PREFERENCE IDENTIFICATION AND POLITICAL PARTICIPATION IN ALTERNATIVE VOTING SYSTEMS

by

Emilia J. Suggs

A Dissertation Submitted in Partial Fulfillment of the Requirement for the Degree of Doctor of Philosophy in Economics

Middle Tennessee State University

May 2021

Dissertation Committee:

Dr. Daniel Smith, Chair

Dr. Adam Rennhoff

Dr. Steven Sprick Schuster

ACKNOWLEDGEMENTS

No research is accomplished as a solidary endeavor. We are the products of many experiences and conflicts of ideas, and require the help of others to navigate the unfamiliar and arrive at our proper destinations. Throughout this process, I have received a great amount of support and assistance, and wish to properly acknowledge those who have helped me along the way.

I would like to extend my deepest gratitude to my committee chair and advisor, Dr. Daniel Smith, for his advice and guidance in developing this research project. Your expertise and support has been instrumental in my development as a researcher and teacher in economics.

I would like to extend my sincere thanks to Dr. Adam Rennhoff and Dr. Steven Sprick Schuster for serving on my committee and assisting me in the development of this project. In addition, I would like to thank Dr. Art Carden and Dr. Joy Buchanan for their feedback.

My special thanks are extended to the faculty and staff of the Political Economy Research Institute and the Department of Economic and Finance at Middle Tennessee State University. Thank you to all the students who took the time to participate in the experimental sections of my work, and the faculty members who assisted me in finding participants.

Finally, I would like to thank my parents, family, and friends for their support and encouragement throughout this process.

ABSTRACT

The purpose of this dissertation is to explore how modifying the structure of voting systems and procedures can identify and change people's political behavior. By changing the voting rules or election structures faced by individual political agents, we ultimately change the incentive structures that underlie their decision making and should expect changes in political outcomes. The study examines two types of institutional changes and their associated effects on the behavior of political agents. The first study examines the formation of alternative voting rules and their use as a measure of voter preferences. The study defines the rules of a two-stage multivoting system and evaluates the performance of this system against traditional voting mechanism using experimental data gathered from college students. In the second part, the study examines the political entry decisions of political party candidates in state assembly general elections. Two approaches to political entry are presented: the first assumes that the probability of observing a candidate of a specific political party is dependent upon characteristics about a district's election and demographic characteristics, while the other assumes candidates make election choices based on the expected payoff they receive by participating in elections.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND TERMS

1P1V	One-Person, One Vote System
2SMV	Two-Stage Multivoting System
COV	Coefficient of Variation
FWP	Fixed Weight Preferences
IP	Indifferent Preferences
QV	Quadratic Voting
SDP	Strictly Dominating Preferences

CHAPTER 1: INTRODUCTION

Public choice theory as a discipline applies economic theory and methods to analyze the behavior of political agents and their interaction in political institutions. Special attention is given to the economic incentives presented to political agents; voters and politicians do not only consider political choices in regards to their political beliefs, but also consider how choice outcomes affect their personal wellbeing. Neglecting to account for the role of incentives in political science leads to incorrect predictions about political behavior.

The origin of much of the current work in public choice theory begins with the work of Duncan Black during the 1940s and 50s. In his 1948 paper, "On the Rationale of Group Decision-making", Black relates utility-maximization theory to committee voting decisions, establishing the connection between economic and political theory. Black further establishes this linkage in his 1958 book, *The Theory of Committees and Elections*, providing a formalization of political science and analyzing how different voting systems result in different election outcomes.

Another major influence on public choice and social choice theory was Kenneth Arrow. In his 1951 work, *Social Choice and Individual Values*, Arrow examines concepts of social welfare in a formal setting, establishing social choice theory as a discipline. Arrow discusses the difficulties in defining a social choice system that satisfies relatively reasonable conditions, including non-dictatorship, unrestricted domain, independence of irrelevant alternatives, monotonicity, and Pareto efficiency. Much of the terminology related to election theory is established by Arrow's work.

Another key figure in the discipline was Anthony Downs, who is widely credited for his work his work in median voter theory. In his 1957 work, *An Economic Theory of Democracy*, presents a theoretical model of political competition, relating the preferences of voters to the emergence of political parties. In a two-party system, if voter preferences follow a unimodal distribution, a median voter can be identified, and political parties with adopt platforms to appeal to this median voter, each party attempting to capture its largest share of voters. This model is similar to Hotelling's (1929) model of market competition in duopolies, again establishing an informal link between economic and political theory. Down's outlines conditions in which the median voter theorem will not hold and suggests possible empirical tests.

Starting in the 1960s, the works of James Buchanan and Gordon Tullock fully established the field of public choice theory. Buchanan and Tullock's joint 1962 book, *The Calculus of Consent: Logical Foundations of Constitutional Democracy*, directly address their intention to use economic analysis to examine political organization and voting, emphasizing the importance of the individual as a political agent. Using this approach, the "public interest" is not a distinct concept but reflects the aggregate preferences of *individuals*. Independent papers by Buchanan and Tullock further defined public choice theory as a distinct discipline in economics. Buchanan and Lee (1986) address vote buying, one of the first preference-intensity revealing alternative voting systems. Tullock (1965) discusses incumbency as a barrier to entry in models of political competition.

The purpose of this work is to examine elections using a public choice approach. Broadly, the study examines two aspects of voting institutions. The first part of this dissertation examines the formation of alternative voting rules and their use as a measure of voter preferences. The study focuses on examining the rules of a two-stage multivoting system and evaluates the performance of this system against traditional voting mechanism using experimental data gathered from college students. In the second part, the study examines the political entry decisions of political party candidates in state assembly general elections. Two approaches to political entry are presented: the first assumes that the probability of observing a candidate of a specific political party is dependent upon characteristics about a district's election and demographic characteristics, while the other assumes candidates make election choices based on the expected payoff they receive by participating in elections. The outline of the work is as follows.

Chapter 2 examines a method of quantifying voter preferences and behavior using a two-stage multivoting (2SMV) model. The 2SMV model gives voters an endowment of additional votes exceeding the number of policies under consideration in a direct democracy-style election. Voters may freely distribute this endowment to any of the policies up for election. Using the 2SMV mechanism, the study provides a methodology for identifying voter preferences and voting behavior within a staged multivoting system. From this methodology, the study defines three types of voting behavior: policy indifference, strictly-dominating preferences, and fixed-weight preferences.

