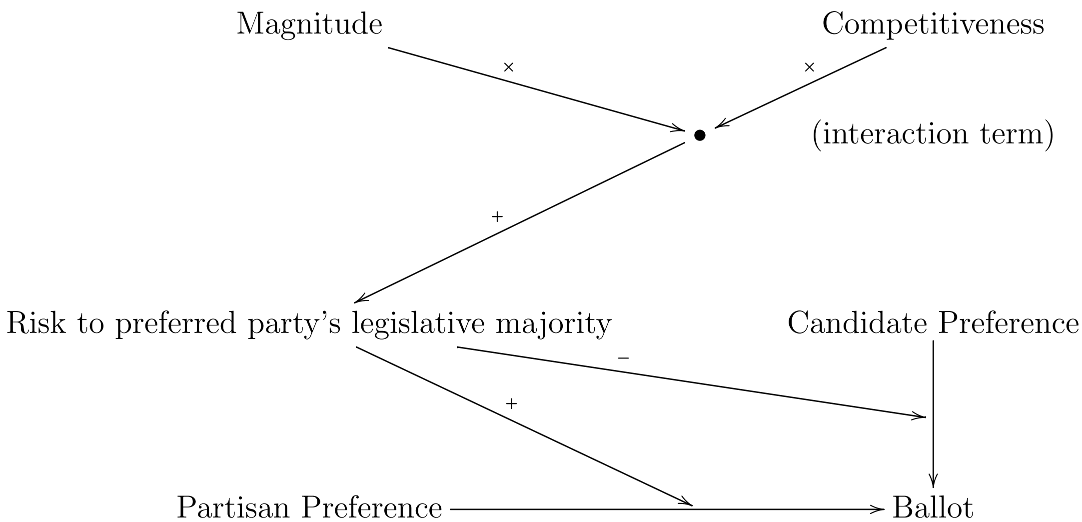
a. **Uncompetitive District**

Candidate	Vote Count: No Split-Tickets	Vote Count: Various Split-Ticket Outcomes			
Republican Candidate #1	**60**	**56**	**56**	**48**	48
Republican Candidate #2	**60**	**56**	**56**	**60**	**60**
Republican Candidate #3	**60**	**56**	**56**	**60**	**60**
Democratic Candidate #1	40	44	52	44	**52**
Democratic Candidate #1	40	44	40	44	40
Democratic Candidate #1	40	44	40	44	40

a. **Competitive District**

Candidate	Vote Count: No Split-Tickets	Vote Count: Various Split-Ticket Outcomes			
Republican Candidate #1	**51**	48	48	39	39
Republican Candidate #2	**51**	48	48	51	**51**
Republican Candidate #3	**51**	48	48	51	**51**
Democratic Candidate #1	49	**52**	**61**	**52**	**61**
Democratic Candidate #1	49	**52**	**49**	**52**	49
Democratic Candidate #1	49	**52**	**49**	**52**	49

Figure 2.2: Model of Ballot Choice, Including Determinants of Party's Risk.



the governing majority, even if every one of a small district's seats changes hands from one party to the other. This is especially true in New Hampshire House of Representatives, which has 400 seats; it is very unlikely that a district with one or two members will affect the governing majority. Therefore, even if a small number of split-ticket votes could determine the outcome of a single low-magnitude district, there is no risk to the voter's partisan interests to dissuade the voter splitting their ticket if they happen to like a candidate from the other party.

This leads to my main theoretical hypothesis, that *as competitiveness and magnitude both increase, split-ticket voting will decrease* (hypothesis one). Of the three hypotheses that I present in this chapter, this is the most important one because it is the only one that, if true, reflects a notable degree of strategic thinking by the voters. While the following two hypotheses are interesting, they do not represent an expectation of strategic voting.

Table 2.4: Risks Associated with Casting a Split-ticket Ballot in an MNTV District.

a. **Uncompetitive District**

Risk Type	Level of Risk: Low Magnitude District	Level of Risk: High Magnitude District
Risk of affecting legislative majority, given district flip	Low	High
Risk of flipping district	Low	Low
Net Risk	**Very Low**	**Low**

b. **Competitive District**

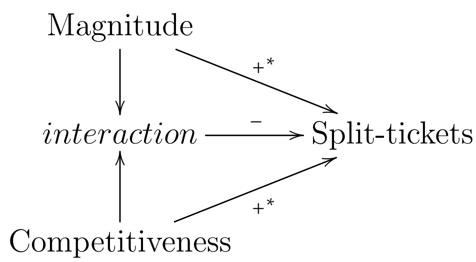
Risk Type	Level of Risk: Low Magnitude District	Level of Risk: High Magnitude District
Risk of affecting legislative majority, given district flip	Low	High
Risk of flipping district	High	High
Net Risk	**Low**	**High**

In addition to the interactive term between competitiveness and magnitude, the independent effects of both variables are of tangential interest as well. To that end, I include two expectations about the simple effects of competitiveness and magnitude to accompany the main theoretical hypothesis. Firstly, I hypothesize that, *when magnitude is at its lowest value in the dataset (2), the effect of competitiveness will be positive* (hypothesis two). This is because, in the absence of all strategic considerations, increased competition in a district means that the electorate is more torn between the parties, which suggests (but does not definitively mean) that a greater share of the district's voters are moderate. One should expect moderate voters to be less inclined to only support one party by casting a straight-ticket and, therefore, more likely to cast split-tickets.

This second hypothesis is different from the main hypothesis, which does not necessarily require that competitiveness has a positive effect when magnitude is low. The main hypothesis could be supported if competitiveness has a negative effect on split-tickets for all possible magnitudes, provided that competitiveness has a stronger negative effect in large-magnitude districts than in low-magnitude districts.

My third and final hypothesis is that *when competitiveness is at its lowest observed value, increased magnitude will increase split-ticket voting* (hypothesis three). If there is little or no strategic risk in casting a split-ticket, then one might expect that voters who can support more candidates simultaneously will be more likely to cast at least one vote for a candidate outside of their preferred party. A voter who can only support two candidates should less inclined to give half of his ballot to his preferred party's opponents, but a voter with eleven votes should be more willing to give nine percent of his ballot to the other party by casting one vote for his preferred party's opponents. Similarly, one might expect a voter with eleven votes to more likely cast a split-ticket on a whim or by accident.

Figure 2.3: Visual Representation of Hypothesized Effects.



*when the other variable is at its minimum.

This third hypothesis is different than the main hypothesis by the same logic that the second hypothesis is distinct from the main hypothesis. The main hypothesis does not require that magnitude has a positive effect on split-tickets in uncompetitive districts. The effect of magnitude can be negative across all possible values of competitiveness yet still support the main hypothesis, as long as the effects of magnitude are less negative in uncompetitive districts than in competitive ones.

A final theoretical consideration is whether or not the competitiveness of the entire chamber conditions the effects of district competitiveness and district magnitude. If a party's expected share of seats in a chamber is so large that even a complete flip of the largest district is unlikely to affect the party's control of the chamber, then by this chapter's own logic, the district competitiveness will be less likely to suppress split-ticket voting in the largest districts.

If a legislative chamber is, as a whole, uncompetitive enough that even the largest district cannot change which party controls the body, then there is less risk overall to the expectant majority party's control. If there is less risk to the party overall, then there is less reason for supporters of either party to cast straight-ticket votes strategically when they have a sincere desire to cast split-tickets, because strategic straight-tickets are responses to perceived threats to the legislative majority of the voter's preferred party. Therefore, in uncompetitive legislatures, competitive high-magnitude districts may not have any special relationship with split-ticket voting; their levels of split-ticket voting might instead be affected by a simple linear combination of, rather than an interaction between, their levels of competitiveness and magnitude.

The aggregate competitiveness of the chamber would need to be accounted for in a study that covered more states and more periods of time. This study, however, only covers one election of one chamber. Because this particular election – the 2016 New Hampshire state house election – was especially competitive and resulted in the chamber majority switching from the Democratic Party to the Republican Party, there is no reason to suspect that a lack of aggregate competition suppressed the hypothesized interactive effects of competitiveness and magnitude. Omitting a consideration of aggregate chamber competitiveness creates bias in favor of the null hypothesis, because the variable should work against the main hypothesis. Its omission, therefore, should not be of any inferential concerns.

2.4 Data and Measurement

The dataset that I use to test my theory is comprised of every multiple-seat district in the 2016 election of the New Hampshire House of Representatives, excluding floterial districts. I exclude single-member districts because such districts only have one seat, which means that the voters in those districts only have one vote. When a voter only has one vote, the voter cannot cast a split-ticket ballot. Therefore, because split-ticket ballots are the dependent variable, single-member districts cannot be included in the dataset.

New Hampshire's floterial districts are house districts that overlap the regular house district map. Every voter is in a regular house district, but only some are also in a floterial district. The boundaries of the regular and floterial districts are incongruent, save for the fact that neither type of district crosses county lines (see figure 2.5). Almost every floterial district is a single-member district and, therefore, already excluded from the dataset. I exclude the few remaining multiple-member floterial districts, to avoid adding an unnecessary level of complication to the model.

These exclusions leave 88 out of 204 districts and 281 out of 400 members. The level of analysis is the district, so the number of observations is 88. All raw data for these units were obtained from the New Hampshire Secretary of State's website.

There are three main variables to derive for each district from the raw data obtained from the Secretary of State: magnitude, competitiveness, and number of split-tickets. The first, magnitude, is the easiest to measure. Magnitude is simply the number of seats in a district, which is the same as the number of winners reported for the district.

To measure the competitiveness of a district, I create a metric for MNTV results that compares the average vote share for candidates of both parties. The metric is bounded by design from zero to 100. Lower values reflect that the district is safely in the hands of either the Democratic or Republican parties. Higher scores mean that the district is more competitive and could more easily change hands between the parties. I calculate this metric as follows:

$$100 - 100\times \left| \frac{\text{mean vote count for GOP candidates in district} - \text{mean vote count for Dem candidates in district}}{\text{ballots cast in district}} \right|$$

Measuring the number of split-tickets cast in a race is more complicated, because data is not available at the ballot level. However, some information about the number of split-tickets cast can be extrapolated from the aggregate vote counts for the candidates in each district. Given the observed vote counts for each candidate and given the number of voters, one can calculate the minimum number of split-tickets that must have been cast to create the observed results. The minimum number of split-tickets can be found by the following formula:

$$\text{highest vote count of GOP candidate in district} +$$
$$\text{highest vote count of Dem candidate in district} -$$
$$\text{ballots cast in district}$$

The number of split-tickets that were actually cast will always be greater than or equal to the returned value. When this formula returns a positive value, it means that the observed vote counts could not have been achieved without split-ticket votes. If this formula returns a negative value or a value of zero, then the observed vote counts could have been achieved without any split-tickets. Furthermore, if the value is negative, then the results can be explained without split-tickets even if the number of voters is decreased by the magnitude of the returned value. For example, if the formula returns a value of negative 5, then the same results could have been achieved even if five fewer voters cast ballots. Negative values can therefore be seen as degrees of freedom between the observed results and an alternative result that requires split-tickets. More practically, negative values of increasing magnitude can be interpreted as decreasing probabilities that any split-tickets were cast at all.

This formula works because the sum of the parties' best vote counts represents the number of voters that are needed to explain the results without any split-tickets. The difference between this sum and the actual number of voters, therefore, is the number of voters who must have cast a split-ticket. Conveniently, this formula works even when one or more voters undervote by casting fewer votes than they are allowed to.

The obvious problem with this metric as a measurement of split-ticket votes is that it is biased downward. The actual number of split-tickets cast in an election will almost certainly be much higher than the metric, very rarely be exactly equal to the metric, and never be lower than the metric. However, if one's goal is to compare the frequency of split-tickets across districts, rather than estimate the exact number of split-tickets in a given district, this metric provides a useful - if crude - tool with which to do so.

In order to compare this measurement of split-tickets across districts of different sizes, I divide it by the number of voters in the district. This provides a measure of the minimum number of split-tickets per voter. Finally, I multiply this value by 100, which creates a variable, "percent split-ticket" that is bounded by construction between -100 and 100.

The competitiveness metric and the split-ticket metric can only be calculated if the number of voters is known. Unfortunately, the exact number of voters in an MNTV election is not inferable from the observed results. In most elections, the number of voters is simply the sum of the vote counts of each candidate. This simple summation does not work in MNTV, however, because each voter can cast more than one vote. Nor can the number of voters be calculated by dividing the number of votes by the number of seats, because each voter can undervote by casting fewer votes than the number of seats. As the amount of undervoting increases, the more negatively biased the $\frac{votes}{seats}$ metric becomes.

To estimate the number of voters, I make the assumption that those who voted in the state senate election are the same voters as those who voted in the state house election. This is a tempting tool for estimating the number of house voters because the Vermont Senate, unlike many state senates, is elected every two years instead of four, and it is elected on the same schedule as the house. Therefore, even if one does not know how many people voted in the house election, they can still know how many people voted in the other state legislative race on the same ballot. The obstacle to using senate turnout as a proxy for the house is that the senate and house districts and incongruous. Fortunately, the nature of reported New Hampshire returns allows for the senate data to be deconstructed to the local level. Because neither small towns nor the wards of larger towns are ever divided between districts at either the house or senate level, the turnout at the local level can then be reconstructed at the house district level and then used in the competitiveness and split-ticket metrics.

Although I use an interaction term between competitiveness and magnitude, I do not rescale either of the variables to center them on their means. This is occasionally done in the presence of interaction terms to make the regression coefficients more interpretable, especially if either of the variables do not have a possible or observed natural value of zero. In this particular case, the natural zero points of both variables are useful substantively interesting. The intuitive expectations laid out in the previous section, predict the effects of both magnitude and competitiveness when the other variable is at its lowest possible value. Mean-centering would make testing these expectations less straightforward, so I leave both magnitude and competitiveness at their observed values.

I present the distributions of competitiveness, magnitude, competitiveness*magnitude, and percent split-ticket in figure 2.4. Despite the fact that the independent variables and their interaction term are not normally distributed, the results in the next section are robust to the removal of the tails and outliers of each variable. These robustness checks are discussed further in the next section. A more geographical representation of the magnitude and partisanship of New Hampshire's districts can be found in figure 2.5.

2.5 Models and Results

2.5.1 Model Specification

I run four OLS models on my data. For all four, the dependent variable is split-tickets per voter. All four models include both magnitude and competitiveness as independent variables. Models number two and four include the interaction between magnitude and competitiveness; models number one and four exclude the interaction for the sake of comparison.

Almost all possibly relevant control variables are not available at the district level. I rely, therefore, on fixed effects at the county level to account for heterogeneity between units. This works in practice because house district boundaries do not cross county lines; each district is in one county. To work analytically, however, this assumes that any relevant control variables are relatively constant across districts in the same county. While this assumption is difficult to test, a cursory examination of the state suggests a degree of homogeneity within each county, at least in terms of urbanization, terrain, and development. To that end, I add county-level fixed effects to models number three and four, which are otherwise identical to models one and two, respectively. The specifications of these models are shown in table 2.5.

Model number four tests both the intuitive expectations and theoretical hypotheses of this chapter, because it accounts for fixed effects and includes an interaction term for competitiveness and magnitude. The significance and direction of β_3, the coefficient for the interaction term, determines whether or not we can reject the null hypothesis, that competitiveness and magnitude decease split-ticket voting together. In models number two and four, in which the interaction term is included, the significances and directions of β_1 and β_2, indicate the effects of magnitude and

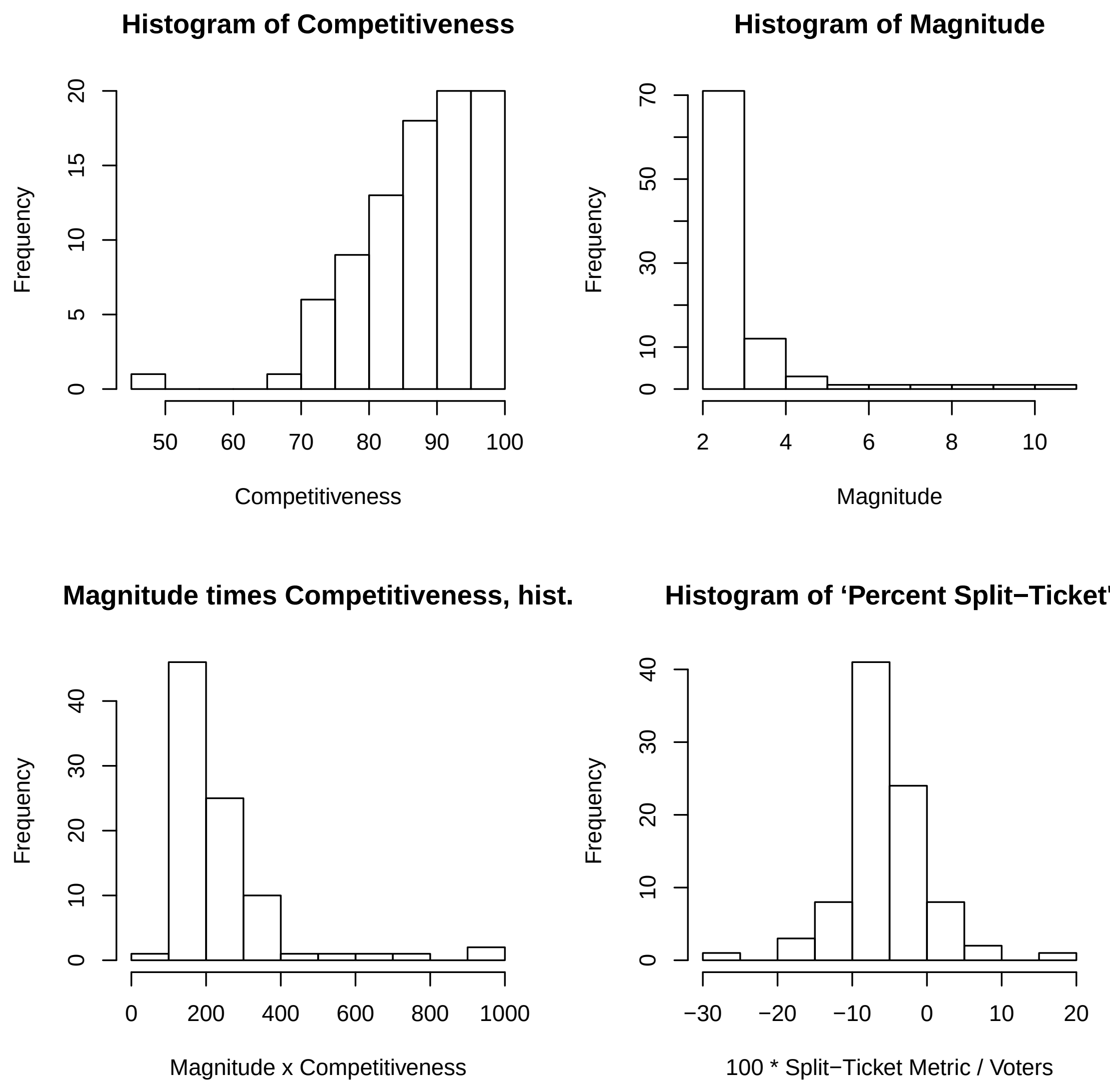


Figure 2.4: The Distributions of the Independent Variables, their Interaction, and the Dependent variable.

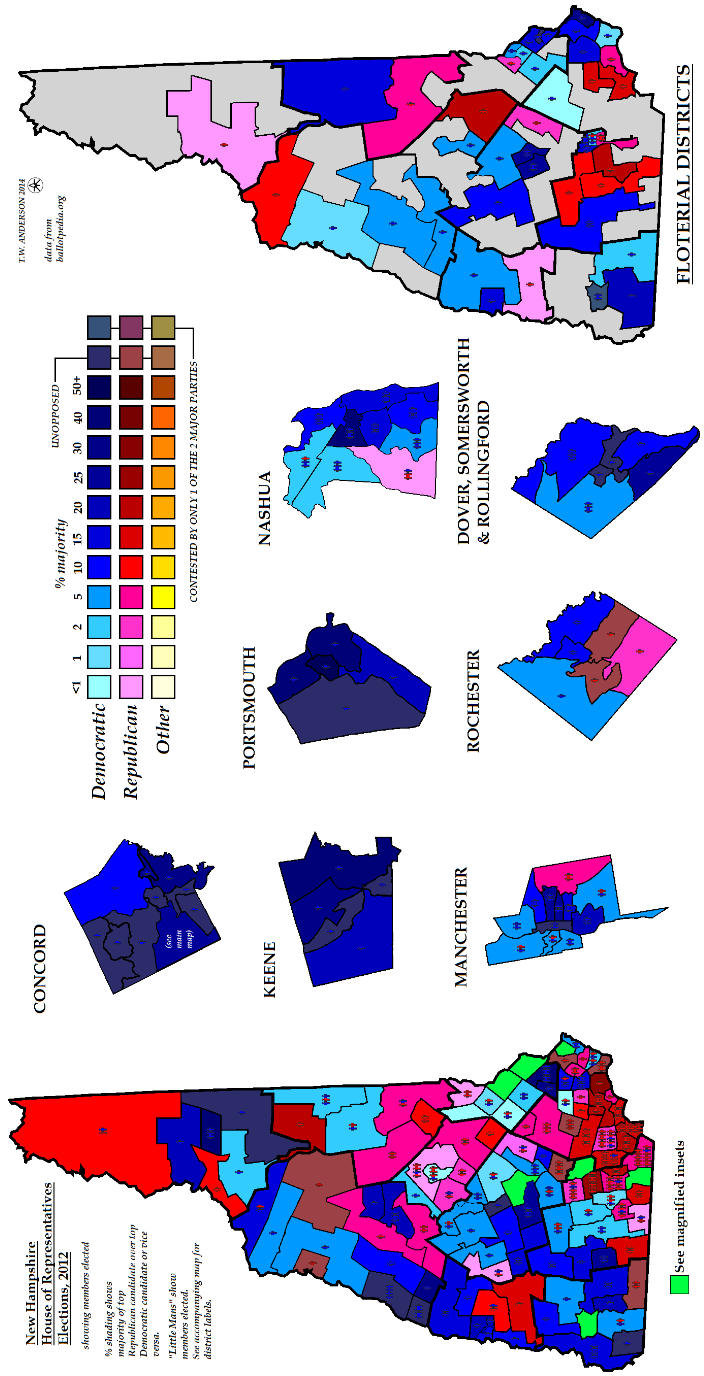


Figure 2.5: A map of New Hampshire's state house districts in 2012, conveying district magnitude and partisan distribution of seats. (T.W. Anderson, 2014)

Table 2.5: The Specification of the Four Models

	Dependent variable:
	'Split Tickets per Voter'
Model 1	$\beta_0 + \beta_1 * \text{Magnitude} + \beta_2 * \text{Competitiveness} + \epsilon$
Model 2	$\beta_0 + \beta_1 * \text{Magnitude} + \beta_2 * \text{Competitiveness} +$ $\beta_3 * \text{Magnitude} * \text{Competitiveness} + \epsilon$
Model 3	$\beta_0 + \beta_1 * \text{Magnitude} + \beta_2 * \text{Competitiveness} + \text{County Fixed Effects} + \epsilon$
Model 4	$\beta_0 + \beta_1 * \text{Magnitude} + \beta_2 * \text{Competitiveness} +$ $\beta_3 * \text{Magnitude} * \text{Competitiveness} + \text{County Fixed Effects} + \epsilon$

competitiveness, respectively, when the other variable has a value of zero. The intuitive expectations of this chapter, as detailed above, suggest that both of these coefficients should be positive and significant, because increased magnitude should increase split-ticket voting in the safest districts and increased competition should increase split-ticket voting in the lowest-magnitude districts.

2.5.2 Results, Findings, and Robustness Checks

The results for these models are shown in table 2.6. Models number two and four both show that the null hypothesis can be rejected. The coefficients for the interaction term are negative and significant in both models. This means that, all things equal, as competitiveness and magnitude *both* increase, a smaller share of voters will cast split-ticket votes. This supports the main hypothesis as well as the theory that, in competitive high-magnitude districts, voters will be less likely to cast split-ticket votes because doing so will endanger their preferred party's control of the house.

Figure 2.6 shows the marginal effects of both competitiveness and magnitude, as found in model number four, and visualizes why the null hypothesis can be rejected. This figure shows that the interaction between magnitude and competitiveness is not merely strong enough to diminish the positive effects of the individual variables, it is strong enough to reverse their positive effects. In two-member districts, increased competitiveness significantly increases split-ticket voting, but in districts with more than five members, competitiveness significantly decreases split-ticket voting. In districts with three, four, or five members, competitiveness has no significant total effects, because for those values of magnitude, the positive simple effects of competitiveness are balanced with the negative interactive effects of competitiveness.

Similarly, in the most safe districts, in which the difference in the average vote share of the GOP candidates and the average vote share of the Democratic candidate is greater than 26 points, increased magnitude has a positive effect on split-ticket voting. When the difference in average vote shares is less than 11 points, however, increased magnitude has a positive effect on split-ticket voting. For values of competitiveness between 11 and 26, magnitude has no significant total effect, because for that range of competitiveness, the simple positive effects of magnitude are balanced with the negative interactive effects of magnitude.

Models number one and three show that, when the interaction term is not controlled for, neither competitiveness nor magnitude have a significant effect on split-ticket voting. This is as expected, because when the interaction is not accounted for, the negative effect of magnitude in competitive districts contradicts the positive effect of magnitude in safe districts, yielding a net insignificant effect overall. Similarly, the positive effect of competitiveness in low-magnitude districts is negated by the negative effects of competitiveness in high-magnitude districts. This can be seen

Table 2.6: The Effects of Magnitude and Competitiveness on Split-ticket Voting, Examined at the District Level.