Using the theoretical framework established in the previous chapter, Chapter 3 uses experimental data collected from college students to evaluate the performance of the two-stage multivoting system in the context of the 2020 Democratic Presidential Primaries, compared to the traditional "one-person, one-vote" (1P1V) system. In this experiment, participants are asked to select the Democratic nominee using both systems. In the 2SMV experiment, participants are given an endowment of 42 votes that can be used to vote for any of the 21 candidates that were running for nomination at the time of the study. In the first stage, participants decide to vote "for" or "against" each of the candidates. In the second stage, they decide how many votes to distribute to each of the candidates. A "candidate score" is constructed for each candidate, the difference in the total number of "for" votes received by the candidate from participants and the aggregate number of "against" votes. The candidate with the highest candidate score is selected as the nominee. The study uses this experimental data to examine political preference intensities and classify voting behaviors differences between political parties and preference groups. In a second part of the chapter, the study compares two-stage multivoting to other alternative voting systems including vote buying (Buchanan and Lee (1986); Hansen (2000)), storable voting (Casella (2003)), and quadratic voting (Weyl (2013)).

Chapter 4 investigates political competition in state assembly elections using two approaches toward modeling political entry. In the first approach, political party entry decisions are influenced by the election, demographic, and economic attributes of election districts. Political party-candidates that "fit" the district's preferences are more likely to enter than others. The second approach models political entry based on the expected payoffs each political party is likely to receive as a result of political party entry. Using these two approaches to modeling political entry, the study estimates the marginal effects that certain district attributes and components of expected payoff have on the probability of entry for three political parties in the United States: Democrats, Republicans, and Independents. Predicted probabilities of entry are tested against observed outcomes in order to evaluate the predictive strength of these two approaches.

CHAPTER 2: IDENTIFYING VOTER PREFERENCES THROUGH TWO-STAGE MULTIVOTING ELECTIONS

One of the primary arguments for the use of democracy as a form of collective decision making is that it is an information-revealing system, used to determine the desires of the collective in an equitable fashion (Black, 1984; Arrow, 1951; Downs, 1957; Young, 1995). Within this framework, the principle of "one person, one vote" (1P1V) serves to ensure that the voting system is equitable between individuals. However, while this mechanism reveals how the majority of voters feel about a certain policy, it does not reveal how intensely voters' value certain policies over others. For a given number of distinct policies in a docket, a voter may value some policies more strongly than others, be indifferent between policies, or simply not care about certain policies altogether. If a voter does not care about the outcome of a certain proposal, having to vote on that policy would be a waste for the individual and to the voting population as a whole. Individuals that vote on policies they are unaffected by do not contribute new information about the preferences of the majority. Under normal circumstances we would expect that these voters refrain from participating in the election. Additionally, the 1P1V system fails to communicate information about how much individuals value some policies over other policies. Even if unconcerned voters refrain from participating, voters with even mild valuations of policies are incentivized to vote. In these regards the 1P1V system introduces potential inefficiency into election outcomes by failing to mitigate the effects of incorrect information on voters' preferences.

Given this concern, a voting system should not only determine how the majority of people value a certain proposal, but reduce the influence of voters who do not care about the outcomes of that proposal. An equitable voting system may not be the most efficient system in the sense of providing information about majority preferences. A voting system should accomplish four things: provide all participants in an election with equal initial representation, provide equal representation at all stages of the voting procedure, incentivize participants to only vote on policies that affect them, and permit participants to express the intensity of their preferences towards certain policies (Schwartz, 2004; Caplan, 2008; Brennan, 2012).

Over the years several voting systems have been designed to reveal the intensity of preferences among individuals. Some of the early literature in preference revealing voting systems focused on vote selling mechanisms. Buchanan and Lee (1986) and Hansen (2000) examine systems of vote selling in which voters with strong preferences may purchase votes from other voters. Under Buchanan and Lee's specification, intensity of preferences is revealed through an individual's choice of whether or not to participate in an election. Assuming there are no transaction costs involved in trades, voters are free to purchase the votes of other voters willing to sell, creating a market for votes. Each individual estimates the value of his vote and is assumed to be willing to sell at any price over his valuation. Vote selling represents an improvement on 1P1V in regards to expressing preference intensity. It is limited, however, in the degree to which a voter can realistically express their interests. Vote selling may also result in some wealthier citizens obtaining greater influence than others since voters draw from personal wealth in order to purchase additional votes.

Weyl (2013) presents a system of quadratic voting (QV) which allows citizens to purchase as many votes as they wish subject to budget constraints, effectively introducing market mechanisms into the voting framework. Under quadratic voting, each individual maximizes his utility from voting with respect to his budget constraint in accordance with a quadratic pricing system. The total cost of buying votes is equivalent to the square of the number of votes purchased. Using their personal wealth, voters choose between purchasing real goods or votes for certain proposals. Individuals that care about certain proposals may purchase additional votes at an increasing total cost, thus increasing their influence for that proposal. As with vote-selling, quadratic voting suffers from potential inequality concerns due to differences in personal wealth. Lalley and Weyl (2017) address this concern by proposing that the quadratic voting mechanism could be altered such that an alternative currency is used as a place of real currency. Under this specification, voters are given endowments of “voice credits” which may be used to purchase votes for certain policies. However, the construction of the quadratic voting system distorts measuring true preference and reduces total representation. Cardenas, Mantilla, and Zarate (2014) further utilize this specification in an experimental study of quadratic voting using evidence from college students.

In contrast to vote purchasing models, Casella (2003) introduced a system that allows votes to be stored over different election cycles. The model considers a single period, single policy election where voters may either vote for or against the proposal. In each period the voter is given one vote that may be used for the given term or saved, forming a budget constraint for the next period. The outcome of the election is determined by majority rule where the option with the greatest number of votes is passed. Under a storable voting system, voters express their preferences for present or future elections through intertemporal choice: either using all votes in the current period or delaying use until subsequent periods.

These voting systems share a common feature of granting individuals more votes, often by redistributing existing votes away from unconcerned voters towards active voters. However, altering the structure of the 1P1V system to grant individuals a larger endowment of votes achieves the same effects. This paper examines a two-stage multivoting (2SMV) configuration for a direct democracy style election with multiple policies under consideration. In this system, voters are given an additional number of votes beyond the number of policies that may be distributed to any of the policies. Given that voters differ in their personal valuations of policies in the docket, they are expected to distribute their votes differently from one another. To preserve equality, they are given the same number of votes. In the first stage of the model, voters decide how many votes to allocate to each policy. During this stage we observe how each voter weights the policies relative to one another. In the second stage they vote on the alternatives inside each policy, often "for" or "against" the proposal. Thus, voters make both global and local choices during an election.

The two-stage multivoting system seeks to balance voter equity and election efficiency. As long as the voters are given the same initial endowment of votes, equality is preserved even as the number of votes extends beyond the number of policies in the docket. 2SMV differs from quadratic voting and storable voting systems since each citizen is endowed with the same number of real votes. At each stage of 2SMV, voters have the same aggregate influence. Voters with strong preferences towards a single policy are free to allocate all their remaining votes to that policy, while voters who value policies equivalently may split the votes evenly across the docket. This property of 2SMV increases the influence of voters with strong preferences towards certain policies while minimizing the influence of unconcerned voters. An additional benefit of this system is that it presents a way to quantify voter preferences and voting behavior.

This study describes the construction of the two-stage multivoting model and the voter's choice problem, emphasizing how 2SMV reveals preferences. Significant discussion is given to the 2SMV systems ability to act as a measuring of quantifying intensities of preference and identifying types of voting behavior.

Theoretical Framework

The rules of a simple majority rule 1P1V system require that each voter receives one vote per policy. The "one person, one vote" system provides a reasonably equitable and simple way of determining how the voting population perceives a given policy; however, it neglects information about the intensity of voter preferences towards individual policies. The two-stage multivoting system gives voters additional votes to be used on the policies of their choice. Under the proper specification, this system reconciles the equity concerns of a vote-selling system while allowing voters to express preference intensities better than in a single-vote system. This section presents a model of voting choice under a 2SMV system, addressing how this system reveals individual preferences and possible voting behaviors we may observe for various sizes of the endowment.

The section proceeds as follows: the first subsection defines the two-stage multivoting mechanism and choice problem, and provides guidelines for 2SMV elections, the second subsection examines how the 2SMV mechanism identifies the intensity of voter preferences, the third subsection defines types of voting behavior likely to occur in 2SMV elections, and the fourth subsection describes preference aggregation and election outcomes in 2SMV systems.

The Two-Stage Multivoting Model

In this model, voters receive an initial endowment of k votes that can be used in an election with p policies in a docket P. Voters may freely distribute votes for each policy v_p in the docket according to their preferences. In the first stage of the model, voters make a global decision regarding how many votes to distribute to each policy. In the second stage, voters use those policy-votes v_p to make local decisions on the secondary-stage alternatives s in each policy.

Choices in the second stage may either be binary or multinary in construction. If there are only two choices within a policy, the 2SMV model restricts voters to allocate their policy votes to only one of choices, allowing otherwise would yield contradictory information on preferences. In the case of multinary alternatives, some information may be gleaned from allowing voters to distribute their inner-policy votes v_p across the alternatives according to their inner-policy preferences, such that there is a unique vote count for each alternative within a policy $v_{p,s}$. Whether or not a free distribution of policy-votes at the alternative level is appropriate depends on the characteristics of the alternatives.

For example, suppose that the inner-policy choice involved selecting between three candidates (Democratic, Republican, and Independent) for a senate seat. The voter may strongly prefer the Democratic candidate, but in the event they were unlikely to win, would prefer the Independent candidate over the Republican candidate. Allowing the voter to distribute their inner-policy votes may yield new information regarding aggregate choices. The election result is that a minority candidate (that would not normally be voted for) could be elected if major party voters do not intensely value their own candidate and are willing to compromise. In contrast, an election planner may want multinary choices to remain mutually exclusive, in which case all policy-votes are distributed to the selected alternative. If inner-policy choice distributions are permitted, the number of votes per policy per alternatives are specified as $v_{p,s}$. Binary second-stage choices are assumed for the remainder of the paper.

General Guideline of Two-Stage Multivoting Elections. A two-stage multivoting system requires that choices be made at both the policy and inter-policy levels given an endowment of votes. This distinguished itself from a general multivoting system in which choices are only made a single level. The two-stage multivoting system requires the following conditions:

1. J participants vote over multiple policies, p , in a docket, P, such that $p > 1$.
2. Each participant j is endowed with k votes that may be used on any policy in P.
3. Endowments are the same across all participants such that $k_j = k_{\neg j}$ for all j.
4. Decisions are made at a global (policy) level and at least one local (inter-policy) level.
5. The number of policy specific votes (votes given to each policy) equals the sum of the inner-policy votes given to each of the secondary-stage alternatives s, such that:
$$\sum_{i=1}^{s} v_{p,s} = v_p$$
6. If facing binary alternatives in a secondary stage, the voter may not distribute votes to both alternatives.
7. Each person must use all their available votes, such that $\sum v_p = k$.

Conditions 1 through 4 establish the system as a two-stage multivoting system. The election must have multiple policies under consideration for there to merit additional votes and requires a tradeoff between policies for the voter. The second condition states that votes may freely distribute votes to any of the policies under consideration. The system is a two-stage multivoting system (as opposed to a singular multivoting system) if choices are made at both the policy level (how to distribute votes between competing policies) and the inner-policy level (how to vote on the alternatives for a particular policy). The system is not a two-stage multivoting system if choices are only made at a single level.

Conditions 5 through 7 restrict voting behavior during the two-stage voting procedure. At the inner-policy level, the number of votes distributed to a policy v_p constrains the number of votes that can be distributed to the alternatives for that policy, $v_{p,s}$. The sum of votes distributed to each alternative must equal the number of policy-votes. If the second stage choices are multinary the voter may distribute votes across the alternatives. If the choices are binary, then the voter is restricted to only vote for one of the alternatives. Finally, the voter must use all their available votes endowed to them. This preserves the equity condition by ensuring that all voters have the same initial representation and maximizes their respective impact on the election.