	Dependent variable:			
	'Split-Ticket per Voter'			
	(1)	(2)	(3) fixed county effects	(4) fixed county effects
Magnitude	−0.394	15.943**	−0.311	16.753**
	(0.368)	(6.386)	(0.400)	(6.746)
Competitiveness	0.089	0.579***	0.089	0.611***
	(0.071)	(0.203)	(0.078)	(0.220)
Competitiveness:Magnitude		−0.190**		−0.198**
		(0.074)		(0.078)
Constant	−12.129*	−54.296***	−11.859	−54.562***
	(6.396)	(17.584)	(8.060)	(18.567)
Observations	88	88	88	88
R^2	0.034	0.104	0.133	0.201
Adjusted R^2	0.011	0.072	0.007	0.074
Residual Std. Error	5.931 (df = 85)	5.745 (df = 84)	5.940 (df = 76)	5.739 (df = 75)
F Statistic	1.473 (df = 2; 85)	3.235** (df = 3; 84)	1.059 (df = 11; 76)	1.575 (df = 12; 75)

Note 1: *p<0.1; **p<0.05; ***p<0.01

Note 2: The reference unit for the fixed effects models is Belknap County

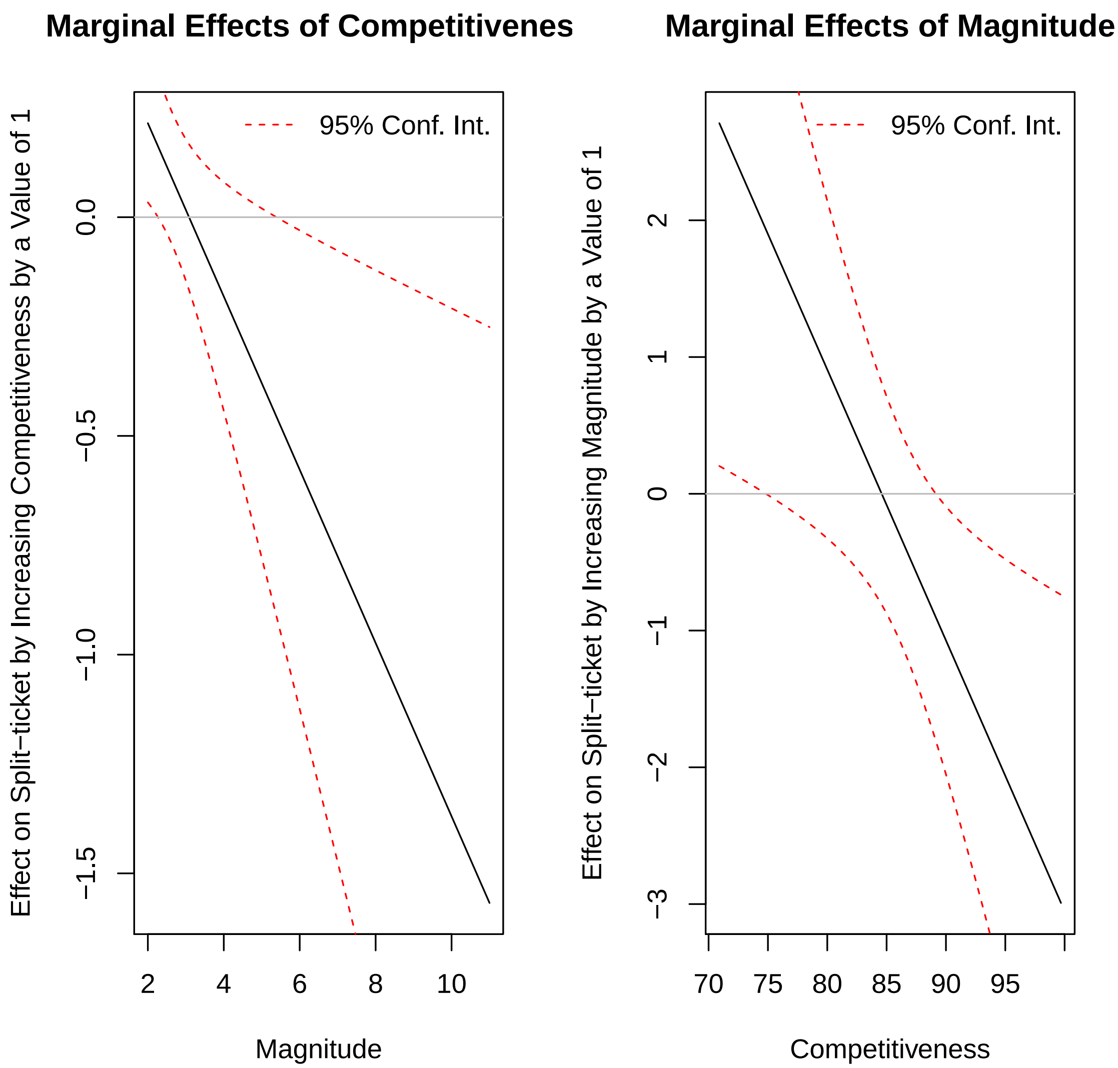


Figure 2.6: The Marginal Effects of Competitiveness Given Varying Levels of Magnitude (left) and the Marginal Effects of Magnitude Given Varying Levels of Competitiveness (right).

in the marginal effects plot in figure 2.6, which shows that 95 percent confidence intervals are not outside of zero when magnitude is at its mean of 3.04 or when competitiveness is at its mean of 87.37.

The marginal effects plots, along with the results from both the second and fourth models, also support the intuitive expectations of hypotheses two and three that I included alongside my theoretical hypothesis. First, when competitiveness is at its lowest value, magnitude has a simple positive effect on split-ticket voting. This suggests that, when all strategic incentives are accounted for, voters are more likely to cast a split-ticket ballot if they are able to support a greater number of candidates simultaneously. Similarly, a simple positive effect is found for competitiveness when magnitude is at its lowest value. This suggests that more competitive districts have more moderate voters who are more likely to split their tickets between parties when all strategic considerations are accounted for, even if they only have a small number of votes to cast. Note that both of these conclusions are separate from conclusions regarding the interactive effect, because the main hypotheses could have been supported even if either or both of the coefficients for the individual variables were negative or insignificant.

The significance of the results are very sensitive to, but not completely dependent upon influential data points. None of the observations have a Cook's distance greater than one, yet removing the most influential observations impacts the significance of the interaction term between competitiveness and magnitude. Cheshire's 11th district is the most influential observation, because it is the least competitive district by far. Removing this district causes the interaction term to lose its significance, even though it preserves the direction of the effect. It would be problematic if the results relied entirely upon one observation for its significance. However, even upon the removal of Cheshire's 11th district, evidence of a significant relationship remains.

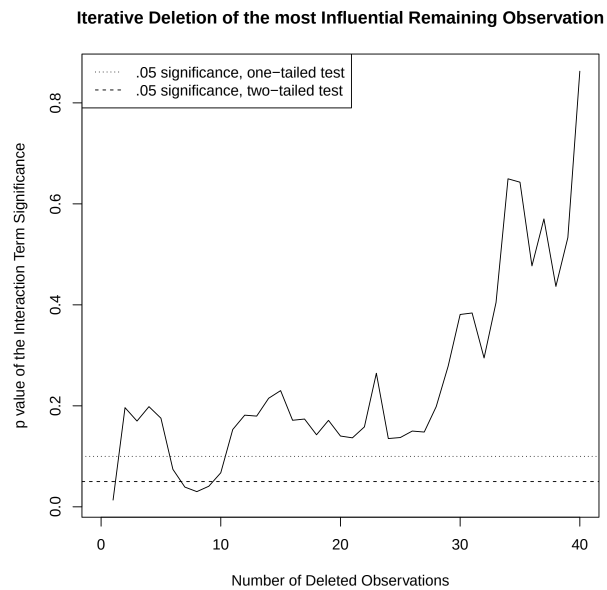


Figure 2.7: The p-values of the Interaction Term of Magnitude and Competitiveness, as the Most Influential Remaining Datapoint is Iteratively Removed from the Dataset.

If one continues to iteratively remove the most influential remaining data point, significance quickly returns to the results. Furthermore, the sign of the coefficient remains negative until 40 of the 88 observations have been eliminated. This suggests that the relationship does indeed exist. This iterative deletion is show in figure 2.7. Note that, in this plot, a gradual upward trend in p-values should be expected as observations are removed and n decreases. Most importantly, this suggests that the significance of the model will become more stable and more robust to outliers as more data is added to the dataset from more years.

2.6 Conclusion

These results show that voters' tendency to cast straight-ticket ballots in MNTV systems correlates with their strategic systemic incentives to do so instead of casting a straight-ticket vote. This suggests that voters, at least in New Hampshire, are aware of the unique and somewhat complex incentives of their unusual electoral system, despite the fact that the system is somewhat unusual even within the context of New Hampshire politics because their state senate, governor, and congressional delegation are all elected using a simple single member plurality system. Furthermore, because the systemic incentives of MNTV are dependent upon the context of each election, particularly the competitiveness and magnitude of the district, voters are able to adapt to MNTV's changing strategic incentives over time. This is reassuring at a normative level, because it indicates that voters may have a more sophisticated understanding of the broader electoral contexts in which they vote.

Chapter 3

Strategic Party Formation and Candidate Entry in Vancouver, BC

3.1 Introduction

In British Columbian municipal elections generally and in Vancouver elections especially, party systems are highly volatile and unstable. Parties rise, merge, split, and fall quickly. Candidates form parties, switch parties, and abandon parties frequently. Since 1938, 19 different parties and 17 independents have received at least one tenth of the vote in either a mayoral or city council election. Yet, of these parties, only three (the Non-Partisan Association, the Coalition of Progressive Electors, and the Green Party) existed for more than 20 years.[1] Of those three, 50-year-old COPE appears to be approaching the end of its lifespan. Only the NPA has existed the entire 80-year span of the dataset, while the other remaining long-lived party–the Green Party–is 34 years old. The average lifespan of a Vancouver municipal party is only 6.58 years or 3.90 elections while the median lifespan is just one election. Lest

[1] The Civic Action Association existed over the span of 21 years, but it only contested elections in 1945, 1964, 1966. By all reckoning, it is two separate parties with the same name, one in 1945 and one in the mid-sixties.

this give the impression that Vancouver's history is characterized by meager small parties challenging the perennial NPA, even the shortest-lived of these parties, along with a fair number of independents, have had many significant victories over the past century while the NPA has suffered its share of defeats.

The rapid rise and fall of Vancouver's municipal parties combine with several other trends and patterns that together create the specific and theoretical questions this chapter aims to answer. The first trend is an increase in the effective number of parties by all metrics, though some metrics show a sharper and accelerating increase while others show a slower and even decelerating increase. The second trend is a gradual yet steady decline in the success of the NPA over time. Thirdly, while the NPA has proven to be the most resilient party, it has had a problem with candidate defections throughout its history. The final trends regard independent candidates: a recent increase in the number of independent mayoral candidates, a generally volatile yet directionless trend of the quantity of independents, and a decreasing trend of independent electoral success.

These trends are all directly or indirectly tied to each other. The goal of this chapter is to explain how these trends affect each other and how they are affected by outside forces. Some of these explanations are idiosyncratic, arising from random features and decisions of the city, its voters, and its candidates. Underneath this noise, however, are patterns driven by multiple nontransferable voting that illuminate how this particular electoral system shapes and drives the incentives of the actors involved. Specifically, this chapter finds that party systems under MNTV elections become destabilized as the vote share of the largest party increases, which manifests in more new candidates running for office, more candidates running as independents, and more new parties. Conversely, if the vote share of the largest party increases in a prominent concurrent election, such as the mayoral race in the case of Vancouver, the party system becomes more stable.

3.2 Background of Political Parties and Electoral Systems in Vancouver

Vancouver was incorporated in 1886. The city started out with a mayor elected at-large, a ten-member council elected by MNTV in five two-member districts, and a three-member school board elected by at-large MNTV; all fourteen elected offices have one-year terms. Over the next 35 years, many minor changes were made to the government, most of which gradually expanded the number of elected city officials. By 1912, the City Council had 16 members, each elected by MNTV in two-member districts for one-year terms; the school board had seven trustees elected by at-large MNTV for one-year terms; and the now-elected Park Board had five members elected by at-large MNTV for staggered two-year terms. Following a reduction in council size to eight in 1916, each elected in single member districts, Vancouver briefly adopted proportional representation for its city council from 1921 through 1923. Following a return to single member districts for council elections in 1923, in 1928 the Council, School Board, and Park Board are increased to 12, nine, and seven members, respectively.

This chapter's temporal scope begins in 1936, when Vancouver first adopted at-large MNTV elections of its city council. 1936 is also the year that municipal parties first appeared and it happens to be the earliest year that reliable data becomes readily available at the candidate level. It is, therefore, both a substantially and practically good starting point for examining MNTV-driven behavior in the city.

Political parties first emerged in Vancouver city politics in the 1930s, first with the Co-operative Commonwealth Federation (CCF), a branch of the national party of the same name, then shortly thereafter by the center-right Non-partisan Association

(NPA). Despite winning the mayorship in 1938 and 1940 (and beating the NPA's mayoral candidate in 1944, coming in at a close second to an independent), the CCF had very low electoral success across the board until its eventual death in 1953. Incidentally, the poor performance of the CCF is echoed in following parties that were local branches of provincial and national parties, namely the New Democratic Party,[2] which have generally struggled to compete against parties indigenous to Vancouver. The only exception to this general rule is Green Party, which currently holds seats on all three boards due to an electorate that is highly divided between multiple small parties.

The era from 1936 to 1966 is characterized both by especially high party instability, even compared to the rest of Vancouver's municipal party, and by the dominance of the NPA, especially in non-mayoral races. Parties of this era (with the exception of the relatively steady NPA) rose and collapsed quicker and more unpredictably than the parties that came after (see figure 3.1). By contrast, parties since 1966 usually follow a simple lifespan; they are created, slowly rise to their strongest point, then fall to their inevitable death at either relatively slow yet constant pace (see The Electoral Action Movement, the Coalition of Progressive Electors, or Vision Vancouver). This phenomenon was partially driven by the frequent election cycle (yearly as opposed to less frequent election thereafter) and is explored more in the theoretical and empirical sections of this chapter.

Vancouver's second distinct political era began in 1968 and was characterized by diminished NPA dominance, the rise of stronger, longer-lived challengers, and a significant increase in the effective number of parties at any given time. Its beginning was marked by both a significant change to the electoral system, as the entire city council was placed on the same cycle thus increasing the magnitude to ten, but also

[2] The NDP was, itself, a direct descendant of the CCF.

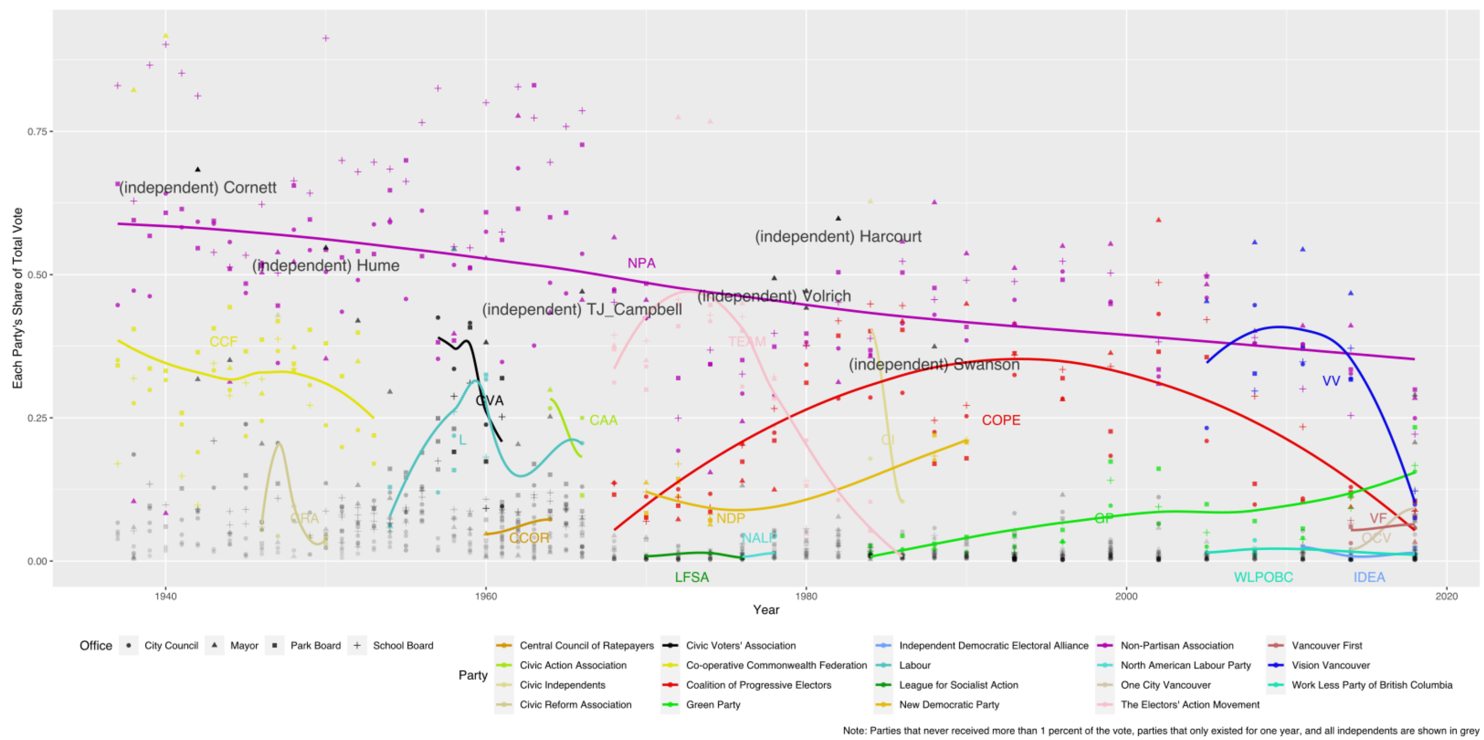


Figure 3.1: The Electoral Performance of Vancouver Municipal Parties over Time.

to the party system, as every pre-1968 party besides the NPA ceased to exist after 1966 and three other major parties quickly rose to fill the vacuum. These three new challenging parties were the Coalition of Progressive Electors (COPE, 1968 - present, originally the "Committee of Progressive Electors"), the Electors' Action Movement (TEAM, 1968 - 1986), and the New Democratic Party (NDP, a branch of the national party, active in Vancouver from 1970 through 1974 and from 1988 through 1990).

The longer lifespans of NPA's challengers came with greater overlap between their lifespans. This overlap, in turn, lead to a higher effective number of parties Conversely, while its challengers grew stronger, the NPA itself found itself in a much weaker position during this second era than its first. From 1968 to 2002, the NPA only won a majority of the seats in seven of the 16 council races, a plurality of the seats in two others, and split the council in half with COPE in 2002. The NPA only won half of the mayoral races during this period, and lost seven mayoral contests in a row starting in 1972. The NPA faired better in school and park board elections, in which they won a majority of the seats in ten and twelve of the elections, respectively.

It is possible that Vancouver's party system is currently toward the beginning of a third era, one that began somewhere between 2005 and 2014. While the first era was defined by quick burning parties, the second by slow burning parties, and both by the endurance of the NPA, Vancouver's emerging political system seems to be defined both by the emergence of new small parties and the weakening and shrinking of old large parties. The new parties since 2005 include Vision Vancouver, the Work Less Party of British Columbia (WLP), the Independent Democratic Electoral Alliance (IDEA), Vancouver First, and One City Vancouver. Of these, only Vision Vancouver, which won both the mayorship and council majority in 2008, 2011, and 2014, has had much absolute success at the polls. However, the other four smaller new parties have still had significant *relative* success because there is no longer a strong party in

Vancouver, the maximum vote share for any office was lower in 2018 (30%, cast for NPA's Park Board candidates) than in any other election in Vancouver's history.

The large parties are in crisis and only remain notable because no large party has risen to take their place. By vote share, 2018 was the NPA's fifth worst mayoral race (28% of the vote), second worst council race (25% of the vote), worst school board election (22% of the vote), and worst park board race (30% of the vote). COPE, the longtime strongest challenger to the NPA, did not win a single seat in 2014. Following this blowout, all but two of COPE's candidates left the party and COPE did not even field a mayoral candidate in 2018. Further, 2018 was COPE's worst council race (9% of the vote and only three candidates), second worst school board election (9% of the vote and only two candidates), and fifth worst park board race (10% of the vote and only two candidates).

The last two elections also saw large defections from both the NPA and COPE. On the 2014 ballot, COPE loyalists were outnumbered by COPE defectors; only three of COPE's 19 candidates had previously run under the COPE banner, but three former COPE candidates ran as independents and five ran as Vision Vancouver Candidates. Of COPE's nineteen 2014 candidates, all but one quit the party and quit city politics by 2018. In terms of defections, the span from 2014 to 2018 was the worst inter-election period in the NPA's history. Like COPE, the NPA's candidates were outnumbered by NPA defectors; the NPA ran seven candidates, while five former NPA candidates ran as independents, three ran as Coalition Vancouver Candidates, and two ran Vancouver First candidates.

Changes in Vancouver's party system over time can be seen in table 3.1 and figure 3.2. This chapter uses three measurements of party system fragmentation for each office. N_V and N_S are common metrics, based on the parties' relative shares of votes and seats. Vancouver lends itself to a new metric, N_C, which is based upon the relative share of candidates nominated by each party. Similarly to N_V and N_S, N_C is calculated as $\frac{1}{\sum c_i^2}$, where c_i is the share of candidates nominated by party i (e.g. if party i nominates ten candidates and there are 100 candidates between all the parties, then c_i is .1). In most systems, in which parties might be expected to nominate an equal number of candidates, N_C would simply be the same as the nominal number of parties.[3] In Vancouver, however, in which parties nominate a highly variable number of candidates ranging from one to M, N_C provides a metric for measuring how fragmented the nominees are between the various parties.

Figure 3.3 shows the effective number of independents over time, which is calculated by multiplying the effective number of parties (N_V) by the proportion of votes cast for independent candidates: $I_V = N_V \times \frac{\sum v_{independent}}{\sum v_i}$, where N_V is the effective number of parties with each independent treated as their own party, $\sum v_{independent}$ is the sum of all votes cast for independents, and $\sum v_i$ is the total number of votes cast. Visualizations of the flow of candidates between parties can be seen in figures 3.4 which shows the total number of candidates switching between parties between given elections and 3.5 which allows for the allegiances of individual candidates (albeit unlabelled for the sake of presentation) to be tracked over time.

3.3 Large Parties and Unstable Party Systems

This section theorizes that the party system of a set of MNTV institutions will be destabilized, as measured by the actions of the candidates therein, when the vote share of the largest party increases, thereby creating a pattern wherein the success of one party in a MNTV election yields greater party fragmentation in the next election.

[3] Take, for instance, an election in which five parties compete for ten seats. If each party nominates a full slate of ten candidates, then $N_C = 5$. It should be noted further that in all single-winner elections, wherein each party can only nominate one candidate, N_C will always equal the number of candidates.

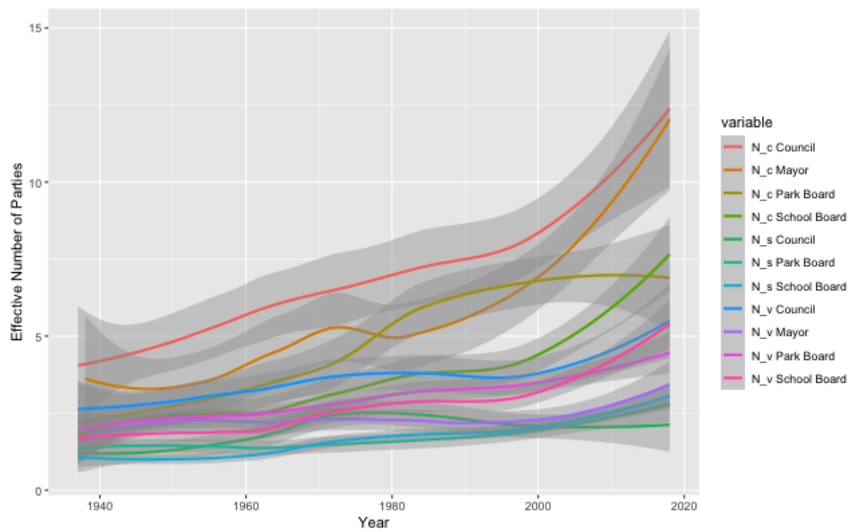


Figure 3.2: The Effective Number of Parties over Time, where N_C is Calculated by the Share of Candidates Nominated by each Party, N_V by the Share of Votes Received by each Party, and N_S by the Share of Seats Won by each Party.

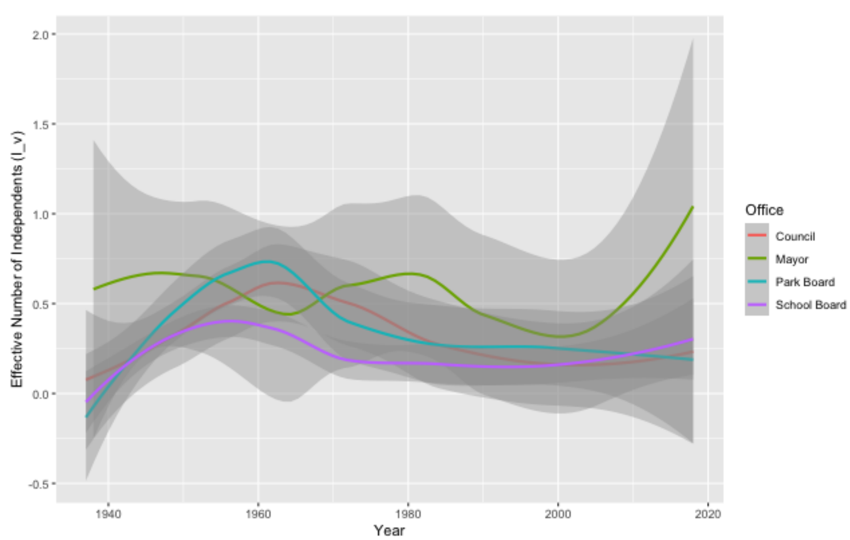
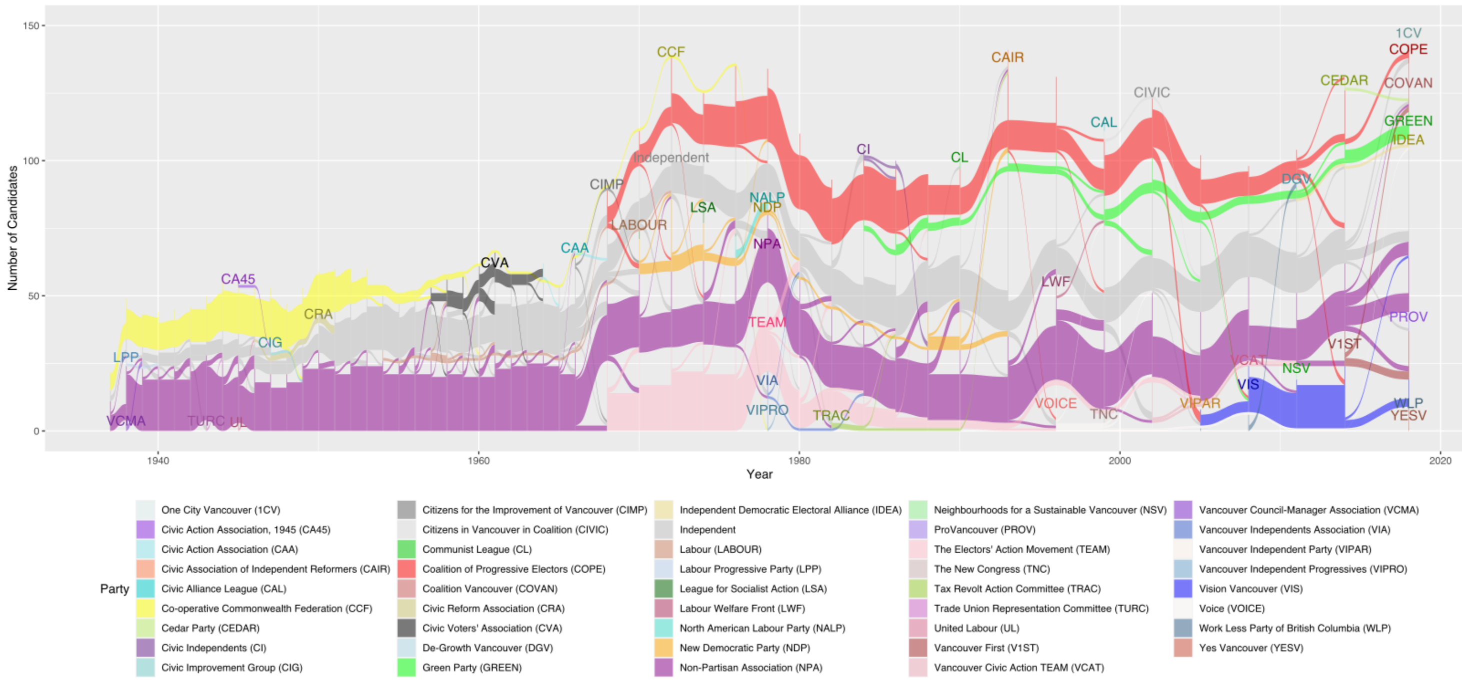


Figure 3.3: The Effective Number of Independents (I_V) over Time.