The first four conditions are necessary for the system to be classified as a two-stage multivoting procedure. In essence, 2SMV requires that choices be made at both global and local levels. However, certain conditions may be modified or included to serve the needs of an election. Below are six recommendations that may be incorporated into the system at the election planner's discretion:

1. The size of the endowment is greater than the number of policies in the docket, such that $k > p$.
2. Election outcomes are determined by majority rule. The collective decision $D_{p,s}$, the choice collective choice of secondary-stage alternative s, for a given policy p is:

$$D_{p,s} = \max(\sum_{j=1}^{J} v_{p,1,j}, \ldots, \sum_{j=1}^{J} v_{p,s,j})$$

3. Votes may not be traded, sold, or carried into subsequent periods.
4. Each person must vote at least once per policy $v_p \geq 1$.
5. In the case of multinary secondary stage choices:
    a. Require voters to only vote for one of the alternatives for a given policy.
    b. Permit voters to distribute votes across the alternatives in p.
6. Either the number of voters J or the number of policies under consideration p be set such that $J \geq p + 1$ to guarantee that a voter with strictly dominating preferences cannot outweigh the preferences of all other indifferent voters.

Recommendation 1 describes the extension of the endowment, such that voters have more votes available to use than there are policies. Should voters value all the policies equally, this recommendation allows them to vote on all policies.

Recommendation 2 addresses the criterion by which proposals are adopted. In a simple majority rule election, the alternative with the greatest number of total votes is adopted. If this is deemed undesirable, the election may stipulate a threshold that total votes must exceed for the alternative to be adopted. For example, the election may have a rule that requires that an alternative receive at least 60% of total votes in order to pass.

Recommendation 3 restricts the use of alternative systems that would result in individual voters holding unequal endowment sizes. This maintains the equity condition between voters.

Recommendation 4 requires the voter to express his preferences towards each of the policies. Without this condition in place voters may not have an incentive to distribute votes towards certain policies. The benefit of putting this condition in place is that it allows us to derive what the election outcomes would have been in a 1P1V system. In a 1P1V system, voters have to vote on each of the policies if they want to maximize their endowment because they are restricted to one vote for each policy. Assuming binary second stage choices, requiring the voter to vote on each of the policies will result in the same decisions as in the 1P1V system.

Recommendation 5 describes the nature of multinary second stage choices. If multinary alternatives are present for some of the policies, the election should specify whether the voter is permitted to distribute votes among the alternatives or is restricted to select a single alternative.

Recommendation 6 pertains to types of voting behavior likely to arise in two-stage multivoting elections, and limits the ability of "single-issue" voters to dominate the outcomes of elections if all other voters distributed the expected value of voters given an endowment. See Appendix 1 for an extended discussion.

The Voter's Choice Problem. The voter's choice problem occurs in two stages. In the first stage the voter decides how many votes v_p to allocate to each of their policies in order to maximize his utility from voting on P. In the second stage the voters decide to vote "for" or "against" a given policy (if facing binary alternatives), taking the utility maximizing number of votes for that policy as given from the first stage. This second stage is equivalent to selection under a 1P1V system in the binary case. If the voter faces multinary choices, they may divide votes among the alternatives $v_{p,s}$ or vote in favor of a single alternative based on the election guidelines.

Outcomes of the first stage depend on two factors, which may be codependent on one another. The first regards individual preference and ranking of the policies in P. Different assumptions about human behavior define the outcome of the choice problem. The second factor pertains to the number of votes k allotted to each voter. Differing the number of votes changes the optimal allocation dependent on voter preferences.

Measuring the Intensity of Voter Preferences

In a direct democracy setting, voters are tasked with voting on the policies in P. The outcomes of the policies presumably have different impacts on members of the voting base, resulting in some voters having stronger preference intensities than other voters on the same policy. For example, a policy regarding the choice of a school board member would directly affect the guardians of school age children within the district, and either have indirect or no effect on voters without school age children.

Additionally, voters who are not directly (or negatively) affected policy may have subjective preferences that increase their preference intensities towards certain policies. For example, wealthy individuals may vote in favor of a higher income tax with the purpose of increasing the supplementing the incomes of poor individuals because they are motivated by unobserved moral principles or social incentives. By increasing the number of votes available to each person, voters rank policies based on their own judgments, distributing votes according to their individual policy preference intensities. Voters now face a choice problem of how to maximize their utilities from voting through the choice of v_p given the constraint of k.

Among the policies, voters rank the policies according to their valuation of their relative importance. The individual may strictly prefer, prefer, or be indifferent to one policy relative to another. This ranking provides an initial set of decision rules regarding how a voter will allocate the remaining votes after they initially votes on each of the policies. Any allocation chosen by the voter must be consistent with this ranking, satisfying both transitivity and completeness conditions following Arrow (1951).

Given a fixed endowment of additional votes and the voter's subjective ranking, there are a limited number of possible feasible, consistent allocations of votes among policies. The ranking itself establishes the voter's decision rules but is insufficient in determining which of the allocations the voter will choose. An ordinal ranking alone cannot determine the intensity of their preferences; however, the outcomes of a two-stage multivoting model can reveal these preference intensities to the election planner.

Consider the outcome of an individual's choice under an initial endowment of k. By dividing $v_{p,k}$ by k we obtain the percentage of total votes allocated to p, represented by the weight $\alpha_{p,k}$. If it can be assumed that the values of $v_{p,k}$ represent the utility-maximizing quantities of that endowment, we can interpret the values of $\alpha_{p,k}$ as weights reflective of the individual's preference intensity towards a policy, relative to the policies within the docket. It is worth noting that the observed weighting at a given endowment may not hold true over all possible endowments of k, but is expected to converge as values of k become sufficiently large. As the endowment size increases, policy-weight estimates become more precise, resulting in some difference in weight estimates and small and large size endowments.