Notes: If a party nominated a total of 2 or fewer candidates during its entire existence, then its candidates are treated as independents.
The height of a verticle increase or decrease represents the number of candidates within a party who ran for office for the first or last time, respectively.
If a candidate does not run for office for more than 20 consecutive elections, then they are treated as a new candidate if they run for office again.
For shorter gaps between candidacies, candidates retain their most recent preceeding partisan affiliation.

Figure 3.4: Gross Exchange of Candidates between Parties.

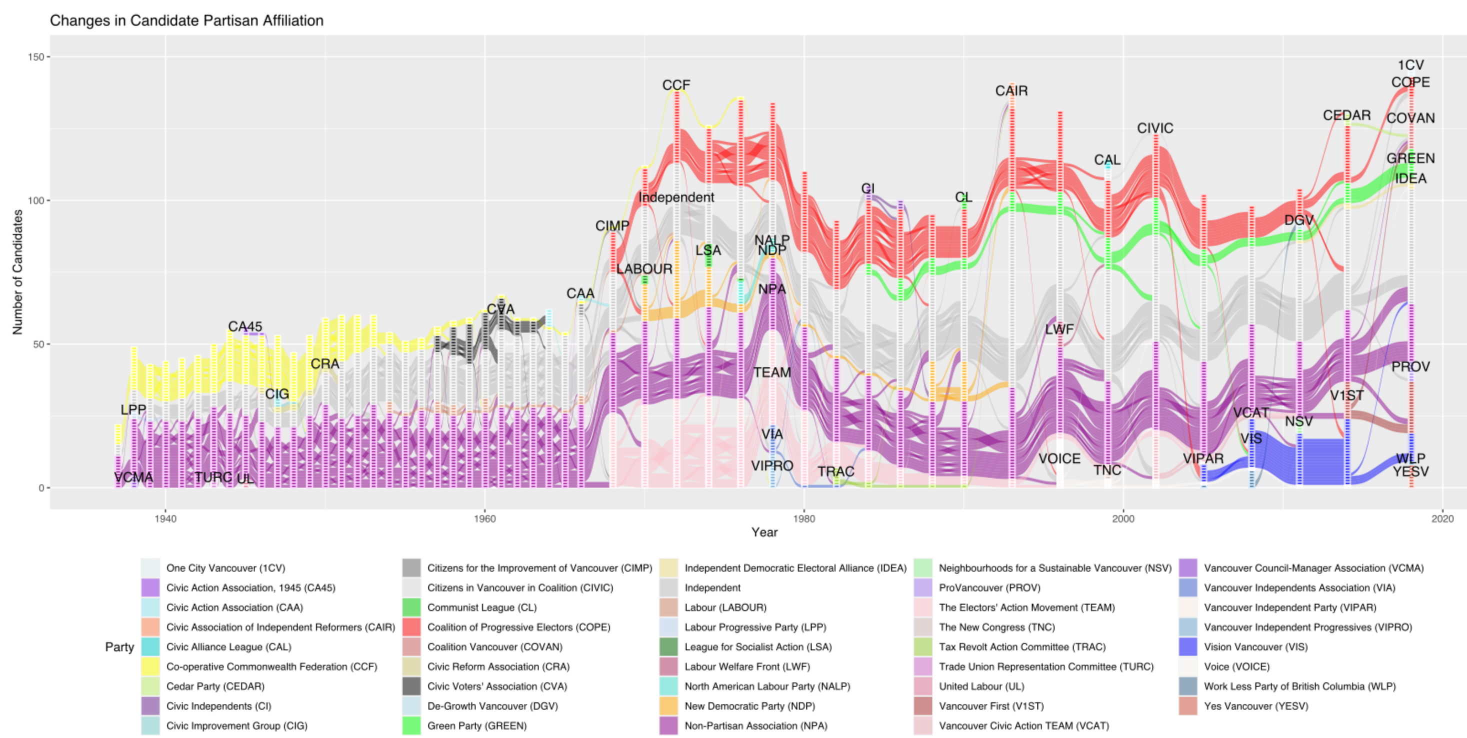


Figure 3.5: Changes in Allegiance of Individual Candidates.

Office	Measurement	First Era 1936-1966	Second Era 1968-2002	Third Era 2005-2018
City Council	N_C	5.02	7.26***	9.77
	N_V	2.97	3.64***	4.42
	N_S	1.44	2.19***	2.21
Mayor	N_C	3.50	5.69***	9.40**
	N_V	2.17	2.26	2.73
School Board	N_C	2.34	3.70***	5.84
	N_V	1.85	2.81***	4.21
	N_S	1.06	1.74***	2.64*
Park Board	N_C	2.82	5.85***	6.20
	N_V	2.30	3.19***	3.86
	N_S	1.40	1.67**	2.47*

T-test statistical significance of difference from previous era: *p<0.1; **p<0.05; ***p<0.01

Table 3.1: The Effective Number of Parties for each Era of Modern Vancouver Political History

This section further theorizes that the party system of an MNTV institution can be stabilized by a larger vote share in the concurrent election of a prominent single-winner election, which in the case of Vancouver is the election of the mayor.

3.3.1 Theories of Party Lifespans

This chapter builds on two bodies of existing academic literature. This first concerns the rise of new parties and the stability of party systems. The second and smaller body concerns the behavior of parties, candidates, and voters in MNTV systems. The current intersection of these two veins of research is limited—as far as this author is aware—to Crisp and Demirkaya (2019) who found (unsurprisingly) that MNTV elections produced more candidates than otherwise similar SMD elections of the same offices in Brazil.

The existing literature on new parties focuses on three main topics: the first defines and conceptualizes party lifespans generally and new parties specifically; the second categorizes new parties into different types; and the third explains the birth of parties. In this first topic, Pedersen (1982) provides the seminal work on party lifespans, while critiquing and building upon earlier authors such as Janda (1970) and Sartori (1976). Pedersen argues that the lifespan of a party can be conceptualized best by discrete thresholds that parties pass as they grow (or fall behind as they decay): the thresholds of declaration (i.e. publicly launching the party), authorization (i.e. getting on the ballot), representation (i.e. winning at least one seat), and relevance (a more nebulous threshold, first suggested by Sartori (1976)). He argues that, contrary to previous conceptions that parties either have a bell-shape lifespan or rise and plateau indefinitely, one party may rise, fall, then rise again numerous times. This chaotic multi-modal lifespan is seen in various parties in Vancouver, such as the New Democratic and Labour Parties, which both fell below all thresholds in the middle of their lifespans before rising once again before their eventual deaths. Figure 3.6 shows that parties in Vancouver do follow a gentle bell-curve throughout the course of their life in terms of the number of candidates within the party, while figure 3.7 shows no clear pattern in vote share through the course of an average party's life.

Pedersen's most important contributions are his warnings to the researchers of party lifespans, both against focusing only on parties that cross the higher thresholds of representation and relevance and against thinking of any party, no matter how large, as anything other than a "mortal" entity with a finite lifespan. These warn-

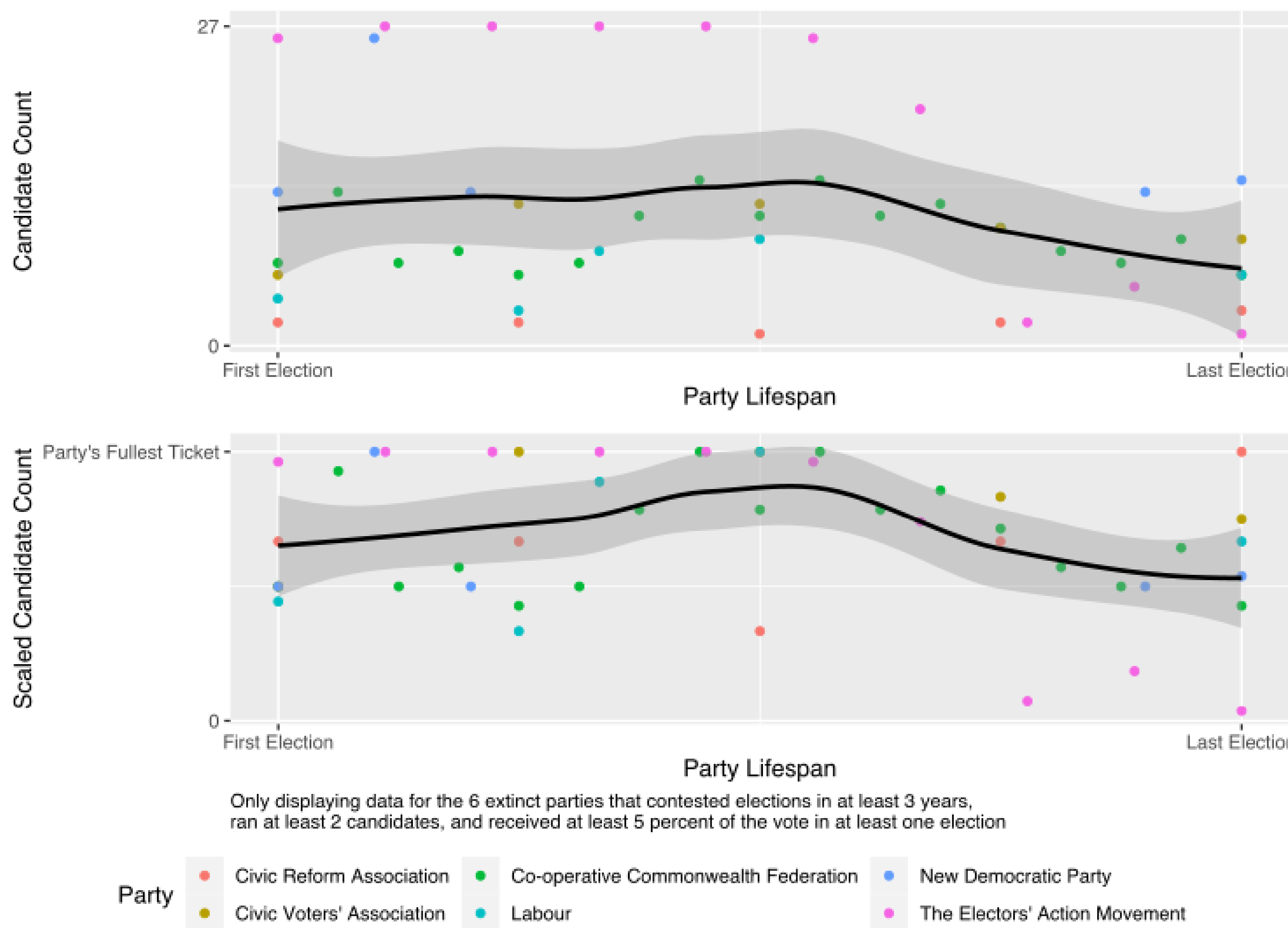


Figure 3.6: The Standardized Lifespan of a Party, Based on the Number of Candidates Nominated.

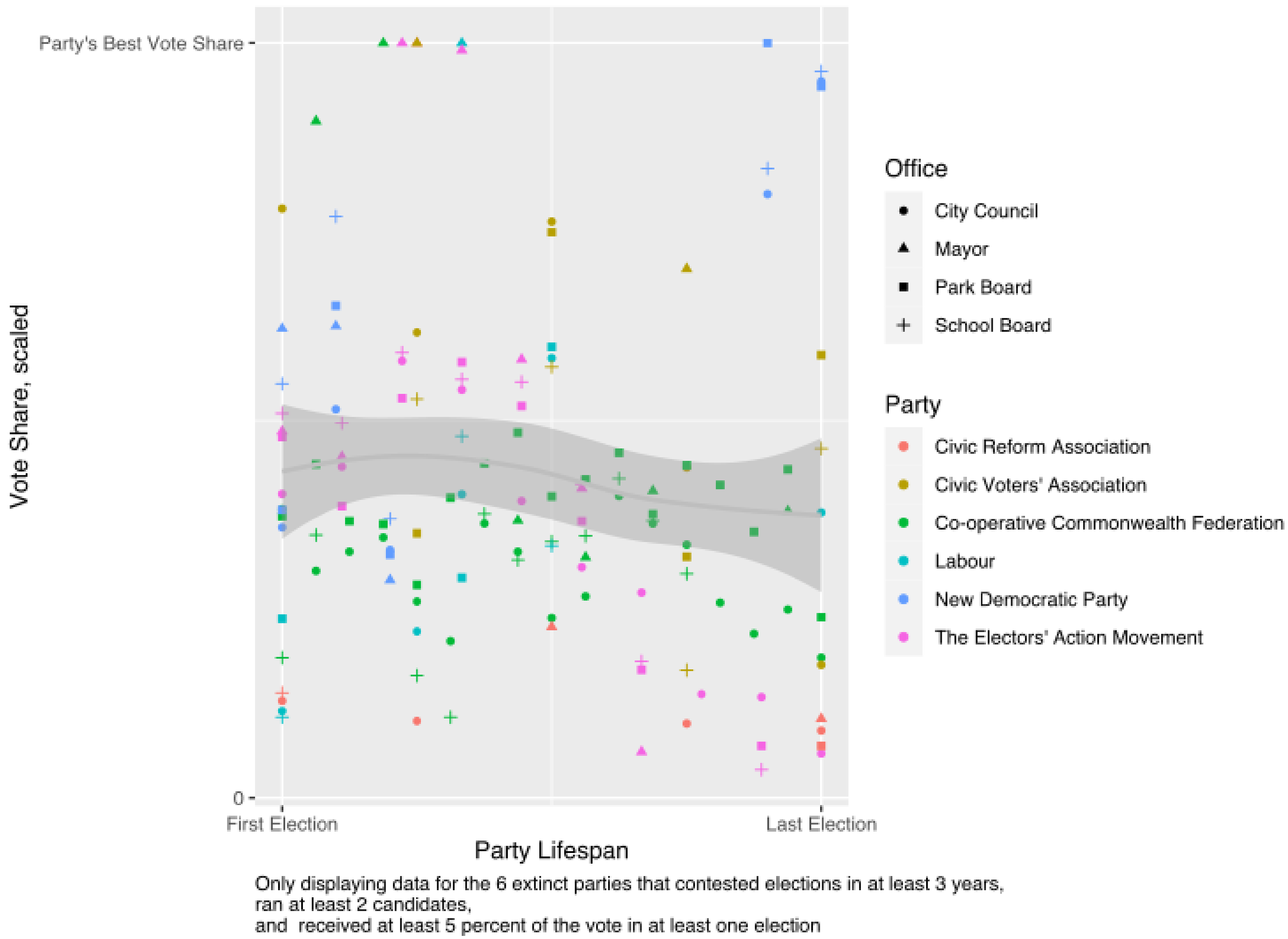


Figure 3.7: The Standardized Lifespan of a Party, Based on Vote Share

ing touch upon deeper measurement issues regarding the lifespan of parties, namely survival bias and truncation. Firstly, especially small parties that never cross the thresholds of authorization or representation, or perhaps even of declaration, may illuminate by contrast the factors that fuel a successful party yet may escape detection by researchers (a concern which is also raised by Hug, 2001, p. 15). Secondly, while all parties might be mortal, very few of the largest parties in the world's oldest and most stable democracies have lived long enough to die of old age. Therefore, it simply may be too early—perhaps by several centuries—to fully understand the lifespan of large parties.[4]

Pedersen (1982) and his contemporaries, such as Janda (1970) and Sartori (1976), and even earlier scholars on party systems and party lifespans, such as Kirchheimer (1966) and Key Jr. (1955, 1959), did not provide a strong operational definition for what constitutes and does not constitute a new party. This distinction is less clear cut than one might expect, both in Vancouver and throughout both the developing and developed world. As Krouwel and Lucardie (2008, p. 279) argue, there are four basic ways in which a new party can come into existence: first, a party can transform into a new party in name, fundamental policy positions, or both; second, two or more parties can merge together into one party; third, a new party can split from an existing party; and finally, a new party with no strong connection to an existing party can arise from the electorate. Different scholars generally count the latter two or three as valid new parties, but even the decision to include merged parties as new parties can double the count of new parties in the same geographical

[4] Some scholars have studied the life and death of young parties, which avoids the truncation problems of studying party death generally. Buelens and Hino (2008), for instance, examine the effect of government participation on the surivial of new parties. They find that participating in a governing coalition is no more or less detrimental to the electoral future of new parties compared to older established parties.

and temporal scope compared to counts that only count schisms and wholly new parties (Ignazi, 1996; Mair, 1999; Hug, 2001; Krouwel and Lucardie, 2008). One could argue that another fifth time of new party might exist in the form of "reborn" parties that cary the name and ideals of a dormant or dead party, which is seen repeatedly in Vancouver.

Barnea and Rahat (2011, p. 306) provide the first best definition of a new party that is both conceptually grounded and operationally workable. They build upon Key Jr. (1942) three-faces of the party (the party-in-the-electorate, the party-as-organization, and the party-in-government) to develop eight possible criteria of "newness":

> [1] Is the name genuinely new or does it contain an 'old' party name?. . . [2] How different is the 'new' party platform from the old party/ies platform/s?. . . [3] How different is the 'new' party electoral base from preceeding parties?. . . [4] Is the party registered as new?. . . [5] Were the party institutions separated and differentiated from those of the old party/ies?. . . [6] Does the 'new' party have new activists or did they 'immigrate' to it from the old party/ies?. . . [7] Are the top candidates new (non-incumbents)? Did most or all of them come from a single party?. . . [8] How different are the 'new' party's policies from the old party's/ies' policies?

Barnea and Rahat eliminate and elaborate upon these potential criteria, ending with a parsimonious definition of a new party as "a party that has a new label and that no more than half of its top candidates (top of candidate list or safe districts) originate from a single former party" (p. 311) under the logic that a party can change under the other six criteria of "newness" without fundamentally becoming a new party. This definition informs this chapter's own identification of new parties in Vancouver, but it is ultimately insufficient for the chaotic nature of the city's party system. Under Barnea and Rahat's definition, the two Civic Action Associations would be seen as two temporally separated instances of the same party, rather than as the two distinctly different parties they clearly are as discussed elsewhere in this chapter. More problematically, their definition does not account for certain types of breakaway parties. For instance, if twenty candidates from a party of 100 candidates form their own new party but do not recruit any additional candidates, then the 20-member breakaway party would not be counted as a new party because more than half of their candidates originate from the same original party. This might not be a problem in the examination of most party systems, but it does present a problem in the case of Vancouver and in any other system in which small breakaway parties are common. Most notably, it precludes the classification of Vision Vancouver as a new unique party, because four of its six original candidates were defectors from COPE.

While the works on identification are instrumental to this chapter's methodology, the other two topics regarding new parties–classification and explanation–provide the foundation for this chapter's theory. The final topic of new party literature–classification–manifests in two approaches. The first and most common approach is to divide and study new parties according to their ideology. Some works in this vein often focus on emergent leftist and green parties (see, for instance, Kitschelt, 1989; Inglehart and Andeweg, 1993; Müller-Rommel, 1995), while others focus on emergent right-wing parties (see, for instance, Husbands, 1988, 1992; Betz, 1990, 1994), and others on regional and ethnic parties (see, for instance: Urwin, 1983; Levi and Hechter, 1985; De Winter, 1995; Newell, 1998; Buelens and Dyck, 1998; Müller-Rommel, 1998). While this approach provides insight into the nature of parties of particular ideologies, including Vancouver's own Green Party which defies the city's recent trend of decaying parties, the division of new parties by ideology does not provide a parsimonious or strong understanding of new party entry generally. Further, focusing on a particular type of new party both introduces self-selection problems and greatly limits the number of observations in a given study (Hug, 2001,

p. 73).

The second and more useful way to classify new parties is by their strategic goals and strategic relationship to other parties. Building on earlier authors such as Rochon (1985) and Lucardie (2000), Krouwel and Lucardie (2008) provide a useful classification for new parties. They argue that there are four motivators (ideological motivations, personal ambitions, needs or frustrations, strategic or tactical considerations, and altruistic-societal goals) that create five types of new parties (p. 284):

1. prophetic parties aiming to drastically change the political and social system by propagating a coherent set of ideals and values not articulated by established parties;
2. challengers of established parties that try to maintain or renew the ideology of the latter, possibly in a more extreme or in a more moderate variety;
3. advocates of particular interests neglected (in their perception) by established parties, (elsewhere also defined as 'prolocutors');
4. reform parties that try to change or purge the political system and political culture without an explicit ideology (purifiers or pragmatic reformers);
5. idiosyncratic parties, reflecting the personal inclinations and ideological eclecticism of the founders.

They find that challengers of established parties are the most successful, at least in terms of the number of candidates elected. This categorization of new parties illuminates the rise of new parties in a given system is a complex one driven by many different types of actors with different types of motivations. As this chapter argues and finds below, the ambition and success of challengers are largely responsible for Vancouver's party system instability.

Early theories on the origins of parties can be grouped into historical-situation theories, developmental theories, and institutional theories (LaPalombara and Weiner, 1966, p. 7). Historical theories, as summarized by (LaPalombara and Weiner, 1966), argue that "historical crises not only often provide the context in which political parties first emerge but also tend to be a critical factor in determining what pattern of evolution parties later take. They are often historical turning points in political systems" (p. 14). These crises can be sudden and dramatic, such as a war, or slow and mundane, such as an increase in education rates. Parties arise as elites respond or fail to respond to these crises. These theories are often built in case studies (Chambers, 1966; Rustow, 1966; Sartori, 1966; Kirchheimer, 1966; Binder, 1966; Emerson, 1966; Grodzins, 1966) and, while illuminating, often fail to provide a unifying causal explanation of party formation across different countries beyond the fact that formation is driven by crisis.[5]

Early institutional works theorize that political parties rise from and are strongly influenced by formal governmental institutions and rules, such as legislatures and electoral systems. Institutional theories can be traced back to Duverger (1955) who theorized that many parties start as blocs within a legislature, which then organize locally to ensure their reelection. Duverger also suggested that factions may mobilize outside of the legislature as challengers to the status quo, creating "external" parties that are more centralized and ideologically coherent than their "internal" party counterparts.

More modern institutional works, among which this chapter counts itself, also account for nongovernmental institutions, such as the parties themselves, and informal institutions, such as the party system. As summarized by Barnea and Rahat (2011), contemporary institutional research on new parties can be further divided into three types: "One centered on the development and change of party models or

[5] For further critique of the use of case studies in building theories of party formation, see Hug (2001).

types[6]... Another analyzed development and change at the intra-party level[7]... And, finally, the analysis of stability, continuity and change at the party system level has become a central concern for scholars of party politics." (p. 304). This chapter builds on this final thread, building a theory concerning the stability and dynamics of the party system, driven by strategic considerations of politicians who are considering entering the system, forming a new party, or changing parties.

3.3.2 A Theory of Challengers, Parasites, and Instability

This chapter's theory can be summarized by the two following propositions. Firstly, party systems in elections in which MNTV is employed will be *destabilized* as the vote share of the largest party in the MNTV election increases. Secondly, party systems in elections in which MNTV is employed will be *stabilized* as the vote share of the largest party/candidate in a prominent concurrent single-winner election increases.

A large vote share for one party in an MNTV election destabilizes the party system through two causal mechanisms. The first, the "parasite effect," causes independents and small parties to enter the system, each attempting to draw support from the largest party. The second mechanism, the "weak challenger effect," in which new parties rise in an attempt to replace the existing minority party or parties as the main challenger to the majority party.

The parasite effect can be understood by the mathematic threshold required for an independent or small parties (hereafter referred to as "parasitic spoilers" or "parasites") to steal a seat that would otherwise be won by the largest party. As the gap

[6] See, for instance, Aldrich (1995) who argues that "the major political party is the creature of the politicians, the ambitious office seeker, and officeholder. They have created and maintained, used or abused, reformed or ignored when doing so has furthered their goals and ambitions" (p. 4). In this vein of research, the political party is wholly endogenous and can be understood by understanding the actors within it.

[7] See, for instance, Bolleyer (2013) who argues that the death of a new party is driven by its internal characteristics and structure.

between the largest and second largest parties' vote shares increases, spoilers can win a seat with both fewer votes and a smaller share of votes from the largest party (this pattern holds until the largest party is twice as big as the second-largest party, at which point the absolute number of votes required by the spoiler begins to increase while the required share of defectors from the largest party remains constant). This proposition is true for all simple plurality-based elections, both with one winner and with many, which makes sense conceptually because when the magnitude of an MNTV election is reduced to one, it collapses into a single-winner plurality election. While an equal number of defectors from the largest party is required for a spoiler to win in an MNTV election, regardless of magnitude, the degree of defection required from each defector (i.e. supporters of the parasite candidate who would otherwise support the largest party) to secure victory decreases as magnitude increases. Therefore, the parasite effect is much more powerful in MNTV elections than single-winner elections and more powerful in MNTV elections of higher magnitude than those of lower magnitude.

The mathematics of the parasite effect can be understood by considering a simple two-party single-winner election then generalizing to systems with more parties and more winners. First, for all following examples, let a "parasite" be more precisely defined as a spoiler candidate who targets and attempts to draw votes from one specific candidate—or, in the case of a single-winner election, the only candidate—of the largest party. This makes parasites a specific subtype of spoiler candidate; non-parasitic spoilers are those who change or might change the outcome of an election merely by running, but who do not target a specific candidate of the largest party. Next, for the simplest starting example, take a single-winner election in which the larger party is supported by v_1 percent of the voters, the smaller party is supported by $v_2 = 100 - v_1$ percent of the voters, $v_1 > 50$, and $v_1 > v_2$. If a parasite enters this system and attempts to take votes away from the largest party's candidate, they will successfully spoil (i.e. change the outcome) if they convince at least $200 - \frac{10000}{v_1}$ percent of the largest party's voters to defect. More relevantly, if the parasite wants to win rather than merely change the outcome, they need to win a larger minimum share (p) of the largest party's voters. This minimum share p must be large enough that the parasite's vote count (v_P) rises above both the new diminished vote share of the largest party (v_1') and unchanged vote share of the second party (v_2). Put mathematically, the parasite will win if both $v_P > v_1'$ and $v_P > v_2$ (hereafter referred to as the first and second thresholds, respectively). These two thresholds are transformed below to explore the share of the largest party's voters the parasite needs to convince to defect. Both thresholds are expressed as equalities rather than inequalities, to find the minimum (p) rate of defection that results in the parasite's victory and to analyze the relationship between changes in v_1 and changes in p. The transformations require the following equalities: $v_2 = 1 - v_1$ because there are only two parties, so any voters who do not support the largest party in the initial setup support the second-largest party; $v_P = p \times v_1$ by definition; and $v_1' = v_1 - p \times v_1$ which simply means that the largest party's diminished vote share is comprised of those voters who did not defect to the parasite.

The first threshold is transformed as follows:

$$
\begin{aligned}
v_P &= v_1' \\
p \times v_1 &= v_1 - p \times v_1 \\
2 \times p \times v_1 &= v_1 \\
p &= \frac{v_1}{2 \times v_1} \\
p &= \frac{1}{2}
\end{aligned}
$$

The second threshold is transformed as follows:

$$\begin{aligned} v_P &= v_2 \\ p \times v_1 &= 1 - v_1 \\ p &= \frac{1}{v_1} - \frac{v_1}{v_1} \\ p &= \frac{1}{v_1} - 1 \end{aligned}$$

These transformations spell out mathematically what may be intuitively obvious in this simple example. The first shows that a parasite must convince at least half of the largest party's supporters to defect. The second speaks more directly to the parasite effect; v_1's position in the denominator shows that as the vote share of the largest party increases, the parasite requires fewer defectors to beat the smaller of the two parties.[8]

To continue forward, I will first generalize these statements to account for additional parties and then for additional seats. This will ultimately show that the parasite should be stronger as magnitude increases. First, to account for all parties, let v_{3+} be the sum of all votes shares for parties that are neither the largest nor second-largest parties. Under this new construction, $v_2 = 1 - v_1 - v_{3+}$. This does not affect the first threshold, because any parasite must still take at least half of the largest party's votes. Adding additional parties does, however, affect the second

[8] This can be seen analytically by taking the derivative of $p = \frac{1}{v_1} - 1$ with respect to v_1 to show that p decreases for all values of v_1: $\frac{dp}{dv_1} = \frac{-1}{v_1^2}$.

threshold as follows:

$$v_P = v_2$$
$$p \times v_1 = 100 - v_1 - v_{3+}$$
$$p = \frac{1}{v_1} - \frac{v_1}{v_1} - \frac{v_{3+}}{v_1}$$
$$p = \frac{1}{v_1} - \frac{v_{3+}}{v_1} - 1$$

This generalized second threshold shows that, even if additional parties are added, an increase in the largest party's vote share will still decrease the share of defectors from the largest party that the parasite requires to be elected (p).[9] Importantly, it also shows that p also decreases as v_{3+} increases for any given value of v_1.[10] For practical purposes, this simply means that as the number of votes cast for especially small parties increases, a potential parasite can beat the second largest party with an increasingly smaller percent of the voters from the largest party.

To generalize these thresholds to multiple-winner elections, let $c_{i,j}$ be the j^{th} candidiate of the i^{th} party where $i = 1$ represents the largest party, $i = 2$ represents the second-largest party, and so forth. Further, assume that each party has an initial given share of the vote that is divided equally between each of its candidates, such that the vote count of each candidate, $v_{i,j}$, is equal to the average vote count of the given party, v_i. Next, let the parasite candidate target one specific candidate of the largest party, $c_{1,x}$ with initial vote count $v_{1,x}$. To win, the parasite still needs to beat every candidate of the second-largest party (i.e $v_P > v_2$, as in single-winner elections),

[9] This can be seen analytically by taking the derivative of $p = \frac{1}{v_1} - \frac{v_{3+}}{v_1} - 1$ with respect to v_1 to show that p decreases for all values of v_1 given all possible values of v_{3+}: $\frac{dp}{dv_1} = \frac{v_{3+}-1}{v_1^2}$; given $0 < v_{3+} < 1$, $\frac{0-1}{v_1^2} < \frac{dp}{dv_1} < \frac{1-1}{v_1^2}$; $\frac{-1}{v_1^2} < \frac{dp}{dv_1} < \frac{0}{v_1^2}$; and, finally, $\frac{-1}{v_1^2} < \frac{dp}{dv_1} < 0$. Therefore, $\frac{dp}{dv_1}$ is always negative.

[10] This can be seen analytically by taking the derivative of $p = \frac{1}{v_1} - \frac{v_{3+}}{v_1} - 1$ with respect to v_{3+}: $\frac{dp}{dv_{3+}} = \frac{-1}{v_1}$.

but only needs to beat one candidate $c_{1,x}$ of the largest party (i.e. $v_P > v'_{1,x}$). The threshold required for the parasite to beat $c_{1,x}$ can be transformed similarly as the first threshold in single-winner elections, but with $v'_{1,x}$ substituted for v'_1, because each vote taken from $c_{1,x}$ increases the parasite's vote count by one. Therefore, the parasite requires the same share of the majority party's voters to be elected, even when more than one seat is contested.

While the parasite candidate requires the same percent of defectors from the largest party, regardless of magnitude, the *degree* of defection required from each defector decreases as magnitude increases. When magnitude is one, the parasite can only win if each defector casts their entire vote for the parasite instead of for the largest party. As magnitude increases, the parasite only requires each defector to cast a fraction of their total vote for the parasite instead of for the largest party. Specifically, the parasite requires $\frac{1}{m}$ or one "m^{th}" of each defector's vote, where m is the district magnitude. One could say that the *quantitative* threshold (p) for a parasite's victory is constant across magnitude for a given distribution of votes between parties, but the *qualitative* threshold (q) decreases as magnitude increases, where q is the share of each defector's vote required by the parasite to win. As magnitude increases to infinity, q decreases to zero:

$$\lim_{m \to \infty} q = \lim_{m \to \infty} \frac{1}{m} = 0$$

Thus far, the parasite effect has described the viability of an individual spoiler candidate to win by targeting one specific candidate of the largest party, with increasing ease as magnitude increases. As magnitude increases, it also becomes increasingly viable for multiple independent candidates to each target specific and unique candidates of the majority party as parallel parasites. Parallel parasites become increasingly viable as magnitude increases; as magnitude increases to infinity, the burden on the supporters of the majority party between one and multiple parasites becomes equally insignificant.[11] Taken one step further, it would be in the interest of the parallel parasites to coordinate their efforts to ensure that each targets a different majority party candidate or that, as a group, they together target a group of majority party candidates equal in number to the number of parasites. Therefore, if the magnitude increases large enough, the parasite effect should be expected to produce parties—albeit ones without a full slate—rather than just independent candidates.

Both the parasite and weak challenger effects describe phenomenon wherein new parties and candidates rise in response to larger vote shares of the system's largest party. While the parasite effect predicts the rise of small parties and independents that precisely target specific candidates of the majority party, the weak challenger effect predicts the rise of fuller-slate parties that target and attempt to supplant existing minority parties. The logic of the weak challenger effect is less mathematical and more practical than the parasitic effect. Simply put, as the size of the largest party increases, then the maximum size (and, by extension, the average and expected sizes) of all other parties decreases. If the existing challengers to the majority party are weak and small, then there is more incentive and opportunity for opponents of the majority party to enter the system as a new challenging party. For example, if the vote distribution between two parties is 60/40, then there is little incentive for a new challenger to enter the race, because if the new party wishes to successfully establish itself as the predominant challenger to the larger party, then it would need to bring

[11] For example, if a voter has 100 votes, there is little difference between two parasite candidates each asking for one vote from a supporter of the majority party. Mathematically, for any k number of parasite candidates, the share of the majority party's supporters' votes required to elect all parasites diminishes to zero as magnitude increases to infinity:

$$\lim_{m\to\infty} q = \lim_{m\to\infty} \frac{k}{m} = 0$$

in or syphon off enough voters to supplant the existing party with 40 percent of the vote. Conversely, if the vote distribution between two parties is 90/10 or between three parties is 80/10/10, then a new challenging party only needs to receive ten percent of the vote to position themselves as the predominant challenger for the new party.

The final theorized effect suggests that the party system of an MNTV election can be stabilized by a prominent single-winner election that is highly proximate to the MNTV election in both time and jurisdiction. Specifically, I suggest that strong domination by one party in the concurrent single-winner election decreases the number of new parties and independents, including those who would challenge the majority party and those who would challenge the minority parties. The dominance of a single party in the single-winner election signals to potential parasites that the major party is strong and may be difficult to divide, while it signals to potential new opposition parties that supplanting the existing opposition parties will only lead to an inevitable defeat against the majority party. This is more or less a highly localized application of a presidentialization argument, that the presence of a popularly elected executive will influence the nature and election of the legislature, especially if the elections are contemporaneous. In the case of Vancouver, specifically, the prominent concurrent single-winner election is that of the city's mayor, which is the central contest in the city's politics.

In conclusion, entry in MNTV elections can be summarized as follows. As the highest vote share in the election of the given office increases, more parties and independents enter the system. And as the highest vote share in the election of the mayor increases, fewer parties and independents enter the system. These can be expressed as the following tractable hypotheses:

Hypothesis One: *Higher maximum vote shares in a given MNTV election will* *cause instability in the party and candidate systems in the following election of the same body.*

Hypothesis Two: *Higher maximum vote shares in the mayoral election will decrease instability in the party and candidate systems of the MNTV bodies.*

These two hypotheses are further divided into four more specific hypotheses (1a-1d and 2a-2d), each measuring a different element of instability:

a More new candidates,

b More independents,

c More new parties, and

d More new candidates in new parties

3.4 Data, Methods, and Findings

The basic empirical plan of this section is to study one unit (i.e. Vancouver) over time, rather than comparing multiple units at the same time (or close to the same time). This contrasts with most other works on new parties, but it avoids several problems associated with cross-sectional studies of this particular topic, namely the difficulty of creating a unifying measurement of new parties between countries in which parties may look very different. Additionally, it can be difficult to control for important yet unobserved independent variables related to party formation between units (e.g. ballot-access rules or societal cleavages), which can lead to omitted variable bias. A time-series analysis of one unit, however, controls for unobserved independent variables inherently, so long as the variables are relatively static over time.[12] Focusing on one unit further allows for parsimonious measurement rules

[12] Even non-static unobserved variables can be automatically accounted for in a time-series analysis if they change at a relatively constant rate, as their effect is captured by estimated trend.

for the various variables, especially if each variable is drawn from a common datasource.[13]

3.4.1 Modeling

This chapter uses three main binary categorical dependent variables to measure candidate entry and party formation. The four variables captures whether or not each candidate falls into the following categories. The first three of these are: candidates running for the first time ("new candidates"), candidates running in new parties ("new party candidates"), and candidates who are both running for the first time and running in a new party ("new candidates in new parties"). All operationalized forms of these variables only measure the number (or relative number) of candidates in each category but not the number of candidates within a specific party. This is driven by the theoretical justification that the underlying effects are driving candidates into politics and into new parties but not into specific new parties. Therefore, for instance, the rise of two new two-member parties is functionally equivalent to the rise of one new four-member party. This is similar to the methodology employed by Rosenstone, Behr and Lazarus (1984) and later justified in more depth by Hug (2001).

While these categories are generally meant to directly explore party formation, none directly explore party death. From both a qualitative and quantitative perspective, the birth of a party is usually a much more clear and tractable phenomenon than its death. This is especially true when examining the quantity of candidates in a party throughout its life. While new parties generally do not nominate a *full* slate of candidates in their first election, parties of any future consequence do usually nominate a *large* slate of candidates in their first election. Therefore, in both the

[13] For a further discussion of the pros and cons of cross-sectional and time-series analysis in the context of party formation, see Hug (2001, pp. 75-76).

data and casual observation the formation of a party is usually observable as a large slate of candidates running under a new partisan affiliation for the first time, thus providing a clear and definitive election year that can be identified as its year of genesis. Conversely, party deaths vary greatly and can be much more gradual than party births. While one might definitively say that a given year was the final year that any candidate ran under a given party's banner, the practical death of the party may have been much earlier. The final official years of political party may be marked by one or two failing candidates stubbornly clinging to a party that once elected full slates to office. The quick births and slow deaths of parties is illustrated in figure 3.6. Overall, the variability and general lengthiness of party deaths make questions regarding the end of party lives difficult to explore directly. Nonetheless, a limited degree of understanding about the final years of a party's life can be inferred indirectly, by exploring retirements within and partisan defections from failing parties, which is beyond the scope of this chapter.

This also includes supplemental models that explore four additional binary categorical dependent variables: candidates running under a different partisan affiliation than their previous campaign ("new allegiances"), candidates running as independents who did not run as independents in their previous campaign ("new independents" when including independents who are also new candidates and "new independents & returning candidates" when excluding new candidates, thereby only including independents who recently defected from a party), and candidates running for office for the final time ("final elections").

The first major issue when defining these categories is how to treat independent candidates. Conceptually, an independent can be considered as either the singular member of their own party or as a candidate who has opted out of the party system entirely. Under the former conceptualization, variables that affect party formation should have a similar effect on the number of independents, because if the number of multiple-member parties increases then so should the number of single-member parties. Under the latter conceptualization, however, an increase in the causes of party formation may be expected to decrease the number of independents under the logic that an increase in the number of parties is either associated with or partially caused by independents forming and joining new parties. This chapter makes two steps to account for both conceptualizations of independent candidates in case either is fundamentally flawed. Firstly, this chapter introduces a fourth binary categorical dependent variable which measures whether or not a candidate is independent. This variable isolates the effect of the independent variables upon candidate independence and thereby helps contextualize the models for the other three dependent variables. The inclusion of independents as a dependent variable in the study of new parties is further justified in its own right, because as Weeks (2008) argues, "a study of independents can teach us a lot about why new parties emerge, since they are the next stage down from parties on the evolutionary chain of party formation" (p. 154). Secondly, all models concerning the partisan dependent variable (i.e. "new party candidates" and "new candidates in new parties") are duplicated so that one includes independents in the dataset and the other does not. Models that include independents treat each independent as a member of their own unique party.

There are two conceptually and statistically different paths to gaining traction on these categories as dependent variables, thereby explaining the underlying elements of party system change they represent. The first path is to predict the *probability* of a given candidate being in a given category. The second path is to predict the *quantity* of candidates within a given category. This chapter employs a set of "candidate-centric" models to explore the former and a set of "election-centric" models explore the latter. The results of the candidate-centric models show the effect of the independent variables on the *relative* number of candidates in a given category, while the election-centric models show their effect of the *absolute* number of candidates in a given category.

The candidate-centric models explore behavior at the individual candidate/election year level, i.e. the unit of analysis is a candidate in a given election cycle. Each candidate-centric model uses logistic regression; their dependent variables are the probability of a candidate being in their respective category, such as being a new candidate or a member of a new party. These models can account for the attributes of individual candidates and their parties, such as how long the candidate has been in office, the strength of their party in the previous election, and the candidate's relative performance within the party in the previous election. For the sake of comparison and robustness, for each candidate-centric model that includes variables related to political strength of the candidate and their party, a second model is included without those variables.

The second set of models explores the aggregate behavior of political actors in a given election year, i.e. the unit of analysis is the election year. These models use Poisson regression to predict the number of candidates in their respective category in a given year.[14] For example, some of these year-centric models examine the number of independents in a given year, while others examine the number of candidates running under a new party. These models cannot account for attributes of individual candidates, but they are fundamentally more robust against unobserved heterogeneity between years because each year only accounts for a single observation.

While the candidate-centric models can account for candidate-level attributes, they are vulnerable to unobserved heterogeneity between election years because mul-

[14] For a similar application and justification of Poisson models in the study of new parties, see Hug (2001).

tiple candidates are observed each year. This heterogeneity can be generally addressed by including year-level variables, such as district magnitude and the lagged max vote shares, and party-year-level variables, such as the vote share of the given candidate's party in the pervious year. Many of these year-level and party-year-level variables are of substantive interest to this chapter and should be included in any case. Any remaining unobserved heterogeneity between election years can be addressed by using either fixed-effects or random-effects for each election year. The use of fixed-effects is suboptimal due to the large number of election years. Further, random-effects are generally preferred to fixed-effects in any case, if the effect in question is not of substantive value. Unobserved heterogeneity between years is also a potential problem for the election-centric models, albeit less so than in the candidate-centric models because the election-centric models only have three observations per year (one for each office) while the candidate-centric models have one observation per candidate. Therefore, random-effects for each year are used in all models to capture all remaining unobserved heterogeneity that is not captured by the year-level variables.

These models are duplicated, such that one set only contains candidates running for mayor while the other set contains all other candidates, so that the effects of the independent variables can be compared between the MNTV elections and the plurality mayoral elections, thereby allowing the hypotheses to be tested. The alternative to using two sets of models would be to use interaction terms within one model to capture the difference in effects between mayoral races and other races. This alternative, however, is infeasible because many important control variables for the MNTV elections, such as magnitude, are constants for all mayoral elections, which leads to perfect multicollinearity if all candidates are grouped into one model.

3.4.2 Data, Measurement, and Independent Variables

Most data used in this chapter is taken from Madden (2003), who privately compiled historical election results from Vancouver and donated his work to the city's archives. I spot-checked his data against old issues of *The Vancouver Sun*, which generally reported election results shortly after each election. It was impossible, however, to build a dataset using the *The Vancouver Sun* alone due to critical gaps in the paper's archives. Election data from 1996 onward was obtained directly from Vancouver's election online archives.

Two main independent variables of interest are included in all models. The first is "Lagged Maximum Vote Share, given institution" which simply measures the highest vote share won by any party for the given institution in the previous election. The second is "Lagged Maximum Vote Share, mayoral election", which is the highest vote share won by any candidate in the previous mayoral election. Note that these two variables are the same in models regarding mayoral races.

Where possible, the candidate-centric models include a slate of three control variables pertaining to the recent political strength of the candidate and the candidate's party in the previous election year. The first two of these variables are the candidate's party's vote shares in the previous elections of both the given board and the mayoral race. For periods during which mayoral elections were held every other election year, data from a party's previous mayoral contest is drawn from two election years prior. A third variable, "Relative Candidate Strength" measures a candidate's performance *within* their party in the previous election of the given institution, which is measured using data from the previous election of the given office as follows, $\frac{\text{the candidate's vote count} - \frac{\text{the party's vote count}}{\text{the party's candidate count}}}{\text{the party's vote count}}$, which scales the variables across parties so that a relatively strong candidate within a small party has a similar score to a relatively strong candidate in a large party. This third variable allows the models to determine and control for whether or not candidates that are more or less popular than their comrades are more or less likely to engage in any of the actions captured by the dependent variables (e.g. perhaps a candidate who outperforms their party average will be more inclined to shed their partisan identification and seek office as an independent). Note that "Relative Candidate Strength" will be always be zero for candidates who were independents in the previous campaign. By construction, "Relative Candidate Strength" cannot be included for mayoral races because a party's singular mayoral candidate will always match the party's average for the mayoral race. When data is missing for these three variables, because either the candidate or party is new, the value of these lagged variables default to zero. For the lagged party vote share variables, defaulting to zero reflects that a party cannot receive votes before it exists. For the candidate's relative strength variable, defaulting to zero reflects that a new candidate is necessarily a candidate who neither over-pefomred nor underperformed relative to other candidates. None of these five variables are used in models exclusively concerning new candidates, because their values would all default to zero. As mentioned in the modeling subsection, for every candidate-centric model that includes these five variables, there is a second otherwise identical model without these five variables.

"Lagged Final Candidacies, given institution," which is simply a count of candidates from the previous election of the given institution who have no further candidacies in the dataset, is one of the most important and powerful control variables. This variable assumes that a candidate's exit from city politics both creates an opening for more candidates and introduces a degree of instability to the party system. "Lagged Final Candidacies, given institution" is used in every model except for those in which the number of final candidacies is itself the dependent variable.

"Lagged Final Elections, given institution" is used in every model except those that are used to predict final elections in a given year. Supplemental models predicting candidate exit use "First Elections, given institution," which counts the number of new candidates in the given election of the given institution, as a control variable, under the inverted logic that new candidates crowd out veteran candidates and cause retirements.

All models include year as a continuous control variable. This allows the models to control for linear trends in the data. For the candidate-centric models, this trend is included along with random-effects, which accounts both for a trend and random deviations away from the trend each year.

The magnitude of the given institution is included as a control variable in each model. This accounts for the possibility that a larger number of contestable seats would cause a larger quantity within most candidate categories of interest. For models regarding mayoral elections, the magnitude of the City Council—the most prominent multiple-member body—is included.

Finally, in the candidate-centric models, the number of years that a candidate has been in city politics is included as a control variable. This variable is valued at zero for a candidate's first election, then increases by one for each year (not election year) thereafter. This accounts for the possibility that veteran politicians may be more or less likely than their newer counterparts to conduct any of the underlying actions that are captured by the dependent variables (e.g. switch parties, become an independent, or retire).

It should be noted that each numeric independent variable in all models are rescaled by subtracting its mean and dividing by its standard deviation to aid convergence of the random-effects model.

3.4.3 Results: Patterns in Vancouver's Chaos

The results from the models generally support the hypotheses that an increase in the highest party vote share within a given MNTV election destabilizes the party system in the following election cycle by increasing the absolute and relative numbers of new candidates entering the system, independents, candidates entering the system to run in new parties, and candidates running in new parties. With the exception of three, all candidate-centric models exploring MNTV elections are significant at the .01 level (tables 3.2 ans 3.3). The election-centric models generally support the findings of the candidate-centric models (table 3.4). The effect of the lagged maximum party vote share in a given office is positive and significant at the .01 level for all models, except for model 10, which predicts the number of new candidates. The relevant marginal effects are illustrated in figure 3.8 and the accuracy of the predictions is shown in figure 3.9.