The merit of a two-stage multivoting system is that it reveals the intensity of preferences among various policies. Individuals that care greatly about a particular policy are free to allocate more votes toward that policy at the costs of the other policies in the docket. However, this mechanism is dependent on the observed relative weighting of each policy being equivalent to the corresponding utility-maximizing vote. If there is reason to believe that the revealed intensity of preference is equal to its true preference, then the individual outcomes of a multivoting system will better reflect voter valuations of these policies.

Suppose that for a given initial endowment, k, there exists a set of utility-maximizing policy weighting vector A^* that reflect the individual's strength of preference intensities towards the policies in P. Voting in the election produces a set of observed voters for each policy $V_k = (v_{1,k}, \dots, v_{p,k})$ and associated weighting vector $A_k = (\alpha_{1,k}, \dots, \alpha_{p,k})$. If revealed preference reflects the true preference intensities a voter has towards the policies in the docket, then the observed set of weights A_k equals the optimal set A^*. That is to say, for all the policies in P, if the endowment size k is sufficiently large the observed 2SMV policy-weight values will equal the voter's true relative preferences towards each policy, such that:

$$\lim_{k \to \infty} \alpha_{p,k} = \alpha_p^*$$

Identifying Voting Behavior in 2SMV Elections

In the two-stage multivoting system (2SMV), intensities of preference are identified by allowing the free distribution of votes from a predetermined endowment. As such, the election planner of a two-stage multivoting election faces a choice of what size endowment voters should receive prior to voting. The choice of endowment size depends on two factors: the number of policies under consideration and the number of voters.

Voting outcomes in any election with a flexible distribution of votes are dependent on the size of the endowment of votes. In traditional, one vote per policy election, the planner views only the qualitative, rather than quantitative, preferences of the voter in regards to the policy. With the condition of inflexible vote distributions removed, increasing the endowment size of votes given to voters allows them to distribute votes according to the intensity of their preferences relative to the set of policies. From this distribution a unique weighting vector can be formed, with each weight-element conditional on the endowment size.

If the 2SMV system is to act as a method of identifying types of voting behavior, the observer must assume these weighting vectors are independent of the endowment size. Violating this condition results in ambiguous preferences, leaving the observer unable to distinguish which vector truly represents the voter's preferences. Preference stability in a 2SMV system requires that the weighting vector must be invariant between endowment sizes. This establishes that the choice of endowment size has no effect on the preferences of individuals. For example, if the participant distributes 30% of their votes to the tax policy for a small endowment size, this should remain unchanged when the endowment size is increased.

Assuming that preferences are stable, the weighting vectors obtained in a 2SMV election equal the true relative weighting vectors for each voter. Information about voter weighting vectors can be used to identify unique types of voting behavior in multivoting elections. For example, there may be certain voters who only care about the outcomes of a single policy. In a "one person, one vote" (1P1V) election it is not possible to identify these types of behavior because the system mechanism does not distinguish the intensity of voter preferences, only the preferred alternative. Since the 2SMV system produces relative weights of voter policy preferences, it is possible to evaluate the degree to which a voter prefers one policy to the others. The study examines three possible types of normal voting behavior: indifferent preferences, fixed-weight preferences, and strictly dominating preferences.

Indifferent preferences (IP) are characterized by an equal division of weight across all the policies in a docket. For any endowment size, the participant attempts to divide votes equally across the policies. If the endowment size is not divisible, it would be assumed that additional votes would be distributed randomly or by an external characteristic, such as the ordering of the policies when presented. Voters with indifferent preferences distribute votes in such a way that the weight given to any policy equals $\alpha_p = 1/P$.

On the other extreme, some voters may only care about the outcomes of a single policy. Voters with *strictly dominating preferences (SDP)* will distribute the entire endowment of votes to the policy that is most highly valued. This results in the highest-valued policy receiving a weight equal to $\alpha_p = 1$ and all other policies receiving a weight equal to $\alpha_{\neg p} = 0$.

Fixed-weight preferences (FWP) are characterized by unequal weights placed on multiple policies in the docket, at least two of which are non-zero. Voters exhibiting this type of behavior rank some policies higher than others, without fully opting out of lower ranked policies. If preferences are stable, these weight values will equal some constant c that will vary across policies, $\alpha_p = \frac{v_p}{k} = c$.

The classification of voting behavior is an expression of the degree of variation present in a participant's weighting vector. The expected weight of any policy in a docket is $1/P$. Calculating the coefficient of variation (CoV) of a person's policy-weighting vector produces a measure of the strength of policy preferences, allowing for the identification of voting behavior. The *policy-weight coefficient of variation (PW_{CoV})* is the ratio of the standard deviation of the weights given to each policy to the mean of those weights. The mean weight of a weighting vector will be the same across all individuals, and its value will equal $1/p$.

$$PWC_j = P\sqrt{\frac{\sum_{p=1}^{P}\left(\alpha_p - \frac{1}{P}\right)^2}{P}}$$

If a voter is indifferent between the policies, each of the policy-weights equals its expected value and the standard deviation of those weights is zero:

$$PWC_{IP} = 0$$

On the other hand, those with strictly dominating preferences will have weight values either equal to zero or one, the two most extreme values weights can take on. The policy-weight coefficient of variation for a person with strictly dominating preferences will equal:

$$PWC_{SDP} = \sqrt{P-1}$$

As weight is divided further across policies, the distance between the weight value and mean necessarily fall, decreasing the standard deviation found within the policy-weight vector. The stronger the preference intensities are for a single policy, the greater the degree of standard deviation. Thus, the policy-weight coefficient of variation of those with fixed weight preferences will fall somewhere between these two extremes:

$$PWC_{FWP} = (0, \sqrt{P-1})$$

The use of the coefficient of variation, as opposed to the standard deviation only, is preferred since we can define a difference between fixed-weight voters with high variation versus low variation in their preferences. Those with highly variant preferences will have a standard deviation in their weights greater than the mean (CoV > 1) while those with more even preferences will have a standard deviation lower than the mean (CoV ≤ 1). A detailed description of the derivation of exact values for policy-weight coefficient of variation calculation can be found in Appendix 2.