The models also show weaker, yet generally supportive results, of the hypotheses that the lagged maximum vote share of the mayoral race stabilizes the party system. The candidate-centric models show that the probability of a given candidate being in a new party or being a new candidate in a new party decreases as "Lagged Maximum Vote Share, mayoral election" does indeed decrease as hypothesized, whether or not independents are accounted for, but most of the results are only significant at the .1 level. Model 8, which does not account for independents but does account for the performance of the candidate and party, does not find any significant effect of the "Lagged Maximum Vote Share, mayoral election" on the probability of a candidate being in a new party. Further, the candidate-centric models show no relationship between the "Lagged Maximum Vote Share, mayoral election" and the probability of a candidate being a new candidate or independent. The election-centric results similarly show no relationship between the "Lagged Maximum Vote Share, mayoral

	Unit of Analysis: the Candidacy of a Given Person in a Given Year *Dependent Variable: the Probability of a Candidate Being in a Given Category* *Distribution: Logistic* *Random Effects Grouping: Year*		
Dependent Variable Category:	—New Candidate —	—	Independent —
Independents Included:	yes	yes	yes
model:	(1)	(2)	(3)
Lagged Max Vote Share in Given Board	0.123** (0.062)	0.188** (0.078)	0.235*** (0.076)
Lagged Max Vote Share in Mayoral Race	−0.029 (0.051)	−0.011 (0.071)	−0.027 (0.068)
Lagged Final Elections in Given Board	0.191** (0.083)	−0.080 (0.104)	−0.084 (0.100)
Lagged Party Vote Share in Given Board		−0.776*** (0.119)	
Lagged Party Vote Share in Mayoral Race		−0.029 (0.116)	
Lagged Candidate Performance within their Party and Given Board		−0.121* (0.065)	
Year	0.100 (0.097)	0.591*** (0.142)	0.629*** (0.136)
Magnitude	0.239** (0.097)	−0.378*** (0.131)	−0.477*** (0.125)
Park Board Candidate (dummy variable)	0.495*** (0.122)	−0.817*** (0.148)	−0.893*** (0.144)
School Board Candidate (dummy variable)	0.316*** (0.116)	−1.167*** (0.156)	−1.304*** (0.152)
‘Years in Politics‘		0.075 (0.049)	−0.151*** (0.048)
Constant	−0.363*** (0.076)	−0.705*** (0.100)	−0.520*** (0.094)
Observations	2,867	2,867	2,867
Log Likelihood	−1,916.076	−1,480.606	−1,560.555
Akaike Inf. Crit.	3,850.151	2,987.211	3,141.109
Bayesian Inf. Crit.	3,903.800	3,064.705	3,200.720
Note:			*p<0.1; **p<0.05; ***p<0.01

Table 3.2: Candidate-Centric Results for all MNTV Elections, part 1

	Unit of Analysis: the Candidacy of a Given Person in a Given Year					
	Dependent Variable: the Probability of a Candidate Being in a Given Category					
	Distribution: Logistic					
	Random Effects Grouping: Year					
Dependent Variable Category:	—New Candidate in New Party —		—	Candidate in New Party		—
Independents Included:	yes	no	yes	yes	no	no
model:	(4)	(5)	(6)	(7)	(8)	(9)
Lagged Max Vote Share	0.419^{***}	0.667^{***}	0.315^{***}	0.340^{***}	0.223	0.304^{*}
in Given Board	(0.086)	(0.193)	(0.087)	(0.085)	(0.187)	(0.175)
Lagged Max Vote Share	-0.147^{*}	-0.530^{*}	-0.157^{*}	-0.160^{**}	−0.553	-0.537^{*}
in Mayoral Race	(0.078)	(0.280)	(0.082)	(0.079)	(0.340)	(0.319)
Lagged Final Elections	0.081	0.172	0.122	0.093	0.037	0.006
in Given Board	(0.105)	(0.219)	(0.111)	(0.107)	(0.214)	(0.207)
Lagged Party Vote Share			-0.877^{***}		-1.004^{***}	
in Given Board			(0.147)		(0.265)	
Lagged Party Vote Share			0.158		−0.207	
in Mayoral Race			(0.134)		(0.272)	
Lagged Candidate Performance			0.034		−0.002	
within their Party			(0.060)		(0.083)	
and Given Board						
Year	0.366^{**}	0.482	0.305^{*}	0.347^{**}	0.751	0.706
	(0.144)	(0.428)	(0.161)	(0.153)	(0.510)	(0.474)
Magnitude	0.191	0.497	0.121	0.043	0.292	0.237
	(0.134)	(0.363)	(0.148)	(0.141)	(0.372)	(0.354)
Park Board Candidate	-0.369^{**}	−0.293	-0.518^{***}	-0.609^{***}	-0.804^{**}	-0.812^{**}
(dummy variable)	(0.162)	(0.366)	(0.164)	(0.160)	(0.357)	(0.344)
School Board Candidate	-1.037^{***}	-1.258^{***}	-1.203^{***}	-1.317^{***}	-1.511^{***}	-1.580^{***}
(dummy variable)	(0.168)	(0.349)	(0.168)	(0.163)	(0.333)	(0.320)
‘Years in Politics‘			-0.540^{***}	-0.868^{***}	0.004	-0.387^{***}
			(0.074)	(0.080)	(0.087)	(0.090)
Constant	-1.153^{***}	-2.984^{***}	-0.954^{***}	-0.813^{***}	-2.897^{***}	-2.591^{***}
	(0.103)	(0.373)	(0.115)	(0.108)	(0.448)	(0.418)
Observations	2,867	2,120	2,867	2,867	2,120	2,120
Log Likelihood	−1,357.481	−544.228	−1,373.613	−1,431.704	−581.003	−639.344
Akaike Inf. Crit.	2,732.961	1,106.456	2,773.225	2,883.408	1,188.006	1,298.687
Bayesian Inf. Crit.	2,786.611	1,157.388	2,850.719	2,943.019	1,261.575	1,355.279
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 3.3: Candidate-Centric Results for all MNTV Elections, part 2

	Unit of Analysis: the Election of a Given Office in a Given Year					
	Dependent Variable: the Number of Candidates in a Given Category					
	Distribution: Poisson					
	Random Effects Grouping: Year					
Dependent Variable Category:	New Candidates	Independent Candidates	—New Candidates in New Parties —		—Candidates in New Parties —	
Independents Counted:	yes	yes	yes	no	yes	no
model:	(10)	(11)	(12)	(13)	(14)	(15)
Lagged Max Vote Share	0.050	0.189^{***}	0.362^{***}	0.733^{***}	0.282^{***}	0.407^{**}
in Given Institution	(0.058)	(0.072)	(0.090)	(0.204)	(0.081)	(0.166)
Lagged Max Vote Share	−0.044	−0.044	-0.162^{*}	-0.631^{*}	-0.137^{*}	-0.621^{*}
in Mayoral Race	(0.055)	(0.062)	(0.088)	(0.339)	(0.080)	(0.370)
Lagged Final Elections	0.136^{***}	0.032	0.104	0.096	0.124^{*}	0.003
in Given Institution	(0.050)	(0.063)	(0.074)	(0.158)	(0.068)	(0.142)
Magnitude	0.483^{***}	0.034	0.463^{***}	0.805^{**}	0.355^{***}	0.455
in Given Institution	(0.089)	(0.103)	(0.133)	(0.365)	(0.124)	(0.337)
Park Board Race	0.028	-0.819^{***}	-0.541^{***}	−0.508	-0.580^{***}	-0.879^{***}
(dummy variable)	(0.099)	(0.115)	(0.149)	(0.338)	(0.134)	(0.308)
School Board Race	-0.184^{**}	-1.343^{***}	-1.198^{***}	-1.387^{***}	-1.199^{***}	-1.490^{***}
(dummy variable)	(0.092)	(0.129)	(0.156)	(0.323)	(0.140)	(0.283)
Year	0.182^{**}	0.499^{***}	0.442^{***}	0.720^{*}	0.390^{***}	0.982^{**}
	(0.093)	(0.109)	(0.141)	(0.435)	(0.133)	(0.455)
Constant	1.891^{***}	2.033^{***}	1.384^{***}	-0.930^{**}	1.692^{***}	−0.607
	(0.081)	(0.083)	(0.116)	(0.433)	(0.105)	(0.459)
Observations	147	147	147	147	147	147
Log Likelihood	−372.265	−317.605	−287.555	−164.192	−317.355	−194.267
Akaike Inf. Crit.	762.530	653.209	593.111	346.384	652.709	406.533
Bayesian Inf. Crit.	789.444	680.123	620.025	373.298	679.623	433.447
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 3.4: Election-Centric Results for all MNTV Elections

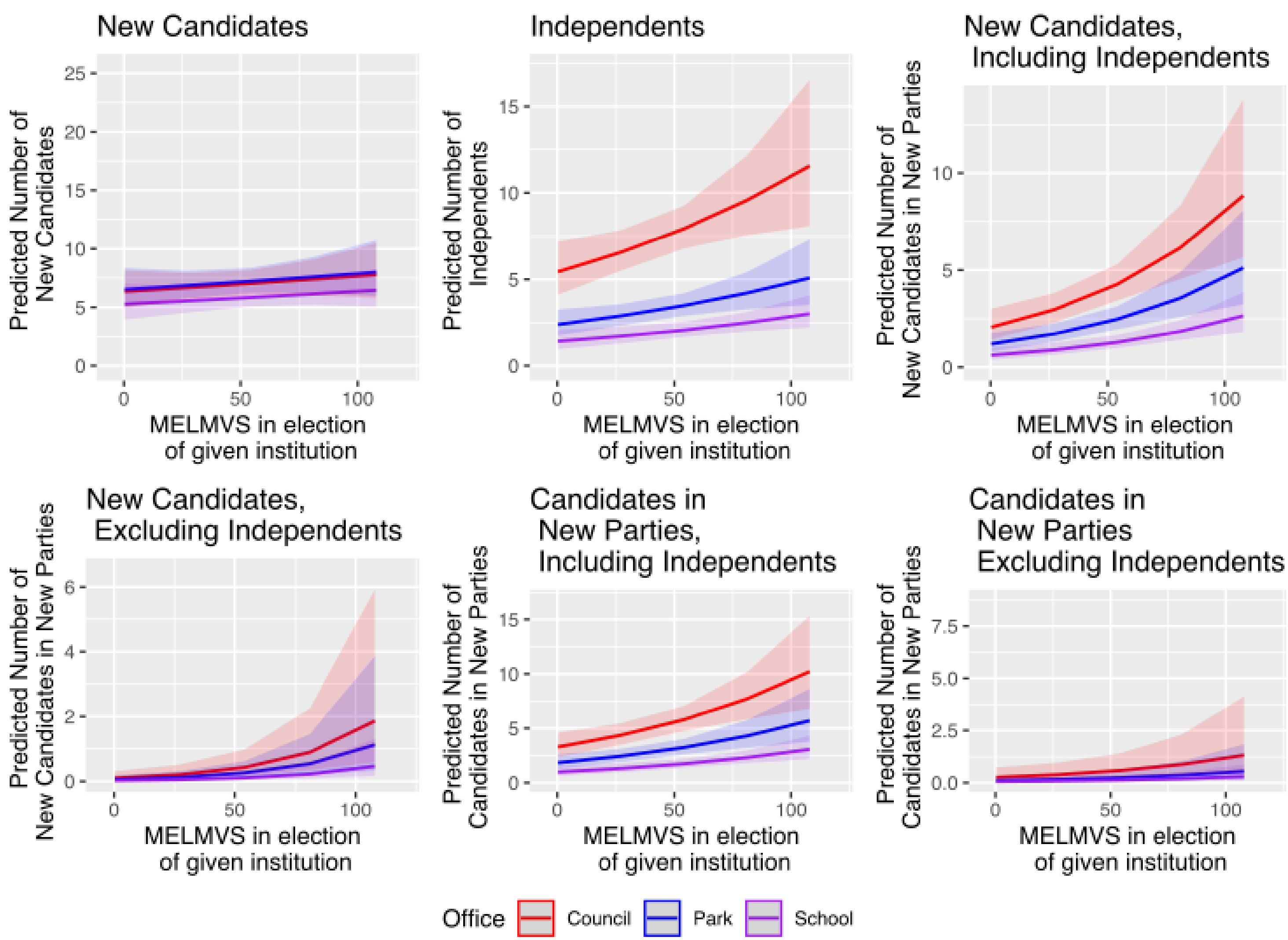


Figure 3.8: The Marginal Effect of the Lagged Maximum Vote Share (MELMVS) in the Election of a Given Institution on the Dependent Variables, in the MNTV elections.

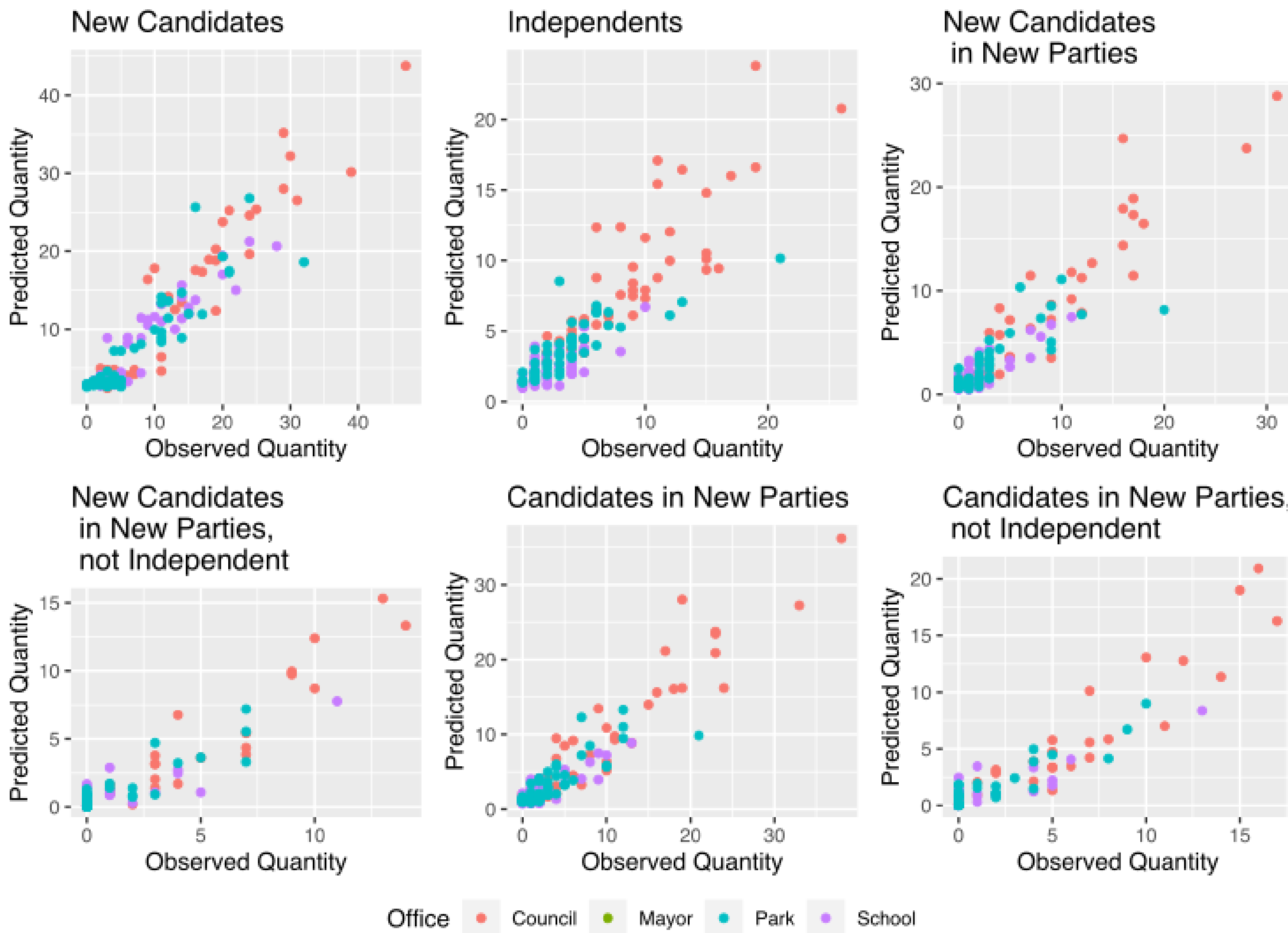


Figure 3.9: The Observed vs. Predicted Quantities of the Dependent Variables.

election" and the absolute number of new candidates and independents, but do consistently show a significant relationship between the independent variable and the number of candidates in parties and new candidates in new parties.

In order to reject the null hypotheses, the results must not only be significant for the MNTV elections, they must be insignificant for the mayoral elections. Tables 3.5 and 3.6 show that "Lagged Maximum Vote Share, mayoral election" is insignificant in every candidate-centric and election-centric model concerning the main dependent variables. The marginal effects (or lack thereof) are shown in figure 3.10. This shows that the MNTV elections are not only acting as predicted, the effects are unique to multiple-winner plurality elections and not to single-winner elections.

3.5 Conclusion

This chapter has shown that the simple concepts of both a spoiler candidate and the spoiler effect in a single-winner plurality election are specific instances of a more generalizable and complex spoiler phenomenon in plurality elections with any number of winners. The viability of spoiler candidates and the strength of the spoiler effect driven by their success increases with both magnitude and the relative size of the largest party. Moreover, however, the chapter has demonstrated that potential candidates and political entrepreneurs are capable of understanding the complex and sometimes counterintuitive opportunities in uncommon and understudied political systems such as MNTV. Taken in combination with chapter 2, this suggests that both the partisan stability of New Hampshire and the partisan instability of Vancouver are driven by rational actors operating under the same set of rules. This raises the question that will be revisited in this dissertation's conclusion, chapter 5: why does MNTV yield party systems with such varying levels of stability?

	Unit of Analysis: the Candidacy of a Given Person in a Given Year Dependent Variable: the Probability of a Candidate Being in a Given Category Distribution: Logistic Random Effects Grouping: Year								
	— New Candidate —	— Independent —		— New Candidate in New Party —		— Candidate in New Party —			
Independents Included:	yes	yes	yes	yes	no	yes	yes	no	no
model:	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)
Lagged Max Vote Share in Mayoral Race	−0.138	−0.153	−0.168	−0.051	0.099	0.130	0.102	0.092	0.038
	(0.227)	(0.183)	(0.179)	(0.228)	(0.459)	(0.195)	(0.193)	(0.354)	(0.351)
Lagged Final Elections in Mayoral Race	−0.085	−1.359	−0.945	−0.725	1.880	−1.587	−1.349	−1.543	−1.148
	(1.087)	(1.030)	(0.960)	(1.054)	(1.993)	(1.086)	(1.042)	(1.639)	(1.562)
Lagged Party Vote Share in Mayoral Race		−0.638***				−0.387*		−0.442	
		(0.193)				(0.201)		(0.270)	
Year	1.012***	0.543*	0.347	1.053***	0.385	0.959***	0.829***	1.102**	0.891*
	(0.337)	(0.300)	(0.277)	(0.336)	(0.617)	(0.326)	(0.308)	(0.520)	(0.481)
Years in Politics		−0.334*	−0.576***			−0.998***	−1.255***	−0.219	−0.398
		(0.173)	(0.183)			(0.275)	(0.267)	(0.262)	(0.293)
Constant	−0.344	−1.181	−0.654	−1.431	−0.433	−1.899*	−1.612	−2.995*	−2.561*
	(1.076)	(1.010)	(0.937)	(1.049)	(1.851)	(1.071)	(1.025)	(1.623)	(1.528)
Observations	188	188	188	188	87	188	188	87	87
Log Likelihood	−111.435	−114.858	−121.835	−111.497	−31.087	−101.700	−103.749	−38.839	−40.430
Akaike Inf. Crit.	232.870	243.716	255.670	232.993	72.174	217.400	219.498	91.678	92.859
Bayesian Inf. Crit.	249.053	266.371	275.088	249.175	84.503	240.055	238.916	108.939	107.654

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

Table 3.5: Candidate-Centric Results for Mayoral Elections

	Unit of Analysis: the Election of a Given Office in a Given Year					
	Dependent Variable: the Number of Candidates in a Given Category					
	Distribution: Poisson					
Dependent Variable Category:	New Candidates	Independent Candidates	—New Candidates in New Parties —		—Candidates in New Parties —	
Independents Counted:	yes	yes	yes	no	yes	no
model:	(25)	(26)	(27)	(28)	(29)	(30)
Lagged Max Vote Share	−0.082	−0.117	−0.039	0.159	0.055	0.151
in Mayoral Race	(0.187)	(0.148)	(0.210)	(0.504)	(0.166)	(0.376)
Lagged Final Elections	0.019	0.145	−0.181	1.491	−0.205	−0.193
in Mayoral Race	(0.481)	(0.450)	(0.535)	(1.275)	(0.472)	(1.052)
Year	1.035***	0.426**	1.132***	0.727	0.974***	1.048**
	(0.193)	(0.173)	(0.216)	(0.523)	(0.183)	(0.415)
Constant	−0.010	0.877***	−0.394	−1.294	−0.024	−1.717**
	(0.365)	(0.328)	(0.413)	(0.915)	(0.353)	(0.798)
Observations	49	36	49	49	49	49
Log Likelihood	−68.505	−59.658	−61.664	−21.479	−71.102	−29.695
Akaike Inf. Crit.	145.010	127.316	131.329	50.957	150.204	67.390
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 3.6: Election-Centric Results for all Mayoral Elections

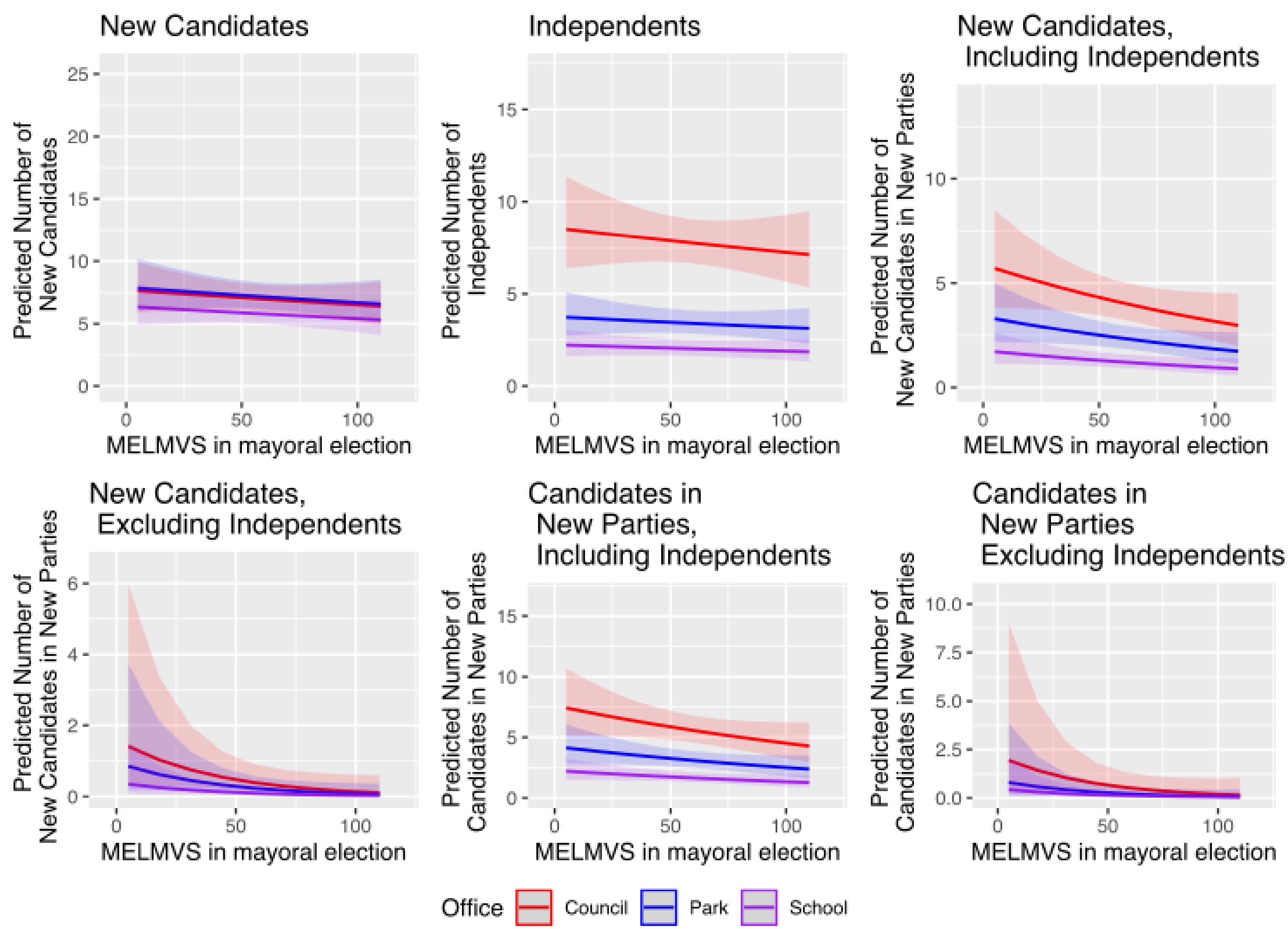


Figure 3.10: The Marginal Effect of the Lagged Maximum Vote Share (MELMVS) in the Mayoral Election on the Dependent Variables, in the MNTV Elections.

Chapter 4

Electoral Reform and System Shock in Vancouver, BC

4.1 Introduction

In 1966, Vancouver's municipal government underwent a broad electoral reform. Prior to the reform, the members of the city's three multiple-member elected bodies–the City Council, School Board, and Park Board–were elected at-large in staggered terms, such that half of each board was elected each year while the mayor was elected every other year. Starting in 1966, the entirety of each board was elected at-large for two-year terms, such that every elected official, including the mayor, served concurrent terms. The long-term effects of this change were much as one would expect; synchronizing all races with the more partisan mayoral races decreased the viability of independent candidates, while the increase in magnitude both increased the fragmentation of the city's party system and made the system more permissive to new candidates. More interestingly, the reform also caused a relatively short-term 22-year-long "shock" over 11 elections that exaggerated several elements of the long-term effects while suppressing others.

This shock was driven by candidate-level behavior but its effects were manifest and observable at the party-level as well. From the period from 1968 to 1990, candidate behavior was significantly more erratic than in previous and following time periods. Partisan candidates during this time were less likely to retire, but more likely to change parties, more likely to found new parties, and more likely to run for mayor. The effects of this erratic candidate behavior resulted in the quick rise of new parties immediately following the reform.This suggests the main thesis of this chapter, that political actors within established political organizations will respond to constitutional change by making bolder and riskier career choices in an attempt to exploit the new system uncertainty. The resulting periods of instability eventually end as political actors come to better understand their strategic options within the new system, power is reconsolidated, and overall uncertainty is reduced.

This chapter has two goals. First, it builds a simple theory of electoral system change, wherein political systems respond to electoral institutional changes with a period of chaotic transition before ultimately settling into a new and stable equilibrium. Second, it measures and identifies the short-term and long-term effects of the 1966 electoral reform. The patterns in the data that reflect these effects are elusive, but are revealed and quantified through a novel time-series neural network designed for the task.

4.2 The Reform and the Collapse of a Political System

The basic theory of this chapter is relatively simple: changes to the rules of an electoral system may produce long-term changes to a political system, namely the behavior of the candidates and parties therein, and eventually lead to a new stable status quo, but in the short-term, the new rules will produce uncertainty, which causes actors to act more erratically, which generates systemic instability, which produces more uncertainty. The short-term effects, therefore, can produce a temporarily self-sustaining equilibrium of chaotic behavior at the individual level that supports and is supported by an uncertain political system at the aggregate level. This temporary equilibrium, or "shock" as this chapter calls it, lasts until enough actors discover and understand how to operate effectively under the new rules, thereby establishing a new stable long-term equilibrium.

While the effects of electoral reform have been studied at all levels of government in countries around the world, most studies approach the topic as a simple comparison between the pre-reform system and post-reform system. To the extent that the literature has suggested a more dynamic long-term effect of electoral reform, such consideration has generally been limited to the idea that the effects of electoral reform may be gradual or slow. Very little consideration has been given to the possibility that electoral reform may produce non-linear effects over time, including the possibilities that reform may produce short-term effects that exaggerate or reverse the long-term effects, despite the fact that studies on the matter are otherwise advanced methodologically. Bowler, Donovan and Brockington (2003, p. 54) provide a rare exception to this deficit, although only at a theoretical level, by suggesting both that candidate entry may increase briefly following the adoption of a cumulative vote ("CV") electoral system at the local level and that long-term increases to party fragmentation would be slow if CV were adopted for US congressional elections. This chapter fills this gap.

In terms of substance matter and scope, this chapter follows in a long line of works on the effects electoral reform. As in this chapter, much of the recent research has been on municipal elections. For instance, research on municipal election reform have found that: both the cumulative vote and limited vote systems produce more minority representation than majoritarian systems (Bowler, Donovan and Brockington, 2003); the effects of electoral reform on women's representation is contingent upon the type of office (Hinojosa and Franceschet, 2012); minor increases in proportionality to municipal elections that are already proportional causes an increase in the number of parties (Fiva and Folke, 2016); and that switching from a single member districts system to the more proportional single transferable vote system also causes an increase in the number of parties (Clark, 2021). And, of course, the body of work studying the effects of electoral reform at the country level by comparing the pre-reform to the post-reform era, rather than through cross-sectional analysis, is even more vast and too expansive to review here.

The types of effect of electoral reform which researchers study can generally be grouped into two categories: representational and systemtic. Representational effects of electoral reform include those regarding changes of the percentage of women or minorities that are elected to office under different electoral systems. In addition to Bowler, Donovan and Brockington (2003) and Hinojosa and Franceschet (2012) discussed above briefly, Norris (2006) provides a good typical example of the study of representation effects in her work on women's representation in the Dutch Parliament.

This chapter is among the works that study the second type of effect of electoral reform: systemtic effects. Systemtic effects include those regarding changes to the number and size of parties (both in elections and in office) and the strategic considerations of all actors in the system. Just as Hinojosa and Franceschet (2012), Fiva and Folke (2016), and Clark (2021) found that increased proportionality yields higher party fragmentation at the local level, many scholars have look for and have generally found similar unsurprising patterns in the election of higher offices (Siavelis, 1997; Sanchez, 2002).

4.3 A Machine Learning Model of Candidate Behavior

The model built in this section is built on three assumptions. Firstly, each candidate's choice after a given election can be explained by four factors: their own political history, the history of their party, the history of the city's political system as a whole, and an idiosyncratic error unique to the candidate at that specific time. Secondly, for each given personal, party, and city political histories, more recent events are generally more relevant than less recent events (e.g. when considering their options, a candidate is likely to be more influenced by their most recent electoral performance than their performance in an earlier election). Finally, while recent events are generally more relevant than older events, different events may produce effects that linger longer than others. For instance, the effects of an extremely lopsided election ten years ago may have a stronger effect on current politics than a more mundane election five years ago.

Given these assumptions, the complex nature of the data and the patterns of interest within the data, this chapter builds a recurrent artificial neural network to examine candidates' behavior following a given election. An artificial neural network, at its most simple, is a series of layers, with each layer comprised of one or more "neurons" or "cells." The neurons in each layer are controlled by the neurons in the previous layer by a combination of weights and biases. The first layer receives the model's input (i.e. numeric independent variables) and the last layer produces the model's estimates of the dependent variables. If the model only has two layers– the input and output layers–then it closely approximates a series of OLS regression models. For instance, if a two-layer model only has one neuron in its output layer, then the weights connecting the input neurons to the output neuron are equivalent to the coefficients of an OLS model while the bias applied to the output neuron is equivalent to the constant term of the OLS model. As layers (known as "hidden layers") are added between the input and output layers, then the neural network can be envisioned as roughly equivalent to an OLS model with many complex interaction terms, although the neural network is much less "linear" than a OLS model because neural networks employ nonlinear "activations" on the output of each neuron (e.g. transforming output with a logistic function to keep all values between zero and one) thereby allowing the model to capture more complex patterns than could be found through the linear combination of interactive terms. The structures of simple neural networks are shown in figures 4.1 and 4.2.

Unlike an OLS model, there is usually no closed-form solution to a neural network. Instead, a random selection of weights are initially selected, then data is "fed" into the model in batches. After each batch, the weights and biases are adjusted slightly to make the model's predictions more accurate. This general process is called "learning" or "training," while the underlying algorithm of using error at the end of the model to adjust weights earlier in the model is called "backpropagation." Batches are fed into the model until a locally optimized set of weights and biases are found.

Because of the time-series nature of the data and assumptions, a temporal consideration must be included in the model. Artificial recurrent neural networks address this consideration by allowing a neuron in a given layer to receive input from both the neurons of the previous layer and from itself at the previous time-step. As with the intra-time weights, the weights across time are trainable, allowing the model to allow more or less information to pass across time-steps.

The ideal type of artificial recurrent neural network to serve this model's specific assumptions is a long short-term memory (LSTM) model, which accounts for differing durations of effects between different inputs (Hochreiter and Schmidhuber, 1997;

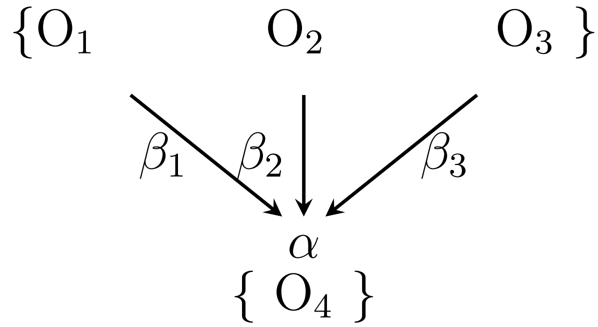


Figure 4.1: A Neural Network Representation of a Simple Linear Regression with Three Independent Variables (O_1, O_2, and O_3) wherein the Predicted Output (O_4) is Equal to $\alpha + \beta_1 O_1 + \beta_2 O_2 + \beta_3 O_3$. Alternatively, this figure can also represent a simple logistic regression if the activation function of the output neuron is a logistic function.

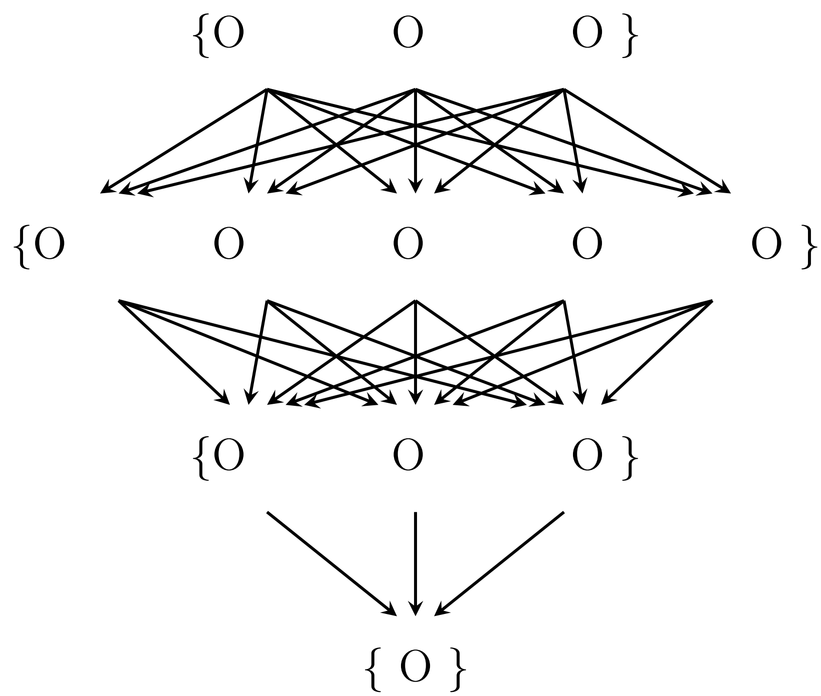


Figure 4.2: A Deep Neural Network with Three Inputs, One Output, and Two Hidden Layers. Each neuron is represented by "O." Each arrow has a trainable weight and each neuron, except for the input neurons, has a trainable bias. The value of a given neuron j is equal to $f\left(b_j + \sum x_i \times w_{i,j}\right)$, where $f()$ is an activation function, b_j is the bias of neuron j, x_i is the value of a given neuron i in the previous layer, and $w_{i,j}$ is the weight of the connection between neurons i and j.

Graves, 2012). An LSTM model is defined by the presence of a layer of LSTM cells within the artificial neural network. Like a neuron in a standard artificial neural network, an LSTM cell receives input from neurons from the next-higher layer at time T and, as with memory cells in a simpler recurrent neural network, an LSTM cell also receives as input the weighted output from its counterpart at time $T-1$. In a simple recurrent neural network, the weight between across time-steps from memory cell to the next is trainable but uniform, making no distinction between information which needs to be remembered over time and that which does not. In an LSTM cell, however, an additional set of weights is trained to allow the cell to distinguish what new input should be "remembered" given the output of the cell at $T-1$ and what output from the cell at $T-1$ should be "forgotten" given its new input. Information that is "remembered" by an LSTM cell is called the cell's "state" and is passed to the corresponding LSTM cell at $T+1$ along with its output. In short, an LSTM cell at time T receives inputs (the current output of the preceding neural layer at time T, the output of the corresponding LSTM cell at time $T-1$, and the state output from the corresponding LSTM cell at time $T-1$) and produces both a regular output and a new state. The general structure of an LSTM cell is shown in figure 4.3. In addition to predicting, and thereby providing a step toward understanding and explaining candidate and party behavior, a well-structured LSTM can also group observations to find important similarities that are not obvious to casual observation or more conventional methods. This analysis by grouping is discussed in greater detail below.

In the context of Vancouver, an LSTM cell can be understood in the following contrived and over-simplified examples. Once an LSTM model has been fitted to the data, an LSTM cell within the party section of the model may be trained to ignore all input regarding the party's performance in the Park Board elections if the

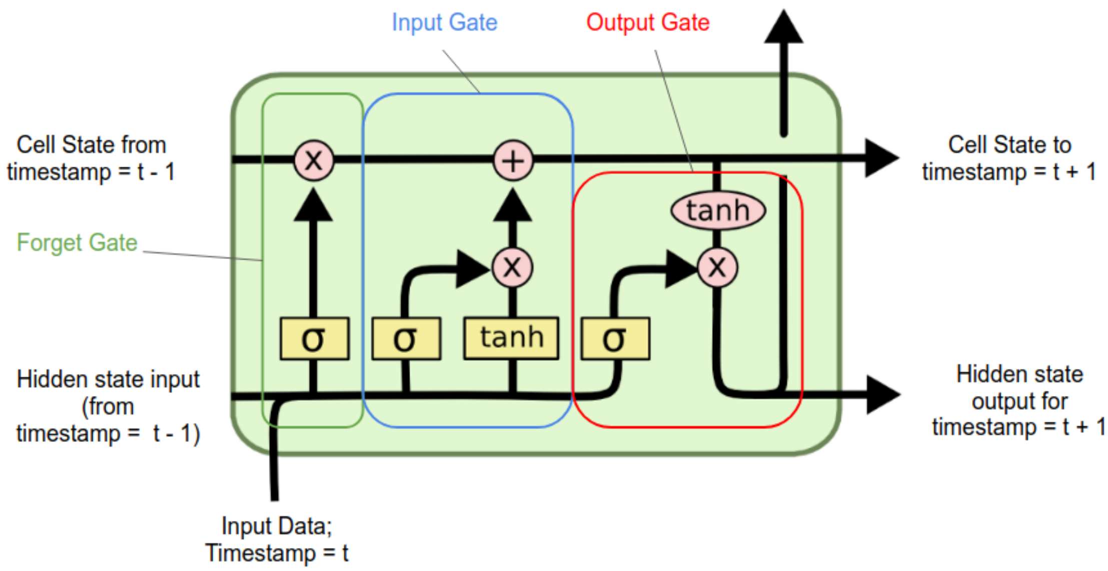


Figure 4.3: A Common Depiction of a Typical Long Short-term Memory (LSTM) Cell. This particular rendering was created by T. J. J. (2020). Each yellow box contains trainable weights that control how the cell reacts to both old information and new input.

party has won a mayoral election relatively recently. Alternatively, an LSTM cell within the system section of the model may be trained to be highly sensitive to all new inputs if a recent election returned a high number of independents into office. Finally, a cell in the candidate section of the model may be trained to forget all previous "memory" of a candidate if a long-time unsuccessful candidate is suddenly elected to the City Council. While these specific examples are merely illustrative, they demonstrate basic types of patterns that LSTM cells can be trained to recognize, most notably that when conditions for a candidate, party, or system are uncertain, new information is valuable, and when conditions are stable, only unusual new input is valuable.

The three sets of variables (candidate, party, and city) are segregated for most of the model. Each set of variables pass through similar hidden layers before joining together and then finally producing one output layer to predict the candidate's decision following a given election. Specifically, the data is first time-distributed so that observations of the same unit over time are fed into parallel layers. Then the variables are segregated into three sets (candidate, party, and city) and each is fed into its own series of sequential dense neural layers, then into two sequential LSTM layers, then into a bottleneck layer that doubles as an encoding layer for the candidate, party, or year.[1] The output from the bottleneck layers are then combined by year, bringing information from the candidate, party, and city together. This combined output passes through four sequential dense layers of eight neurons each, which finally leads to the output layer of six neurons, each representing a different choice the candidate can make.[2] The structure of this model is shown in figure 4.4.

The driving logic of this model's structure is to first isolate important intra-year feature of a given unit (candidate, party, or year) with the initial dense layers, then isolate relevant temporal features across time with the LSTM layers, then reduce the relevant information from a given unit in a given year to its bare minimum number of dimensions with the bottleneck layer, then finally isolate the interactive effects of candidate, party, and year features in the combined dense layers. This structure also has practical advantages as well. Firstly, the initial dense layers greatly reduce the amount of information passing into the LSTM layers, which are computationally expensive. Secondly, the output of the bottleneck layers within the middle of the model provide interesting results in their own by providing a low-dimensional representation of all relevant information from the past and present of given a candidate,

[1] For other examples of encoding layers in a supervised LSTM model, see Mao et al. (2018) and Appiah et al. (2019).

[2] While this model's specific architecture is novel, the general idea of running parallel LSTM networks within a larger network can be seen in and validated by Appiah et al. (2019).

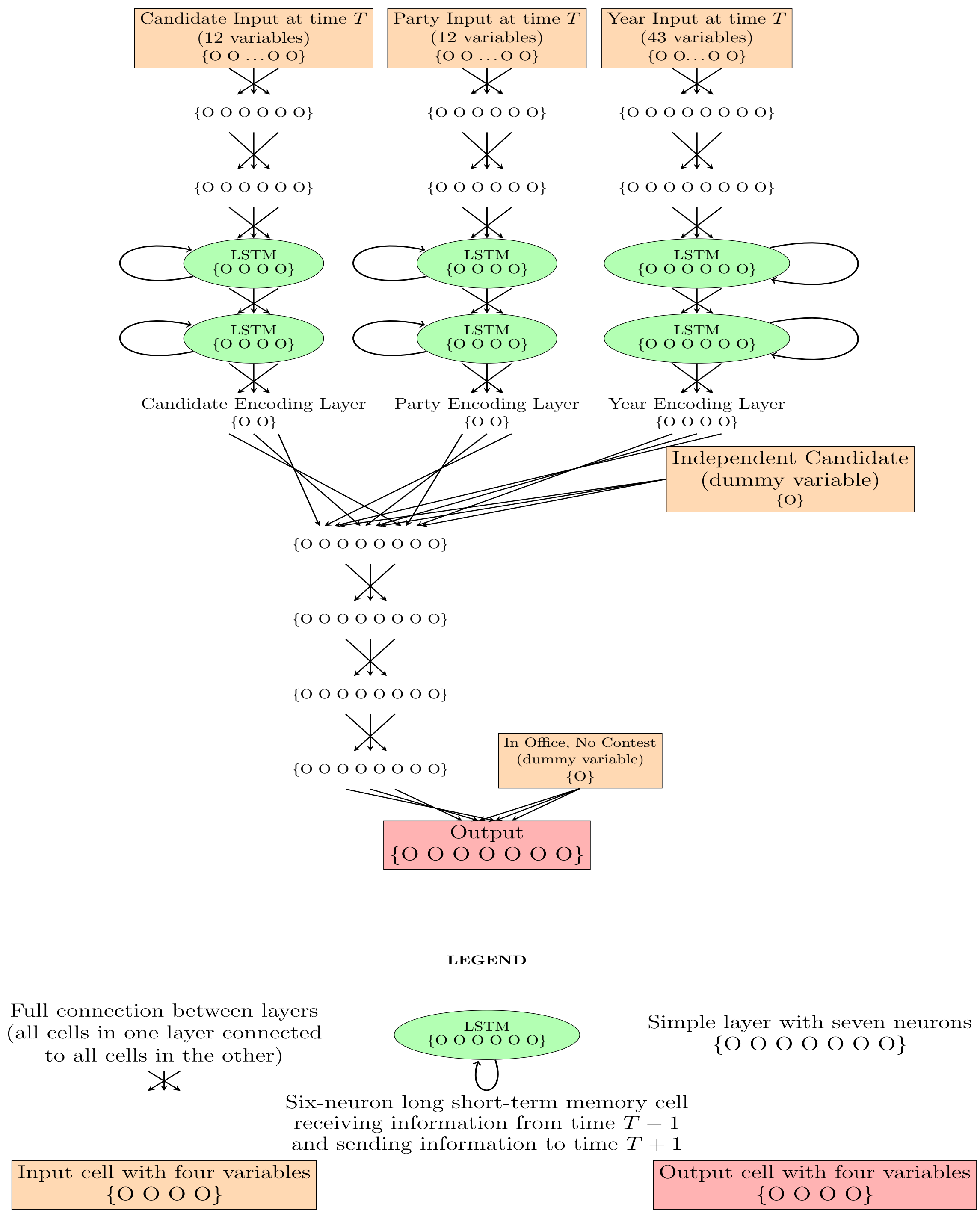


Figure 4.4: The Neural Network Model of Candidate Behavior.

party, or the city. The utility of this bottlenecked information is explored below.

4.3.1 Variables

Following the assumption that a candidate's decision at time T is based upon their own history, their party's history, and the city's history, it is necessary to specify what variables constitute each given history. Each candidate's history is defined by a sequence of candidacies, with each candidacy at time T defined by the following variables: the office for which they are running (which is in turn divided into five binary variables representing each of the four offices plus a fifth binary variable for dormant candidates who are not running for office in the given election but have run for office in the past and will return to politics in the future), a binary variable for whether or not the candidate was elected in the given election, two variables representing the effective size of the candidacy (one based on the number of candidates and another based on their vote share, discussed in greater detail below), the number of years that the candidate has been in city politics, the magnitude of the office for which they are running,[3] a binary variable denoting whether or not the candidate is in a different party than in their previous election,[4] a binary variable indicating whether or not the candidate is an independent, and a final binary variable indicating whether or not the candidate is an incumbent. The party variables all measure the effective size of the party, based upon N_C, N_V, and N_S for all four offices, for a total of 12 variables. Finally, the city-level systemic variables are 11 measurements of fragmentation (N_C, N_V, and N_S for all offices, sans N_S for mayor), the effective size of the independents and the effective size of new candidates (again,

[3] This variable is partially standardized for non-mayoral candidates by subtracting the given magnitude by the given office's lowest magnitude. For example, the lowest City Council magnitude is four, so all City Council magnitude values are reduced by four. This helps train the model because neural networks benefit if their neurons' values can have a value of zero.

[4] With a value of zero for their first election.

both based on N_C, N_V, and N_S for all offices), and the proportionality of the results of all four offices. While there is a high degree of multicollinearity between these variables, this is generally not a problem for neural networks, which are able to reduce the redundancy of input variables as the information passes into deeper and deeper layers of the given network (De Veaux and Ungar, 1994).[5]

This chapter proposes a metric called "effective size" (X_i or, alternatively, the "partial effective number") to measure the strength of an element within a fragmented set. It is calculated such that if the fragmentation of the set, as measured by the effective number formula, is equal to N and the element in question possesses $\frac{1}{N^{th}}$ of the set, then the element's effective size is one. For instance, if N_V is equal to two and a given party has one half of the vote, then the party's effective size is one. This metric is a useful tool for comparing party or candidate across different elections while accounting for varying levels of fragmentation between different years and offices. By contrast, simply comparing the share of the vote would fail to account for the fact that a party with 30 percent of the vote is a good performance in a highly fragmented party system but a poorer performance in a less fragmented one. This tool is also useful for examining the relative combined strength of different types of elements within the set, given its fragmentation. For instance, X_i can be used to measure the effective size of an individual candidate,[6] the effective combined size of independents within the party system (X_{Ind}), or the effective combined size of all new candidates within the party system (X_{New}). In such an application, X_{Ind}

[5] This can be understood, broadly, as each layer condensing the information provided by the layer before it, driven by the decreasing number of neurons in each subsequent layer. In the deeper smaller layers, therefore, information is stored more efficiently, with fewer redundancies and less multicollinearity between neurons of a given layer. It is, however, this same process of condensing information that makes the results of any particular neural network difficult to directly interpret.

[6] When measuring the effective size of an individual candidate in an MNTV election, X_i is multiplied by the magnitude.

or X_{New} is equal to one if the independents or new candidates together control $\frac{1}{N^{th}}$ of a system that is fragmented N ways. The effective size of an element within a fragmented set is calculated as follows, where v_i is the element's share of the vote and N_V is the effective number of parties based on vote shares: $X_{v,i} = v_i \times N_V$. If fragmentation is based on candidate share or seat share, then $X_{v,i}$, v_i, and N_V are replaced by $X_{c,i}$, c_i, and N_C or $X_{s,i}$, s_i, and N_S, respectively.

4.3.2 Technical Hurdles and Considerations

The model was built and trained using TensorFlow in R. The training was conducted in several stages. First, the parts of the network concerning the parties and the system as a whole were pre-trained on aggregate party-level data, while excluding independents. Then the model was "fine-tuned" with the introduction of independents along with a variable to encode whether or not the party is an independent candidate. Finally, the model was expanded to include all candidate variables and then trained on the full data.

With the exception of the bottleneck and output layers, all neurons have rectified linear unit (ReLU) activation functions. As mentioned above, a neuron's activation function is what transforms the sum of its inputs into an output. Without an activation function, the model's output would simply be a complex linear combination of the input and the model would be unable to learn more complex patterns. The ReLU activation function is a computationally cheap, yet common and effective activation function in neural networks Zhang et al. (2018) that returns $max(0, x)$, where x is the sum of the neuron's bias and weighted inputs. The bottleneck and output layers use logistic activation functions $\left(\frac{1}{1+e^{-x}}\right)$ to confine their output to between zero and one. For the bottleneck layers, this makes their respective latent spaces easier to explore visually. For the output layers, this produces predicted output that is interpretable as a probability.

One of the greatest risks in training a machine learning model is over-fitting the model. An over-fit model is one that has learned the data in such detail that it has memorized the specifics of the data instead of identifying more generalizable patterns.[7] This can be avoided by limiting the number of neurons in each layer, as done in this model, thereby lowering the capacity of the model to learn more complex patterns in the noise and forcing it to learn more simple and generalizable patterns instead. Over-fitting can be detected by randomly dividing the data into a training set and validation set. The model is then trained on the former, then it is judged on its ability to predict the validation set. If the model performs well on the validation set, which it has never seen before, then it can be said to have learned generalizable information from the training data, rather than merely learning its noise. In the case of this model, ten percent of the candidates were set aside as validation data. The mean squared errors for both the training and validation sets were both .05, which strongly suggests that the model is not over-fit.

The model uses masking on inputs to account for non-existent data. Masking tells the training algorithm to ignore a given input in certain circumstances, thereby allowing the rest of the model to be trained while leaving the weights for the given input untouched. Masking allows these years to be included in the model training, whereas they would be dropped under more conventional methods. For instance,

[7] To understand an over-fit neural network that has "memorized" its data, consider a model designed to find the relationship between education and income, wherein the individual is the unit of analysis. A well-fit model will include confounding variables to help control for factors such as age and region to help the model discover the true relationship between education and income. If too many variables are included, however, the model may find more powerful ways to predict income by finding a unique pattern for each individual in the data. For instance, the model may conclude that 38-year-old women born on sunny Tuesdays in southern Oregon will earn exactly $135,000, and find similar patterns for every other individual in the data, thereby "memorizing" the data rather than extracting patterns that will hold true beyond the original dataset.

in years without mayoral election, there is no valid input for the system variables related to the mayoral races. Masking also allows a party's vote share in a given race to be ignored if they did not contest it.

The final model includes one important modification to account for one complication in the earlier election years. Prior to 1960, terms of office were staggered such that only half of all offices were up for election at any given time. This means that candidates can exist in an additional state: in office, not up for reelection. This state is problematic for two reasons. Firstly, it preempts the ability of a candidate to choose another state (e.g. an incumbent who is not up for reelection cannot be observed changing parties). Secondly, it makes most candidate-level variables unobservable (e.g. the level of support enjoyed by such an incumbent cannot be measured if they are not up for reelection). As a solution, I add an additional input and an additional output to the model, both of which are coded as one when the candidate is not up for reelection. This new input is inserted in the very last dense neural layer and given initial weights to suppress all output save for the new one. This prevents the model from drawing poor conclusions about the relationships between the regular inputs and outputs during non-contested incumbency years while allowing such years to be incorporated into the candidate's history.

One complication with this model is that the data is inherently temporally complex. Its temporal complexity is driven by three considerations: firstly, that entirety of a single candidate's political decisions throughout their career constitutes the output of a single observation, secondly, that a candidate can change parties throughout their career, and finally, that a candidate's decision at time T is affected by the entire political history of the party with which they are affiliated at time T. As shown in figure 4.4, a candidate's political history can only be predicted by observing the political history of every party with which they are affiliated throughout their career.

This essentially creates a second temporal dimension for each candidate's partisan history.

This second partisan dimension is part of a greater complication of "ragged" (or "jagged" depending on which terminology one prefers) data. Ragged data is data in which some vectors within an observation's data are longer than others (i.e. a cross-section of a candidate's history cannot be represented rectangularly in a spreadsheet with one dimension representing time and the other representing different variables). In this case, the model's assumption is that a candidate's decision at time T is driven by their own history, their party's history, and the system's history, which becomes problematic if a candidate has been in politics for four years, their party has existed for twenty years, and the system (for the purposes of this chapter) has existed for 100 years. This is further compounded if the candidate has been a member of multiple parties, each of which have existed for varying lengths of time.

The general solution to the temporal complexity of the data is simply to "feed" candidates into the training algorithm one at a time, which is considerably slower than using multiple-observation in large batches. Single-observation training batches have the added benefit of optimizing the neural network through regular gradient descent rather than through the stochastic gradient descent of multiple-observation batches. Simply put, this means that neural network is more carefully optimized, because the local optimal solution is searched for with smaller and more precise steps rather than larger less precise steps.[8]

[8] For comparisons between regular and stochastic gradient descent, see Bottou (1991) and Amari (1993).

4.4 Analysis and Results

As with a more conventional model, interpreting the results of a machine learning model is essentially a two-step process: first identify the patterns found by the given method (whether that be the coefficients returned by solving simple ordinary least squares ("OLS") regression model or the weights returned by optimizing a deep neural network), then translate those patterns into substantive analysis. Both steps are more difficult in artificial neural networks than in simple regression models because the patterns of the former are more complicated. Each coefficient in an OLS model can be interpreted as the effect of a one unit increase of its corresponding variable on the dependent variable, all else held equal. However, the input of each variable of a neural network goes through numerous transformations and its effect on the final predicted output is heavily contingent upon the inputs of every other variable, each of which is also repeatedly transformed. This means that the relationships between the input and outputs of even a relatively simple neural network are generally not interpretable in a simple and straightforward manner.