Preference Aggregation and Election Outcomes in 2SMV Elections

The result of each voter using the two-stage mechanism is a collection of j policy-vote distributions vectors $V_{p,s}$ of the dimensions $(p \times s) \times 1$. After each voter has made their choices, aggregate policy preferences are calculated from the voting base's vote distribution vectors. Policy-alternative vote totals are aggregated for each alternative a within policy p, by using calculating the total number of votes given to each alternative within the policy. This is expressed as:

$$p_a = \sum_j v_{p,s,j}$$

For example, suppose that there is a three policy election. Each policy has two alternatives, "yes" or "no". When aggregating policy-alternative preferences for one of the policies, the organizer would add all the first-stage votes, voters distributed to that policy depending on whether they selected to vote for or against the policy. The result is two scalars, one that shows the total "yes" votes given to the first policy during the first stage, and the second a value showing aggregate "no" votes for the policy. In total, for a three policy election with two alternatives unique to each policy, there will be six aggregate values, the sum of all votes to an alternative for each policy.

When aggregate policy-alternative preferences have been calculated, the resulting values are used to determine the collective decision of a policy. The exact criteria used to determine which (if any) alternative is preferred depends on the goals outlined by the election planner. For example, this study assumes that only a simple majority need be met to determine the collectively preferred alternative. However, additional conditions, such as obtaining a certain percentage of total policy votes, may also be applied. The collective decision D_p for a given policy p is $D_p = \max(p_1, \ldots, p_s)$, assuming a simple majority rule. The decision vector D is the set of collective decisions D_p for each policy and is of the dimensions $p \times 1$.

A Numeric Example of Collective Decision-making under 1P1V and 2SMV Systems. The following numeric example illustrates the election process in the two-stage multivoting system and compares how collective decisions may differ between 1P1V and 2SMV systems. All the original guidelines in two-stage multivoting rules section hold, and the choices in the second stage are binary. Since the second-stage outcomes are binary, the choices of the second-stage will equal those made in the 1P1V system, allowing for a comparison.

Table 1 shows the outcomes obtained from a two-stage multivoting election with an initial endowment of 9 votes, three policies, and three voters. The first stage outcomes show how each of the three individuals distributes their votes among the three policies. The second stage outcomes show if the individual voted "yes" or "no" regarding the policy. If we modified the system to a 1P1V specification, only the second stage outcomes would be relevant.

Table 1

Decisions in a Two-Stage Multivoting (2SMV) Election

	First-Stage Decisions			Second-Stage Decisions		
	Policy A	Policy B	Policy C	Policy A	Policy B	Policy C
Person 1	9	0	0	Yes	No	No
Person 2	3	3	3	No	No	No
Person 3	0	2	7	No	Yes	Yes

Notes. First three columns represent the first-stage (global) choices of the three voters on each of the three policies. The second set of columns represents the second-stage (local) choices of the three voters on the policies.

From the first-stage choices vote distribution choices, policy-weighting vectors and policy-weight coefficient of variation estimates can be calculated for each of the individuals based on the methodology in the previous section. The policy-weighting vector for individual j in this example will be $\alpha_{p,j} = ({}^{v_{1,j}}/_{k}, {}^{v_{2,j}}/_{k}, {}^{v_{3,j}}/_{k})$. Since the second-stage choices are binary and mutually exclusive, it is not necessary to include the weight value for the alternative that was not selected. From these estimates, voting behavior types can be identified using each person's calculated policy-weight CoV estimate. In a three-policy election with an endowment size of 9 votes, the maximum

CoV value an estimate can take on is 1.41 in the case of strictly dominating preferences and 0 in the case of indifferent preferences. If the person's CoV value is greater than 1 then they exhibit high-variance fixed weight preferences. Values less than or equal to one are characteristic of low- variance fixed weight preferences.

The policy-weighting vector for the first person is $\alpha_{p,1} = (1, 0, 0)$ and has a policy-weight CoV estimate of 1.41. Given the definitions proposed previously, person 1 exhibits strictly dominating preferences towards Policy A. The second person distributes equal quantities of votes towards all three policies. For the second person's policy-weighting vector is $\alpha_{p,2} = (0.33, 0.33, 0.33)$ and has a policy-weight CoV estimate of 0, indicative of indifferent preferences. Finally, the third person distributes most of their votes towards the third policy; however, gives a small quantity of votes towards the second policy. The policy-weighting vector for the first person is $\alpha_{p,1} = (0, 0.22, 0.78)$ and has a policy-weight CoV estimate of 0.98. The third person's behavior is classified as low-variance fixed weight preferences, since the standard deviation of Person 3's weight values is less than the mean weight value.

Table 2 shows a summary of the collective decisions and election outcomes under both the 1P1V and 2SMV systems.

Table 2

Election Outcomes under 1P1V and 2SMV System Specifications

Policy	"Yes" Votes	"Against" Votes	Total Votes	Outcome
	One-Person, One-Vote (1P1V) Outcomes			
Policy A	1	2	3	Unadopted
Policy B	1	2	3	Unadopted
Policy C	1	2	3	Unadopted
	Two-Stage Multivoting (2SMV) Outcomes			
Policy A	9	3	12	Adopted
Policy B	2	3	5	Unadopted
Policy C	7	3	10	Adopted

Since the second-stage choices are mutually exclusive and binary, the decisions in the second stage are equivalent to those of a 1P1V election. In the 1P1V system, none of the policies manage to pass under a single-vote majority rule system. On Policy 1, Person 1 votes in favor of the policy while Persons 2 and 3 vote against the policy. In a 1P1V election, Policy 1 does not pass. On Policy 2, Persons 1 and 3 vote in against the policy while Person 2 votes in favor of the policy. Likewise, Policy 3 fails to pass under majority rule.

The collective decision making in a 1P1V election expresses what choices are most preferred by the voting base; however, conveys no information regarding the strength of those preferences. If information regarding the intensity of voter preferences is considered important, the election results of the two-stage multivoting system are preferential. In the 2SMV system, policies that are not preferred by the majority may be adopted if minority position voters distribute a sufficiently high quantity of their votes towards the policy in the first stage.

Introducing a 2SMV system gives voters the ability to distribute votes based on what they consider to be the best use of their endowment. As shown by the election results in Table 2, some of the policies that were not adopted in the 1P1V system are adopted in the 2SMV format. Policies A and C are adopted in the 2SMV system, while Policy B remains unadopted.

There are two things to note regarding the 2SMV collective decision making found within the example. First, election results can drastically differ between systems when intensities of preference are incorporated into the model. Consider the preferences of voters for Policy A. While a majority of voters chose to vote against A, the strength of preferences of a single voter with strictly dominating preferences was sufficient to win the election. In general terms, this suggests that the outcome of that particular policy was more important to the first voter than it was to the other voters.

A second observation is that in a 2SMV system, the total votes given by both "yes" and "no" voters varies across policies, indicating that on average, voters had stronger aggregate preferences towards some policies over others. For example, Policy A received 12 total votes whereas Policy B received 5 total votes. Within a 2SMV system, the resulting outcomes of a policy decision are based on the voting strategies of the voters who choose to distribute votes to the policy, rather than all voters as in the 1P1V system. If only one person distributes votes to a policy, the collective decision of that policy will be solely based on that voter's choice.

The choice of a voting system depends on the goals of the election planner. Between 1P1V and 2SMV systems, an initial choice must be made regarding whether the collective decisions should be made on what the majority prefers or by the strength of preferences of those opting into the policy. The benefit of a multivoting system is that it introduces information into the system regarding how intensely people value single policies. Ideally, individuals who are significantly affected by the results of a proposal distribute more votes towards that policy. Individuals who are not affected by a proposal distribute their votes towards policies that matter more to them. This mechanism reduces the effect of voters who do not vote on policies, increasing the likelihood that policies with significant effects (but no significant counteracting effects) for certain subgroups are adopted.

Conclusions

Within the context of direct democracy, "one person, one vote" (1P1V) serves as a benchmark for a voting scheme that reveals preferences in an equitable way. However, while it reveals majority preferences for different policies, it does not reveal how intensely voters prefer some policies over others. For certain policies, voters that place a low value on the importance of the policy may be overrepresented and vice versa. This presents a concern regarding the efficiency of 1P1V systems because the intensity of people's preferences is not adequately revealed in elections.

Several intensity revealing voting systems have been suggested to respond to this concern including vote selling, quadratic voting, and storable vote systems. While these systems reveal preference intensities better than 1P1V, increased preference expression often comes at a cost of reduced equity. Motivated by a desire to balance equity and efficiency concerns, this paper examines a two-stage multivoting system (2SMV) designed to reveal the intensities of voter preferences for different policies. Given an initial endowment of votes exceeding the number of proposals to be voted on, the individual first determines how many additional votes to allocate to each policy consistent with their subjective ranking, and then decides on how to vote on the alternatives within a specific policy.

The discussion here provides a framework for a voting scheme that balances efficiency and equity concerns, as well as a methodology for identifying voter preferences within a staged multivoting system. Using this methodology, three types of behavior are defined: policy indifference, strictly-dominating preferences, and fixed-weight preferences. From these definitions behavior types can be classified and analyzed for both individual and group populations, providing insight into the strength of a particular person or group's position values.

References

Arrow, K. J. (1951). *Social Choice and Individual Values*. New Haven: Yale University Press.

Black, D. (1948). On the Rationale of Group Decision-making. *Journal of Political Economy*, 23-34.

Brennan, J. (2012). *The Ethics of Voting*. Princeton: Princeton University Press.

Buchanan, J. M., & Lee, D. R. (1986). Vote buying in a stylized setting. *Public Choice*, 3-15.

Buchanan, J. M., & Tullock, G. (1999). The *Calculus of Consent*. Indianapolis: Liberty Fund, Inc.

Buchanan, J. M. (1954). Individual Choice in Voting and the Market. *Journal of Political Economy*, 334-343.

Caplan, B. D. (2008). *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton: Princeton University Press.

Cardenas, J. C., Mantilla, C., & Zarate, R. D. (2014). Purchasing Votes without Cash: Implementing Quadratic Voting Outside the Lab. Working Paper.

Casella, A. (2005). Storable votes. *Games and Economic Behavior*, 391-419.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 135-150.

Hansen, R. L. (2000). Vote Buying. *California Law Review*, 1323-1372.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 41-57.

Kochin, M. S., & Kochin, L. A. (1998). When Is Buying Votes Wrong? *Public Choice*, 645-662.

Lalley, S., & Weyl, E. G. (2017). Quadratic Voting: How Mechanism Design Can Radicalize Democracy. *American Economic Association Papers and Proceedings*.

Posner, E. A., & Weyl, E. G. (2014). Quadratic Voting as Efficient Corporate Governance. *81 University of Chicago Law Review 251*.

Posner, E. A., & Weyl, E. G. (2014). Voting Squared: Quadratic Voting in Democratic Politics. *Coase-Sandor Institute for Law and Economics* Working Paper No. 657.

Rasmussen, J. (1987). Estimating correlation coefficients: Bootstrap and parametric approaches. *Psychological Bulletin*, 136-139.

Schwartz, B. (2004). The *Paradox of Choice: Why Less is More*. New York: Ecco.

Tullock, G. (1959). Problems of Majority Voting. *Journal of Political Economy*, 571-579.

Weyl, E. G. (2013). Quadratic Vote Buying. Working Paper.

Young, P. (1995). Optimal Voting Rules. *Journal of Economic Perspectives*, 51-64.

Appendix A: Discussion of 2SMV Recommendation 6

Suppose that an election has J voters, P policies, and an endowment size of k. Alternatives, s, are binary for all policies. Preferences are defined such that a voter displays *strictly dominating preferences*, if the policy-weight for a policy is equal to $\alpha_{p,k} = 1$ for the highest-ranked policy, and all other policies receiving a weight equal to $\alpha_{-p,k} = 0$. Voters display *indifferent preferences* if their policy-weight value equals $\alpha_{p,k} = 1/p$ for all policies.

For certain election characteristics, the choice of J or P can result in the decision of a single-voter with dominating preferences counteracting the decision of all other voters with indifferent preferences.