To that end, this section first explores the latent spaces of the candidate, party, and city variables. This leads to an understanding of, from all of the possible differences between two units (e.g. between two parties), what differences and similarities matter in the causal mechanisms that influence candidates' decisions and which do not. A given unit's position in latent space is a distillation of as much relevant information from the past and present of the unit as possible; all input is collapsed into the latent space and all output is derived from the latent space. In an artificial neural network, the latent space is produced by a "bottleneck" layer, with the output of each layer providing the value of a given dimension. The number of dimensions in a latent space is thus determined by the number of neurons in the corresponding bottleneck layer of the neural network, as seen in figure 4.4. The candidate and party sections of the neural network are each bottlenecked down to two-neuron layers, so the candidate and party latent spaces can both be mapped in two dimensions. The bottleneck layer of for the city as a whole consists of four neurons and thus can be mapped with four dimensions or four variables. All output from the model is produced by the combination and interaction of these three latent spaces.

The second step is to draw meaningful and statistically testable conclusions from the neural network, building upon the exploration of the first step. Specifically, this second step uses the latent spaces produced by the neural network to identify a large yet temporary shift in Vancouver's political system that began with the electoral reform of 1968 and lasted until 1990. Then, this substantive effects of this shift on candidate behavior are measured along with their statistical significances.

4.4.1 The Latent Spaces

To the end of fulfilling the first step of the analysis, the two-dimensional latent spaces of the candidacies and parties, along with the four dimensional latent space of the election years are plotted below, in figures 4.5, 4.6, and 4.7, respectively. Each point in figure 4.5 represents a candidacy–a candidate's performance in a given election–while the same is true for parties in figure 4.6. Therefore, any candidate or party that participated in multiple elections will represented by multiple points in their respective figure. The visualization of the election year latent space is plotted as a time-series, to better serve its ultimate purpose of illustrating the rise and duration of the shock period on which this chapter is focused.[9] While the latent dimensions in these three plots have no directly interpretable substantive meaning, the relative positions of the candidates, parties, and years within each plot help illustrate which

[9] This is a simpler yet conceptually similar method of that suggested by Malhotra et al. (2015).

kind of differences between observations do and do not matter.

Figure 4.5 shows that, despite the network receiving two dimensions to train, most relevant information regarding the past and present of a candidate can be consolidated into an almost one-dimensional space. Even a casual survey of the chart shows what one might easily expect: the most important candidate-level variable when determining a candidate's next career step following an election is their performance in that election. Elected candidates are generally grouped near the low end of their latent X-dimension and near the high end of their latent Y-dimension. The general exception to this rule are first-year council candidates who run and win under the banner of major parties, whom the algorithm has trained the network to treat as though they were independent candidates. The main takeaway points from this figure on its own are that a candidate's performance history matters most when predicting their next move, but also that candidate behavior is affected by more than a simple sum of their electoral success and failure as demonstrated by the complex waves and clusters of points throughout the latent space.

As far as the parties are concerned, the trained neural network suggests that the most important difference between different parties at any given time is how large and established they are. In figure 4.6, parties with longer histories, more candidates, more votes, and more seats are found at the lower end of the latent Y-axis, with the strongest and most established of those parties found at the lower end of the latent X-axis. Newer, weaker, and less successful parties, along with most independents, are scattered along the latent Y-axis at the upper range of the latent X-axis. On its own, this figure suggests that there is not much nuance regarding the effect of a candidate's party on their decision following a given election; candidates in large established parties will behave differently than others.

The analysis of the election year latent space is of the most direct concern to

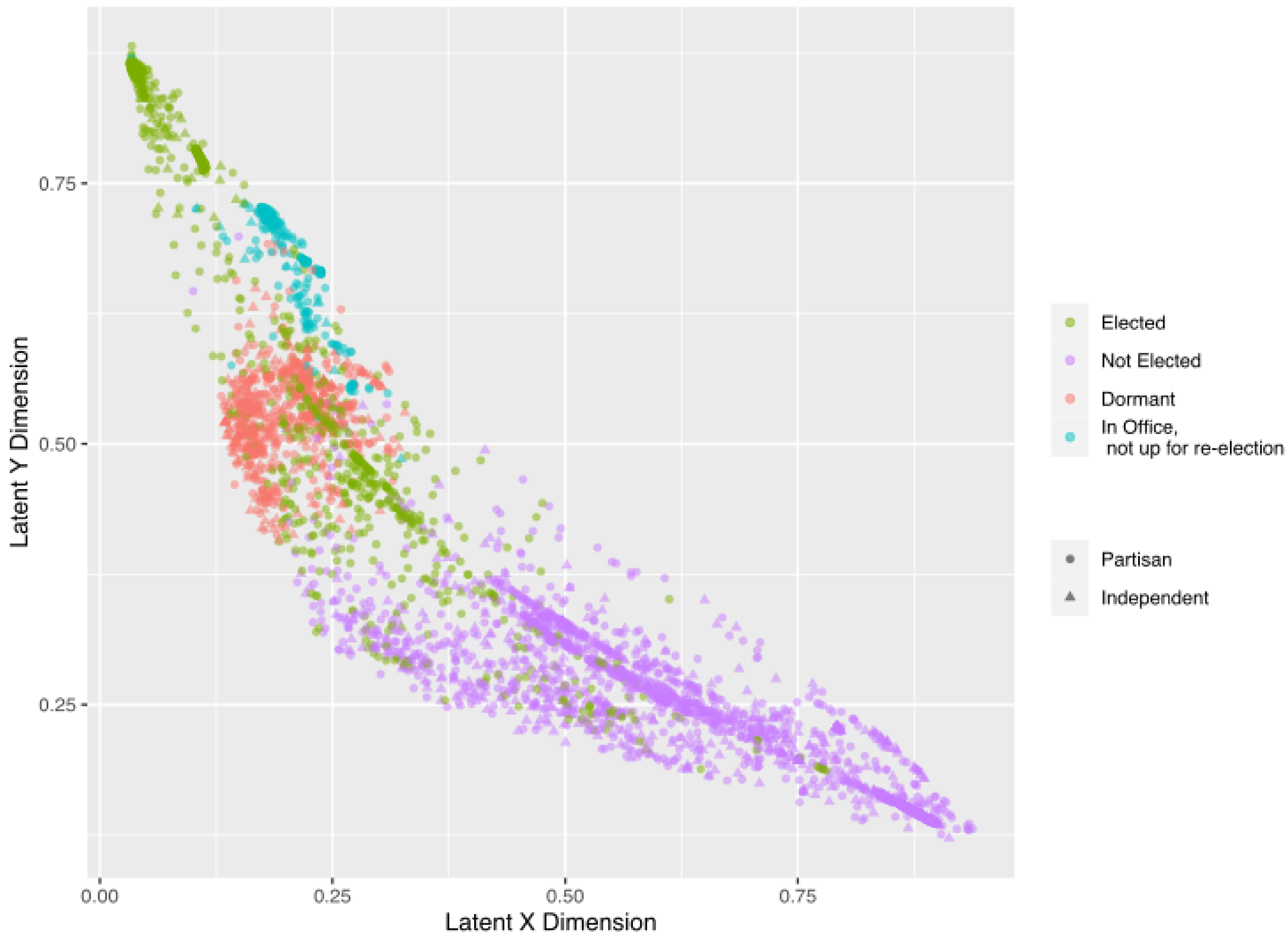


Figure 4.5: The Latent Space of Candidacies. Each point represents a given candidate on a given election year.

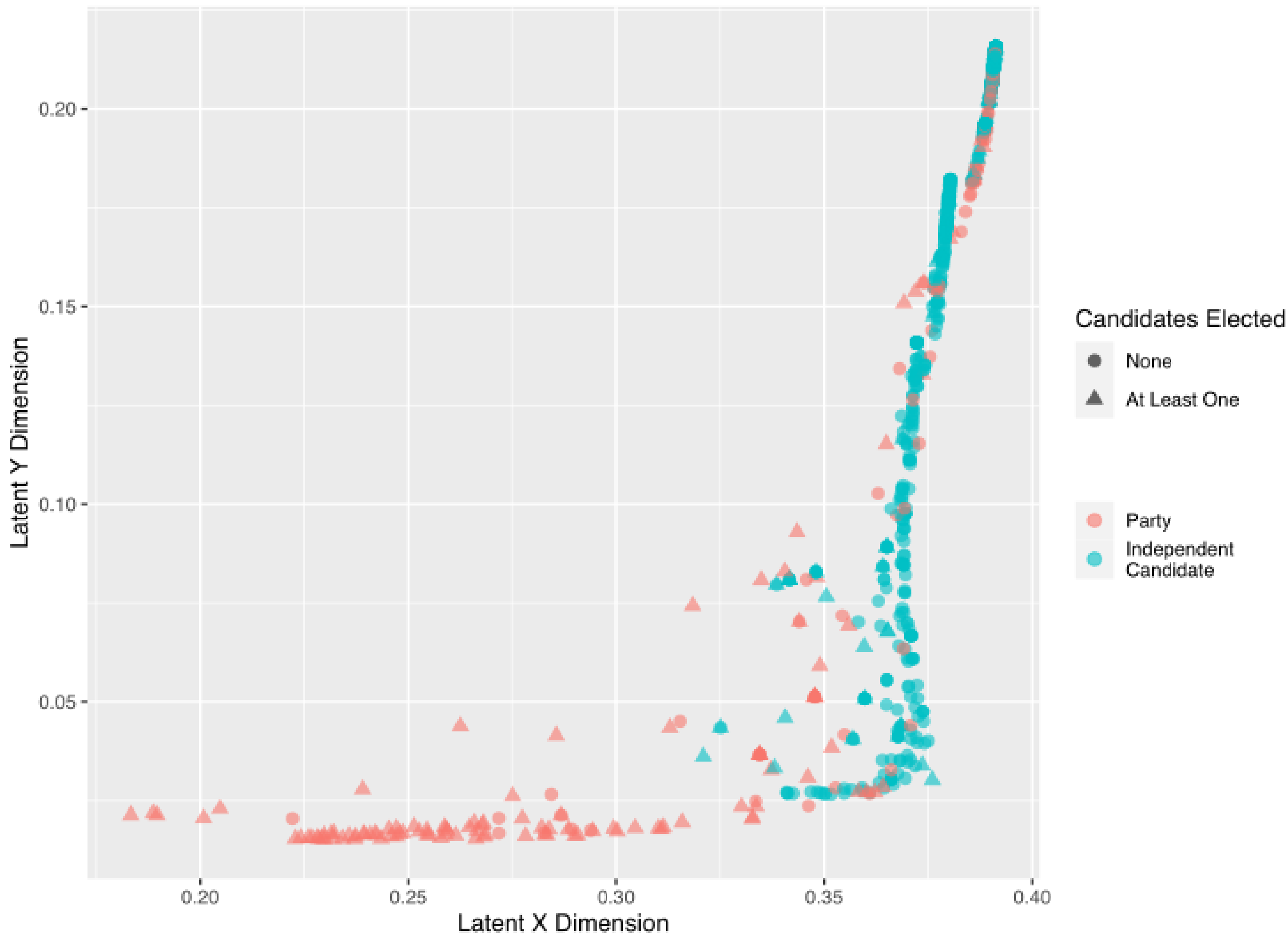


Figure 4.6: The Latent Space of Parties. Each point represents a party/year: a given party on a given election year.

this chapter's analysis, which focuses on the effects of electoral reform on the system as a whole while controlling for candidate and party features. Of most immediate note to this effect, the latent space of the election years show that the 22-year-long shock is not one uniform event from beginning to end. Both a visual account of the latent space, seen in figure 4.7, and a k-means clustering of the latent space shows that the shock can be further divided into two basic phases: a transitional shock phase beginning with the electoral reform of 1968 followed by a peak shock phase. Within the latent space, this transitional period is defined by a large shift across the entirety of the system's bounded first latent dimension accompanied by similar yet smaller shifts across its other three dimensions. The practical interpretation of this transitional phase is of increasing uncertainty and erratic behavior as the stability of the pre-reform political system collapsed over the span of four elections. The second phase, lasting from 1976 to 1990, is the heart of the shock period, characterized first by the absence of stability of the pre-reform era, the dwindling stability of the transitional period, and the relative stability of the post-shock era. This second phase of the shock is the focus of all following analysis and comparison, because the first phase carries lingering effects of the pre-reform era which pollutes attempts to isolate the effects of the shock.[10]

4.4.2 The Effect of the Shock

To the end of deriving meaningful and understandable findings from these patterns to probe this chapter's general examination of post-reform systemic shock, the next step is to simply compare the baseline probabilities of candidates making certain decisions following an election between the different eras. To do so, I compare the predicted

[10] The three-cluster k-means clustering used to identify these periods, which is robust to randomly varied start points, groups 1938-1966 in one cluster, 1976-1990 in a second cluster, and 1993-2014 in a third cluster. 1968-1974, the transitional phase, oscillates between first and third clusters as the system transitions to the second cluster.

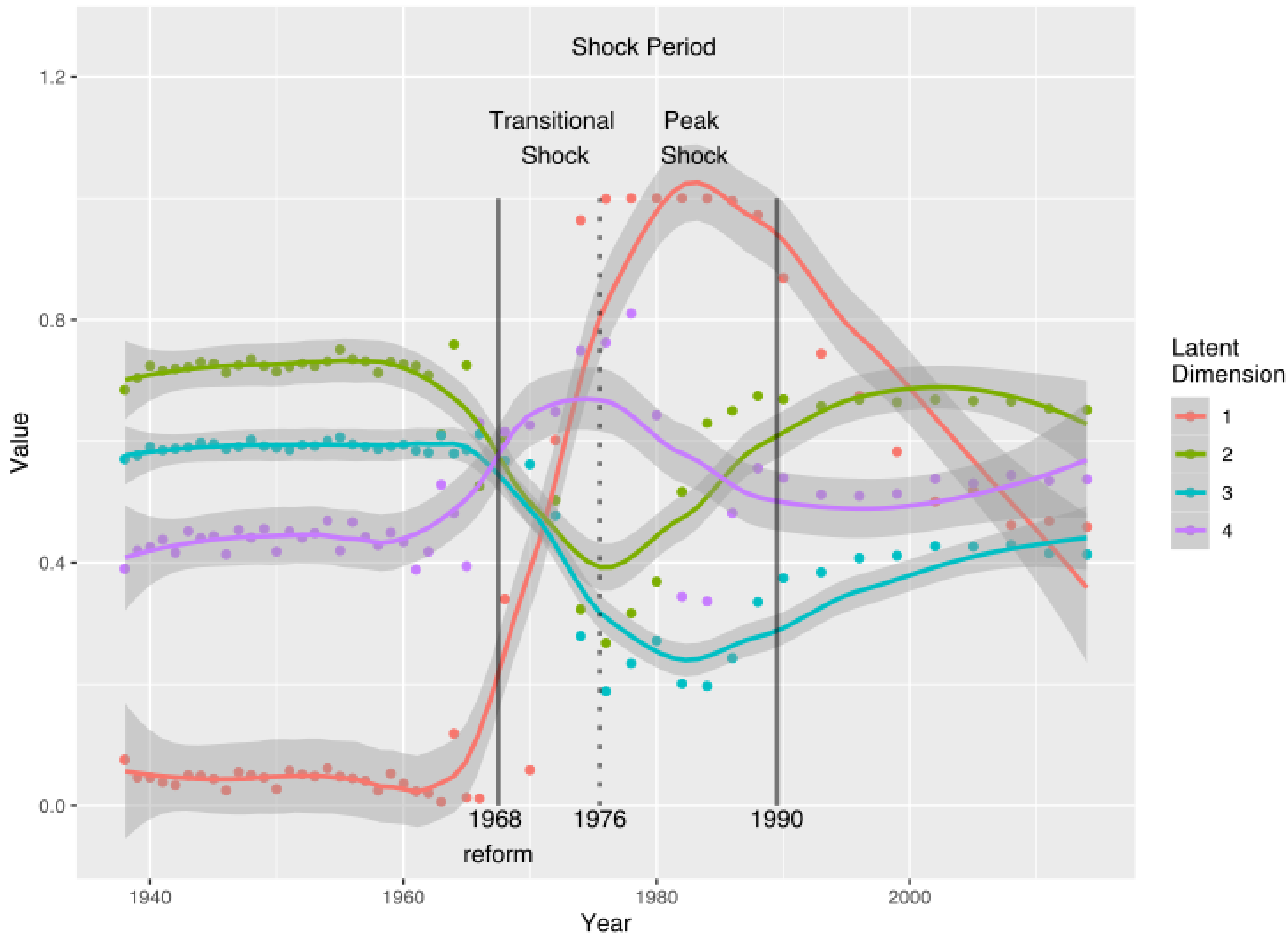


Figure 4.7: The Latent Space of the Political System over Time.

probabilities of given hypothetical candidacy in a hypothetical party between all three eras: the peak shock period (1976-1990), and the preceding and following periods. The given hypothetical candidacy is produced by finding the average position in latent space for all non-dormant and non-independent candidates. This produces a sample candidacy representative of a candidate with a typical past and present political history. The hypothetical party is generated by taking the average position in latent space for party/years in the party latent space, excluding independents. This simulates three counterfactual candidacies, each identical except for the era in which they exist. This methodology is essentially a simple implementation of more advanced machine learning models that similarly use the latent spaces of neural networks to infer causal relationships (Johansson, Shalit and Sontag, 2016; Chen et al., 2019).

To illustrate what these average hypothetical parties and candidates look like, one can simply look at the actual parties and candidates closest to the average positions in their respective latent spaces. The three most "average" parties are Vision Vancouver in 2008, the Civic Voters' Association in 1958, and the Electors' Action Movement (TEAM) in 1970. Each of these three parties at these respective times were young and strong challengers to the existing party establishment, with each winning a significant share of the votes and seats in their given elections. Similarly, the most "average" non-Independent candidates are Anita Romaniuk in 1996, David Cadman in 2002, Raymond Louie in 2008, and Warnett Kennedy in 1974. All four of these candidates, at these respective times, were non-mayoral candidates, relatively new politicians (three were in their second election), and all had lost at least one election previously. Therefore, when comparing the baseline probabilities of the average candidate in an average party, one is comparing a hypothetical candidate and hypothetical party much like the ones described above; the closer a given candidate and party is to these average latent positions, the better the results below will predict their behavior.

Tables 4.1 and 4.2 show the results of this methodology. An average partisan candidate during the 1968-1990 peak shock period had a noticeably lower baseline probability of retiring, but a higher probability of changing parties, co-founding a new party, or running for mayor. The statistical significance of the differences in these predicted probabilities can be approximated by using the standard error formula for linear regression, $\frac{\sqrt{\frac{\sum(Y_i - \hat{Y}_i)^2}{n-2}}}{\sqrt{\sum(X_i - \bar{X})^2}}$, where Y_i denotes whether or not a given candidate chose a given option (e.g. retire), $\hat{Y}_i$ denotes their predicted probability of doing so, n denotes the number of observations, X_i denotes whether or not the given candidate is within the shock period, and $\bar{X}$ denotes the share of candidates within the shock period. When comparing the difference between the shock period and one of the other two periods, the third period is disregarded. This reveals that the differences between the shock period and the other periods are indeed significant at the .05 level, with the exception of the difference between the baseline probabilities of running for mayor in the shock period and pre-shock period which is significant at the .1 level.

4.5 The End of the Shock and the Start of a New Era

This chapter has thus far identified and measured a shock in the political system of Vancouver that began upon the political reform of 1968, reached its zenith in 1976, and lasted until 1990, thereby proving this chapter's thesis that changes to an electoral system can introduce relatively short-term change and chaos to a political system. This leaves to lingering questions. First, how and why did the shock eventually end, leaving the city in its current political state? Second, is the post-shock era any different than the pre-shock era?

Table 4.1: Baseline Probabilities of Candidate Decisions Following a Given Election

Period	Retire	Change Party	Join New Party	Run for Mayor
Pre-Shock & Transitional Shock *1938-1974*	0.665	0.066	0.043	0.023
Peak Shock *1976-1990*	0.461	0.124	0.091	0.052
Post-Shock *1993-Present*	0.581	0.089	0.061	0.033

Table 4.2: Difference in Baseline Probability of Candidate Decisions Between Given Period T and Peak Shock Period: $(P_T - P_{\text{Peak Shock}})$.

Period	Retire	Change Party	Join New Party	Run for Mayor
Pre-Shock & Transitional Shock *1938-1974*	0.204***	-0.059***	-0.048***	-0.029**
Post-Shock *1993-Present*	0.119***	-0.036**	-0.030**	-0.019*

One-tailed T-test, statistical significance of difference from previous era: *p<0.1; **p<0.05; ***p<0.01

The first question is far from simple and a comprehensive answer is beyond the scope of this chapter, because it is difficult to determine whether given events in and around 1990 were either causes or effects of the system reaching a new post-shock equilibrium. Nonetheless, a brief survey of changes in the city's political landscape, at the very least, both illustrates the end of the shock period and raises potential future research questions about the end of transitional political periods caused by institutional change.

The end of the shock period is characterized first and foremost by the consolidation of power by two large parties: the Non-Partisan Association and the Coalition of Progressive Electors. This consolidation included the annexation of the New Democratic Party–the largest third party of the late 80s–into COPE in 1990, along with the collapse of the Electoral Action Alliance and the short-lived Civic Independents, both of which exited the system in 1986. The fledgling Green Party was the only third party to survive the end of the shock period. The fortunes of independents also suffered as the shock period came to and end; Jean Swanson's failed mayoral run in 1988 was the last significant showing an independent in any Vancouver race. The defining characteristics of the shock period, namely greater candidate retention with more erratic behavior, appear to be driven by this partisan consolidation.

To the second question—whether or not the post-shock era is different than the pre-shock era—the general answer is that the system has become more fragmented while current candidate behavior is now similar to pre-shock behavior. In response to the reform of 1968, party fragmentation accelerated as a result of the increased magnitude. This increase is especially pronounced when party fragmentation is measured by the number of candidates in each election (N_C) or the distribution of votes between them (N_V); as the elections are majoritarian, however, fragmentation based on the distribution of seats (N_S) has not noticeably changed. Nonetheless, the general stability of the ordinary city politician's career has increased to the point at which it is neither distinguishable nor statistically different than that of politicians before the reform. As before the reform, a given politician is less likely to change parties, start a new party, or run for mayor, but more likely to retire.

This return of candidate-level stability, even in the face of increased party-level instability, strongly suggests that politicians since 1990 have collectively and generally adapted to the uncertainty that was created by the reform of 1968. Because the measurable shock itself was simply an aggregation of candidate behavior, one could simply say that the shock ended because the political class eventually adapted to the new electoral system. Returning to the first question, this answer is incomplete, because it does not explain why it took 22 years to do so instead of more or less time, nor does it answer why candidate behavior stabilized so suddenly (at least according to the latent space seen in figure 4.7). All together, however, this examination of the end of the shock paints a general picture of eventual return to a degree of normalcy at the individual level once candidates became acclimatized to the new system, despite increasing uncertainty and chaos in the party system as a whole.

4.6 Conclusion

An important takeaway from this chapter is that one should exercise a degree of caution when reviewing old or conducting new research on the effects of electoral reform. If one compares the politics of a pre-reform era with the politics directly following the reform, one might not actually capture the true long-term effects of the reform and may instead capture the short-term effects of instability caused by the change. As for the short-term effects themselves, this chapter has shown that candidate behavior can become increasingly erratic in the elections following the adoption of electoral reform. The obvious question arising from this work is how well the specific findings herein are generalizable beyond increases in magnitude in local MNTV elections in Vancouver. It seems likely, however, that significant changes to an electoral system should engender increased uncertainty and, by extension, increased erratic behavior in actors within the system, regardless of the electoral system, location, or level of government.

Chapter 5

Key Conclusions and Future Research

Together, the papers compiled into this dissertation cary two central themes: strategy and stability. Analysis of New Hampshire showed that voters (or, at least, the politicians marshaling the voters) have a strategic grasp of when and where sincere voting poses a pronounced risk to partisan interests in MNTV systems. Analysis of Vancouver showed that new candidates and party formateurs can strategically time their entry into politics even in the constantly changing and turbulent political landscape of their city, although the predictability of their entry is diminished following changes to the electoral rules.

From these accounts of highly strategic behavior in two different governments, however, come two very different pictures of MNTV systems. In New Hampshire, the two-party system of United States is as stable and as entrenched in the state House of Representatives as it is in the state Senate and governor's office. In Vancouver, however, the party system is in a constant state of change, as parties enter and leave the system abruptly.

What then explains how to systems with such similar rules and similarly strategic actors have such different party systems? The most sensible conclusion is that MNTV, when left to its own devices, creates inherently unstable party systems, but that this instability can be greatly curtailed by the stability of party systems of higher offices. These stabilizing higher offices can be "up the ballot" (i.e. contemporaneously elected with and more prestigious than the given MNTV election), "uphill" (i.e. a more prestigious body in the same government as the given MNTV-elected body, such as a senate relative to a house of representatives), or both. Up-ballot races help stabilize MNTV party systems during the campaign season and uphill races help stabilize the MNTV party system once in office.

In New Hampshire, the party system of the state House of Representatives is deeply embedded in the party systems of higher offices. The New Hampshire House of Representatives election is always down-ballot from the election of the state Senate, governor, and US House of Representatives and often down-ballot from the election of the US Senate and president as well. Once in office, the New Hampshire House of Representatives is also "downhill," so to speak, from the governor and state Senate. There is, simply put, no room for the party system of the state House of Representatives to become as chaotic as that of Vancouver; its party system is too deeply embedded in and driven by that of the state and nation. Strategy in New Hampshire's MNTV elections, therefore is not found in the actions of new parties and political entrepreneurs, but in the considerations and calculations of the major parties and their supporters, as they attempt to project their dominance from the predictable higher elections into the election of the state House of Representatives.

Vancouver, by contrast and with few exceptions, has held its elections out of sync with those of its province and nation. Its political system exists, not in a complete vacuum, but relatively isolated from any up-ballot effects. This largely explains its endemic political parties, which usually have no formal or even informal connection with the parties of British Columbia or Canada. This isolation also likely explains the chaotic nature of the city's party system; without higher offices to stabilize the system, MNTV shows its full chaotic effect. But, as chapter 3 shows, Vancouver does have one up-ballot and uphill office that can bring a degree of stability to the city–the mayor. In the presence of a politically strong (i.e. one with a high share of the vote), even the chaotic party system of Vancouver can stabilize.

While this dissertation could, at some level, be seen as a large comparative case study of the effects of higher offices on the stability of MNTV systems, further research is required to definitively say that MNTV will cause instability unless it is deeply embedded in the party system of a higher office. One possible avenue forward would be to compare multiple cities and their party systems across British Columbia. British Columbian towns all use MNTV to elect their councils and many have endemic political parties, albeit generally fewer than Vancouver. If one were to control for various confounding variables such as population, a cross-sectional study of British Columbian mayoral and city council elections could confirm that up-ballot and uphill races do indeed mitigate the chaos of MNTV elections. Building off of this dissertation in this direction—studying the effects of other offices and other elections on the effects of MNTV—would not merely provide a new dimension to the academic understanding of MNTV. It would also provide context and deeper understanding for reformers seeking to abolish MNTV to increase political competition and improve partisan and ethnic representation.

References

Ahlmann-Eltze, Constantin and Indrajeet Patil. 2021. "ggsignif: Significance Brackets for ggplot2.".

Aldrich, J. H. 1995. *Why Parties?: The Origin and Transformation of Political Parties in America*. Chicago: University of Chicago Press.

Allaire, J.J. and François Chollet. 2021. "keras: R Interface to 'Keras'.".

Allaire, J.J. and Yuan Tang. 2021. "tensorflow: R Interface to 'TensorFlow'.".

Amari, Shun-ichi. 1993. "Backpropagation and Stochastic Gradient Descent Method." *Neurocomputing* 5(4):185–196.

Analytics, Revolution and Steve Weston. 2020. "iterators: Provides Iterator Construct.".

Appiah, Albert Yaw, Xinghua Zhang, Ben Beklisi Kwame Ayawli and Frimpong Kyeremeh. 2019. "Long Short-Term Memory Networks Based Automatic Feature Extraction for Photovoltaic Array Fault Diagnosis." *IEEE Access* 7:30089–30101.

Bache, Stefan Milton and Hadley Wickham. 2020. "magrittr: A Forward-Pipe Operator for R.".

Barnea, Shlomit and Gideon Rahat. 2011. "'Out with the Old, in with the "New"': What Constitutes a New Party?" *Party Politics* 17(3):303–320.

Bates, Douglas, Katharine M. Mullen, John C. Nash and Ravi Varadhan. 2014. "minqa: Derivative-free Optimization Algorithms by Quadratic Approximation.".

Bates, Douglas, Martin Mächler, Ben Bolker and Steve Walker. 2015. "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software* 67(1):1–48.

Bates, Douglas and Martin Maechler. 2021. "Matrix: Sparse and Dense Matrix Classes and Methods.".

Ben-Shachar, Mattan S., Daniel Lüdecke and Dominique Makowski. 2020. "effectsize: Estimation of Effect Size Indices and Standardized Parameters." *Journal of Open Source Software* 5(56):2815.

Betz, Hans-Georg. 1990. "Politics of Resentment: Right-Wing Radicalism in West Germany." *Comparative Politics* 23(1):45–60.

Betz, Hans-Georg. 1994. *Radical Right-Wing Populism in Western Europe*. New York: St. Martin's Press.

Binder, Leonard. 1966. Political Recruitment and Participation in Egypt. In *Political Parties and Political Development*, ed. Joseph LaPalombara and Myron Weiner. Princeton: Princeton University Press pp. 217–240.

Bolleyer, Nicole. 2013. *New Parties in Old Party Systems: Persistence and Decline in Seventeen Democracies*. Oxford: Oxford University Press.

Bottou, Léon. 1991. Stochastic Gradient Learning in Neural Networks. In *Proceedings of Neuro-Nımes '91, Fourth International Workshop Neural Networks and Their Applications*. Nîmes: pp. 12–24.

Bowler, Shaun, Todd Donovan and David Brockington. 2003. *Electoral Reform and Minority Representation: Local Experiments with Alternative Elections*. Columbus, OH: Ohio State University Press.

Brunson, Jason Cory. 2020. "`ggalluvial`: Layered Grammar for Alluvial Plots." *Journal of Open Source Software* 5(49):2017.

Brunson, Jason Cory and Quentin D. Read. 2020. "`ggalluvial`: Alluvial Plots in `ggplot2`.".

Bryan, Jennifer. 2016. "`cellranger`: Translate Spreadsheet Cell Ranges to Rows and Columns.".

Buelens, Jo and Airo Hino. 2008. The Electoral Fate of New Parties in Government. In *New Parties in Government: In Power for the First Time*, ed. Kris Deschouwer. London: Routledge pp. 157–174.

Buelens, Jo and Ruth Van Dyck. 1998. Regionalist parties in French-speaking Belgium: the Rassemblement Wallon and the Front Démocratique des Francophones. In *Regionalist Parties in Western Europe*, ed. Lieven De Winter and Huri Türsan. London: Routledge pp. 51–69.

Canty, Angelo and B. D. Ripley. 2021. "`boot`: Bootstrap R (S-Plus) Functions.".

Chambers, William N. 1966. Parties and Nation-Building in America. In *Political Parties and Political Development*, ed. Joseph LaPalombara and Myron Weiner. Princeton: Princeton University Press pp. 79–106.

Chan, Chung-hong, Geoffrey C. H. Chan, Thomas J Leeper and Jason Becker. 2021. "rio: A Swiss-army Knife for Data File I/O.".

Chang, Winston. 2020. "R6: Encapsulated Classes with Reference Semantics.".

Chen, Peipei, Wei Dong, Xudong Lu, Uzay Kaymak, Kunlun He and Zhengxing Huang. 2019. "Deep Representation Learning for Individualized Treatment Effect Estimation Using Electronic Health Records." *Journal of Biomedical Informatics* 100:103303.

Clark, Alistair. 2021. "The Effects of Electoral Reform on Party Campaigns, Voters and Party Systems at the Local Level: from Single Member Plurality to the Single Transferable Vote in Scotland." *Local Government Studies* 47(1):79–99.

Crisp, Brian F. and Betul Demirkaya. 2019. "Strategic Entry and Strategic Voting in Majoritarian Systems." *Journal of Politics* 82(1):57–71.

Csárdi, Gábor. 2019. "pkgconfig: Private Configuration for R Packages.".

Csárdi, Gábor. 2021. "crayon: Colored Terminal Output.".

Csárdi, Gábor, Kuba Podgórski and Rich Geldreich. 2020. "zip: Cross-Platform 'zip' Compression.".

Dahl, David B., David Scott, Charles Roosen, Arni Magnusson and Jonathan Swinton. 2019. "xtable: Export Tables to LaTeXor HTML.".

Davison, A. C. and D. V. Hinkley. 1997. *Bootstrap Methods and their Applications*. Cambridge: Cambridge University Press.

De Veaux, Richard D. and Lyle H. Ungar. 1994. Multicollinearity: A Tale of Two Nonparametric Regressions. In *Selecting Models from Data*, ed. Peter C. Cheeseman and R. Wayne Oldford. New York, NY: Springer pp. 393–402.

De Winter, Lieven. 1995. Regionalist Parties in Belgium: The Rise, Victory and Decay of the Volksunie. In *Non-State Wide Parties in Europe*, ed. Lieven De Winter. Barcelone: Institut de Ciences Politiques i Socials pp. 23–70.

Dowle, Matt and Arun Srinivasan. 2021. "data.table: Extension of 'data.frame'.".

Duverger, Maurice. 1955. *Political Parties*. New York: John Wiley and Sons, Inc.

Eddelbuettel, Dirk. 2013. *Seamless R and C++ Integration with Rcpp*. New York: Springer.

Eddelbuettel, Dirk and James Joseph Balamuta. 2018. "Extending R with C++: A Brief Introduction to Rcpp." *The American Statistician* 72(1):28–36.

Eddelbuettel, Dirk and Romain François. 2011. "Rcpp: Seamless R and C++ Integration." *Journal of Statistical Software* 40(8):1–18.

Emerson, Rupert. 1966. Parties and National Integration in Africa. In *Political Parties and Political Development*, ed. Joseph LaPalombara and Myron Weiner. Princeton: Princeton University Press pp. 267–302.

Fiva, Jon H. and Olle Folke. 2016. "Mechanical and Psychological Effects of Electoral Reform." *British Journal of Political Science* 46(2):265–279.

Fox, John and Sanford Weisberg. 2019. *An R Companion to Applied Regression*. 3rd ed. Thousand Oaks CA: Sage.

Fox, John, Sanford Weisberg and Brad Price. 2020. "carData: Companion to Applied Regression Data Sets.".

Gagolewski, Marek. 2020. "R Package stringi: Character String Processing Facilities.".

Gaslam, Brodie. 2021. "fansi: ANSI Control Sequence Aware String Functions.".

Genz, Alan and Frank Bretz. 2009. *Computation of Multivariate Normal and t Probabilities*. Lecture Notes in Statistics Heidelberg: Springer-Verlag.

Genz, Alan, Frank Bretz, Tetsuhisa Miwa, Xuefei Mi, Friedrich Leisch, Fabian Scheipl and Torsten Hothorn. 2020. "mvtnorm: Multivariate Normal and *t* Distributions.".

Gilbert, Paul and Ravi Varadhan. 2019. "numDeriv: Accurate Numerical Derivatives.".

Graves, Alex. 2012. *Supervised Sequence Labelling with Recurrent Neural Networks: Studies in Computational Intelligence*. Berlin: Springer-Verlag.

Grodzins, Morton. 1966. Political Parties and the Crisis of Succession in the United States. In *Political Parties and Political Development*, ed. Joseph LaPalombara and Myron Weiner. Princeton: Princeton University Press pp. 303–330.

Hartig, Florian. 2021. "DHARMa: Residual Diagnostics for Hierarchical (Multi-Level/Mixed) Regression Models.".

Henry, Lionel and Hadley Wickham. 2020. "purrr: Functional Programming Tools.".

Henry, Lionel and Hadley Wickham. 2021*a*. "lifecycle: Manage the Life Cycle of your Package Functions.".

Henry, Lionel and Hadley Wickham. 2021*b*. "rlang: Functions for Base Types and Core R and 'Tidyverse' Features.".

Henry, Lionel and Hadley Wickham. 2021*c*. "tidyselect: Select from a Set of Strings.".

Hester, Jim. 2020. "glue: Interpreted String Literals.".

Hester, Jim, Kirill Müller, Kevin Ushey, Hadley Wickham and Winston Chang. 2021. "withr: Run Code 'With' Temporarily Modified Global State.".

Hinojosa, Magda and Susan Franceschet. 2012. "Separate but Not Equal: The Effects of Municipal Electoral Reform on Female Representation in Chile." *Political Research Quarterly* 65(4):758–770.

Hlavac, Marek. 2018. "stargazer: Well-Formatted Regression and Summary Statistics Tables.".

Hochreiter, Sepp and Jürgen Schmidhuber. 1997. "Long Short-Term Memory." *Neural Computation* 9(8):1735–1780.

Hocking, Toby Dylan. 2021. "directlabels: Direct Labels for Multicolor Plots.".

Hug, Simon. 2001. *Alterning Party Systems: Strategic Behavior and the Emergence of New Political Parties in Western Democracies.* Ann Arbor: The University of Michigan Press.

Husbands, Christopher T. 1988. "Extreme Right-Wing Politics in Great-Britain: The Recent Marginalization of the National Front." *West European Politics* 11(2):65–79.

Husbands, Christopher T. 1992. "The Other Face of 1992: The Extreme-Right Explosion in Western Europe." *Parliamentary Affairs* 45(3):267–284.

Ignazi, Piero. 1996. "The Crisis of Parties and the Rise of New Political Parties." *Party Politics* 2(4):549–566.

Inglehart, Ronald and Rudy B. Andeweg. 1993. "Change in Dutch Political Culture: A Silent or a Silenced Revolution?" *West European Politics* 16(3):345–361.

Janda, Kenneth. 1970. *A Conceptual Framework for the Comparative Analysis of Political Parties.* SAGE Publications.

Johansson, Fredrik, Uri Shalit and David Sontag. 2016. Learning Representations for Counterfactual Inference. In *Proceedings of the International Conference on Machine Learning.* pp. 3020–3029.

Johnson, Steven G. N.d. "The `NLopt` Nonlinear-optimization Package." . Forthcoming.

Kassambara, Alboukadel. 2020. "`ggpubr`: `ggplot2` Based Publication Ready Plots.".

Kassambara, Alboukadel. 2021. "`rstatix`: Pipe-Friendly Framework for Basic Statistical Tests.".

Key Jr., Valdimer Orlando. 1942. *Politics, Parties, and Pressure Groups.* New York: Crowell.

Key Jr., Valdimer Orlando. 1955. "A Theory of Critical Elections." *The Journal of Politics* 17(1):3–18.

Key Jr., Valdimer Orlando. 1959. "Secular Realignment and the Party System." *The Journal of Politics* 21(2):198–210.

Kirchheimer, Otto. 1966. The Transformation of the Western European Party Systems. In *Political Parties and Political Development*, ed. Joseph LaPalombara and Myron Weiner. Princeton: Princeton University Press pp. 177–200.

Kitschelt, Herbert. 1989. *The Logics of Party Formation: Ecological Politics in Belgium and Wester Germany.* Ithica: Cornell University Press.

Krouwel, André and Paul Lucardie. 2008. "Waiting in the Wings: New parties in the Netherlands." *Acta Politica* 43(2-3):278–307.

Kuznetsova, Alexandra, Per B. Brockhoff and Rune H. B. Christensen. 2017. "`lmerTest` Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software* 82(13):1–26.

Lang, Michel and R Core Team. 2020. "`backports`: Reimplementations of Functions Introduced since R.".

LaPalombara, Joseph and Myron Weiner. 1966. The Origin and Development of Political Parties. In *Political Parties and Political Development*, ed. Joseph LaPalombara and Myron Weiner. Princeton: Princeton University Press pp. 3–42.

Leal, David L., Valerie Martinez-Ebers and Kenneth J. Meier. 2004. "The Politics of Latino Education: The Biases of At-Large Elections." *Journal of Politics* 66(4):1224–1244.

Lenth, Russell V. 2018. "`estimability`: Tools for Assessing Estimability of Linear Predictions.".

Lenth, Russell V. 2021. "`emmeans`: Estimated Marginal Means, aka Least-Squares Means.".

Levi, Margaret and Michael Hechter. 1985. A Rational Choice Approach to the Rise and Decline of Ethnoregional Political Parties. In *New Nationalism of the Developed West*, ed. Edward A. Tiryakian and Ronald Rogowski. Boston: Allen and Unwin pp. 128–146.

Lucardie, Paul. 2000. "Prophets, Purifiers and Prolocutors: Towards a Theory on the Emergence of New Parties." *Party Politics* 6(2):175–185.

Lüdecke, Daniel. 2018*a*. "`ggeffects`: Tidy Data Frames of Marginal Effects from Regression Models." *Journal of Open Source Software* 3(26):772.

Lüdecke, Daniel. 2018*b*. "`sjmisc`: Data and Variable Transformation Functions." *Journal of Open Source Software* 3(26):754.

Lüdecke, Daniel. 2020. "`sjlabelled`: Labelled Data Utility Functions.".

Lüdecke, Daniel. 2021*a*. "`sjPlot`: Data Visualization for Statistics in Social Science.".

Lüdecke, Daniel. 2021*b*. "`sjstats`: Statistical Functions for Regression Models.".

Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil and Dominique Makowski. 2020. "`parameters`: Extracting, Computing and Exploring the Parameters of Statistical Models using R." *Journal of Open Source Software* 5(53):2445.

Lüdecke, Daniel, Mattan S. Ben-Shachar, Indrajeet Patil, Philip Waggoner and Dominique Makowski. 2021. "Assessment, Testing and Comparison of Statistical Models using R." *Journal of Open Source Software* 6(59):3112.

Lüdecke, Daniel, Philip Waggoner and Dominique Makowski. 2019. "`insight`: A Unified Interface to Access Information from Model Objects in R." *Journal of Open Source Software* 4(38):1412.

Madden, Wayne D. 2003. *Vancouver's Elected Representatives.* Vancouver: City of Vancouver Archives.

Mair, Peter. 1999. "New Political Parties in Established Party Systems: How Successful are they?" *Elites, Parties and Democracy: festschrift for Professor Mogens N. Pedersen* 225:207.

Makowski, Dominique, Mattan S. Ben-Shachar and Daniel Lüdecke. 2019. "`bayestestR`: Describing Effects and their Uncertainty, Existence and Significance within the Bayesian Framework." *Journal of Open Source Software* 4(40):1541.

Malhotra, Pankaj, Lovekesh Vig, Gautam Shroff and Puneet Agarwal. 2015. Long Short -Term Memory Networks for Anomaly Eetection in Time Series. In *Proceedings of the 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning.* Bruges: pp. 89–94.

Mao, Wentao, Jianliang He, Jiamei Tang and Yuan Li. 2018. "Predicting Remaining Useful Life of Rolling Bearings Based on Deep Feature Representation and Long Short-Term Memory Neural Network." *Advances in Mechanical Engineering* 10(12).

Matland, Richard and Deborah Dwight Brown. 1992. "District Magnitude's Effect on Female Representation in U.S. State Legislatures." *Legislative Studies Quarterly* 17(4):469–492.

Microsoft and Steve Weston. 2020. "`foreach`: Provides Foreach Looping Construct.".

Müller, Kirill. 2021. "`hms`: Pretty Time of Day.".

Müller, Kirill and Hadley Wickham. 2021*a*. "`pillar`: Coloured Formatting for Columns.".

Müller, Kirill and Hadley Wickham. 2021*b*. "`tibble`: Simple Data Frames.".

Müller-Rommel, Ferdinand. 1995. Ethno-regionalist Parties in Western Europe: Ampirical Evidence and Theoretical Considerations. In *Non-State Wide Parties in Europe*, ed. Lieven De Winter. Barcelone: Institut de Ciences Politiques i Socials pp. 179–196.

Müller-Rommel, Ferdinand. 1998. Ethnoregionalist Parties in Western Europe: Theoretical Considerations and Framework of Analysis. In *Regionalist Parties in Western Europe*, ed. Lieven De Winter and Huri Türsan. London: Routledge pp. 17–27.

Newell, James L. 1998. The Scottish National Party: Development and Change. In *Regionalist Parties in Western Europe*, ed. Lieven De Winter and Huri Türsan. London: Routledge pp. 105–124.

Norris, Pippa. 2006. "The Impact of Electoral Reform on Women's Representation." *Acta Polit* 41:197–213.

Ooms, Jeroen. 2021*a*. "`curl`: A Modern and Flexible Web Client for R.".

Ooms, Jeroen. 2021*b*. "`rsvg`: Render SVG Images into PDF, PNG, PostScript, or Bitmap Arrays.".

Pedersen, Mogens N. 1982. "Towards a New Typology of Party Lifespans and Minor Parties." *Scandinavian Political Studies* 5(1):1–16.

Perry, Patrick O. 2021. "`utf8`: Unicode Text Processing.".

Pinheiro, Jose, Douglas Bates, Saikat DebRoy, Deepayan Sarkar and R Core Team. 2021. "`nlme`: Linear and Nonlinear Mixed Effects Models.".

Plate, Tony and Richard Heiberger. 2016. "`abind`: Combine Multidimensional Arrays.".

Plummer, Martyn, Nicky Best, Kate Cowles and Karen Vines. 2006. "`CODA`: Convergence Diagnosis and Output Analysis for MCMC." *R News* 6(1):7–11.

R Core Team. 2020. "`foreign`: Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ...".

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Robinson, David, Alex Hayes and Simon Couch. 2021. "`broom`: Convert Statistical Objects into Tidy Tibbles.".

Rochon, Thomas R. 1985. "Mobilizers and Challengers: Toward a Theory of New Party Success." *International Political Science Review* 6(4):419–439.

Rosenstone, Steven J., Roy L. Behr and Edward H. Lazarus. 1984. *Third Parties in America*. Princeton: Princeton University Press.

Rule, Wilma. 1990. "Why More Women Are State Legislators. A Research Note." *The Western Political Quarterly* 43(2):437–448.

Rustow, Dankwart A. 1966. The Development of Parties in Turkey. In *Political Parties and Political Development*, ed. Joseph LaPalombara and Myron Weiner. Princeton: Princeton University Press pp. 107–136.

Sanchez, Omar. 2002. "An Open-Ended Transition: The Effects of Electoral Reform in Italy." *Journal of European Area Studies* 10(2):259–281.

Sarkar, Deepayan. 2008. *`Lattice`: Multivariate Data Visualization with R*. New York: Springer.

Sartori, Giaovanni. 1966. European Political Parties: The Case of Polarized Pluralism. In *Political Parties and Political Development*, ed. Joseph LaPalombara and Myron Weiner. Princeton: Princeton University Press pp. 137–176.

Sartori, Giaovanni. 1976. *Parties and Party Systems: A Framework for Analysis*. New York: Cambridge University Press.

Schauberger, Philipp and Alexander Walker. 2020. "`openxlsx`: Read, Write and Edit xlsx Files.".

Siavelis, Peter. 1997. "Continuity and Change in the Chilean Party System: On the Transformational Effects of Electoral Reform." *Comparative Political Studies* 30(6):651–674.

Squire, Peverill. 1992. "Legislative Professionalization and Membership Diversity In State Legislatures." *Legislative Studies Quarterly* 17(1):69–79.

T. J. J., Ryan. 2020. "LSTMs Explained: A Complete, Technically Accurate, Conceptual Guide with Keras." *AnalyticsVidhya* .

Tierney, Luke. 2020. "`codetools:` Code Analysis Tools for R.".

Trounstine, Jessica L. and Melody Ellis Valdini. 2008. "The Context Matters: The Effect of Single Member vs At-Large Districts on City Council Diversity." *American Journal of Political Science* 52(3):554–569.

Urwin, Derek W. 1983. Harbringer, Fossil or Fleabite? "Regionalis" and the West European Party Mosaic. In *Western European Party Systems*, ed. Hans Daalder and Peter Mair. Beverly Hills: Sage pp. 221–256.

Venables, W. N. and B. D. Ripley. 2002. *Modern Applied Statistics with S*. 4th ed. New York: Springer.

Weeks, Liam. 2008. Independents in Government: a Sui Generis Model? In *New Parties in Government: In Power for the First Time*, ed. Kris Deschouwer. London: Routledge pp. 137–156.

Weingessel, Andreas. 2019. "`quadprog`: Functions to Solve Quadratic Programming Problems.".

Welch, Susan. 1990. "The Impact of At-Large Elections on the Representation of Blacks and Hispanics." *The Journal of Politics* 52(4):1050.

Wickham, Charlotte. 2018. "`munsell`: Utilities for Using Munsell Colours.".

Wickham, Hadley. 2007. "Reshaping Data with the `reshape` Package." *Journal of Statistical Software* 21(12):1–20.

Wickham, Hadley. 2011. "The Split-Apply-Combine Strategy for Data Analysis." *Journal of Statistical Software* 40(1):1–29.

Wickham, Hadley. 2016. *`ggplot2`: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wickham, Hadley. 2019. "`stringr`: Simple, Consistent Wrappers for Common String Operations.".

Wickham, Hadley. 2020. "`modelr:` Modelling Functions that Work with the Pipe.".

Wickham, Hadley. 2021*a*. "`ellipsis`: Tools for Working with ....".

Wickham, Hadley. 2021*b*. "`forcats`: Tools for Working with Categorical Variables (Factors).".

Wickham, Hadley. 2021*c*. "`tidyr`: Tidy Messy Data.".

Wickham, Hadley and Dana Seidel. 2020. "`scales`: Scale Functions for Visualization.".

Wickham, Hadley and Evan Miller. 2021. "`haven`: Import and Export 'SPSS', 'Stata' and 'SAS' Files.".

Wickham, Hadley and Jennifer Bryan. 2019. "`readxl`: Read Excel Files.".

Wickham, Hadley, Lionel Henry and Davis Vaughan. 2021. "`vctrs`: Vector Helpers.".

Wickham, Hadley, Max Kuhn and Davis Vaughan. 2020. "`generics`: Common S3 Generics not Provided by Base R Methods Related to Model Fitting.".

Wickham, Hadley, Romain François, Lionel Henry and Kirill Müller. 2021. "`dplyr`: A Grammar of Data Manipulation.".

Wickham, Hadley and Thomas Lin Pedersen. 2019. "`gtable`: Arrange 'Grobs' in Tables.".

Wilke, Claus O. 2020. "`cowplot`: Streamlined Plot Theme and Plot Annotations for `ggplot2`.".

Wood, S. N. 2003. "Thin-plate Regression Splines." *Journal of the Royal Statistical Society (B)* 65(1):95–114.

Wood, S. N. 2004. "Stable and Efficient Multiple Smoothing Parameter Estimation for Generalized Additive Models." *Journal of the American Statistical Association* 99(467):673–686.

Wood, S. N. 2011. "Fast Stable Restricted Maximum Likelihood and Marginal Likelihood Estimation of Semiparametric Generalized Linear Models." *Journal of the Royal Statistical Society (B)* 73(1):3–36.

Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R.* 2nd ed. Chapman and Hall/CRC.

Wood, S. N., N. Pya and B. Saefken. 2016. "Smoothing Parameter and Model Selection for General Smooth Models." *Journal of the American Statistical Association* 111:1548–1575.

Xie, Yihui. 2014. `knitr`: A Comprehensive Tool for Reproducible Research in R. In *Implementing Reproducible Computational Research*, ed. Victoria Stodden, Friedrich Leisch and Roger D Peng. Chapman and Hall/CRC.

Xie, Yihui. 2015. *Dynamic Documents with R and knitr*. 2nd ed. Boca Raton, Florida: Chapman and Hall/CRC.

Xie, Yihui. 2021*a*. "knitr: A General-Purpose Package for Dynamic Report Generation in R.".

Xie, Yihui. 2021*b*. "xfun: Miscellaneous Functions to Support Packages Maintained by 'Yihui Xie'.".

Zeileis, Achim and Gabor Grothendieck. 2005. "zoo: S3 Infrastructure for Regular and Irregular Time Series." *Journal of Statistical Software* 14(6):1–27.

Zeileis, Achim, Jason C. Fisher, Kurt Hornik, Ross Ihaka, Claire D. McWhite, Paul Murrell, Reto Stauffer and Claus O. Wilke. 2020. "colorspace: A Toolbox for Manipulating and Assessing Colors and Palettes." *Journal of Statistical Software* 96(1):1–49.

Zeileis, Achim, Kurt Hornik and Paul Murrell. 2009. "Escaping RGBland: Selecting Colors for Statistical Graphics." *Computational Statistics and Data Analysis* 53:3259–3270.

Zhang, Yu-Dong, Chichun Pan, Junding Sun and Chaosheng Tang. 2018. "Multiple Sclerosis Identification by Convolutional Neural Network with Dropout and Parametric ReLU." *Journal of Computational Science* 28:1–10.