The expected number of votes $v_{p,s,IP}$ by indifferent voters for the first alternative, $(s = 1)$, is:

$$E[v_{p,s,IP}] = 1/P * k * (J - 1)$$

The number of votes given to the second alternative, $(s = 2)$, by the single voter with dominating preferences is:

$$E[v_{p,2,SDP}] = k$$

To prevent the single voter from dominating the all the indifferent voters, the following inequality can be written:

$$E[v_{p,s,IP}] > E[v_{p,2,SDP}]$$

$$1/P * k * (J - 1) > k$$

$$(J - 1) > (k * P) / k$$

$$(P < J - 1) \vee (J > P + 1)$$

By setting the number of policies P to be less than $J-1$, the choices of a voter with dominating preferences cannot overtake those of the remaining voters if they are assumed to have indifferent preferences. With additional assumptions regarding the preferences of the voting populations, this can further be refined.

Appendix B: Discussion of Preference Intensity Coefficient of Variation Classifications

Policy-Weight Coefficient of Variation

Suppose that within an election there are P policies under consideration. The preference intensity of policy p for voter j is represented by the policy-weight α_p, which may take on values equal and between 0 and 1. Given P policies the sum of weight values is equal to 1, such that:

$$\sum_{p=1}^{P} \alpha_p = 1$$

A vector A of the dimensions $(P \times 1)$ consisting of person j's policy-weights can then be constructed. The average value of the policy-weighting vector is then:

$$\mu_A = {}^{\sum_{p=1}^{P} \alpha_p}\!/_{P} = {}^{1}\!/_{P}$$

The standard deviation of A is:

$$\sigma_A = \sqrt{\frac{\sum_{p=1}^{P}\left(\alpha_p - {}^{1}\!/_{P}\right)^2}{P}}$$

Thus the coefficient of variation for the policy-weighting vector A is:

$$PWC_j = {}^{\sigma_A}\!/_{\mu_A}$$

$$PWC_j = P\sqrt{\frac{\sum_{p=1}^{P}\left(\alpha_p - {}^{1}\!/_{P}\right)^2}{P}}$$

Indifferent Preferences

A voter exhibits *indifferent preferences (IP)* if $\alpha_p = 1/P$ for all values of p. The exact value of the policy-weight coefficient of variation for is as follows:

$$PWC_j = P\sqrt{\frac{\sum_{p=1}^{P}\left(\alpha_p - {}^{1}\!/_{P}\right)^2}{P}}$$

From the definition of *indifferent preferences* the policy-weight value for each policy p is equal to $1/p$:

$$= P\sqrt{\frac{\sum_{p=1}^{P}\left({}^{1}\!/_{P} - {}^{1}\!/_{P}\right)^2}{P}}$$

$$= P\sqrt{\frac{0}{P}}$$

$$PWC_{IP} = 0$$

Strictly Dominating Preferences

A voter exhibits *strictly dominating preferences (SDP)* if $\alpha_p = 1$ for the highest ranked policy in P and thus $\alpha_p = 0$ for all other p. The exact value of the policy-weight coefficient of variation for is as follows:

$$PWC_{CoV} = P\sqrt{\frac{\sum_{p=1}^{P}\left(\alpha_p - {}^{1}\!/_{P}\right)^2}{P}}$$

Let $p = 1$ represent the policy that is rated highest for a person with strictly dominating preferences. Since $\sum_{p=1}^{P} \alpha_p = 1$, the policy-weight values for all other policies in P must equal 0. Thus, the exact value of the policy-weight coefficient of variation for a person with strictly dominating preferences must equal:

$$= P\sqrt{\frac{\left(1 - {}^{1}\!/_{P}\right)^2 + \sum_{p=2}^{P}\left(0 - {}^{1}\!/_{P}\right)^2}{P}}$$

$$= P\sqrt{\frac{\left(1-{}^{1}/_{P}\right)^2+(P-1)\left(-{}^{1}/_{P}\right)^2}{P}}$$

$$PWC_{SDP} = \sqrt{P-1}$$

Fixed Weight Preferences

Fixed-weight preference are defined as preference intensities that converge to exact values at sufficiently large endowment sizes; however, $\alpha_p \neq 1$ for any p or $\alpha_p \neq \frac{1}{P}$ for all p. If preferences are stable, these weight values will equal some constant c that will vary across policies, $\alpha_p = \frac{v_p}{k} = c$.

Suppose the following the policy-weight vectors, $A_{j,p}$, in a P policy election for two voters, 1 & 2, are as follows:

$$A_{1,p} = a_{1,1}, a_{1,2}, \ldots, a_{1,p}$$

$$A_{2,p} = a_{2,1}, a_{2,2}, \ldots, a_{2,p}$$

Let m represent the subscript of the policy with the largest weight value in $A_{j,p}$. Since $\sum_{p=1}^{P} \alpha_p = 1$, then $\sum_{p \neq m}^{P} \alpha_p = 1 - \alpha_{j,m}$.

Assume that $\max(\alpha_{1,p}) > \max(\alpha_{2,p})$, then $[1-\max(\alpha_{1,p})] < [1-\max(\alpha_{2,p})]$.

In an election with P policies, the expected weight that each policies receives is $\alpha_{j,p} = {}^{1}/_{P}$. Since $\alpha_{j,p} = [0,1]$ the following must hold true about the squared difference between $\alpha_{j,p}$ and the expected value:

$$(\max(\alpha_{1,p}) - {}^{1}/_{P})^2 > (\max(\alpha_{2,p}) - {}^{1}/_{P})^2$$

$$([1 - \max(\alpha_{1,p})] - {}^{1}\!/_{P})^2 > ([1 - \max(\alpha_{2,p})] - {}^{1}\!/_{P})^2$$

Thus the following holds true about the sum of squared differences for the two voters:

$$(\max(\alpha_{1,p}) - {}^{1}\!/_{P})^2 + ([1 - \max(\alpha_{1,p})] - {}^{1}\!/_{P})^2$$
$$> (\max(\alpha_{2,p}) - {}^{1}\!/_{P})^2 + ([1 - \max(\alpha_{2,p})] - {}^{1}\!/_{P})^2$$

Let SSD_1 represent the sum of squared differences for voter 1 and SSD_2 represent the sum of squared differences for voter 2. The following inequality holds regarding the voters' policy-weight coefficient of variation estimates:

$$P\sqrt{\frac{SSD_1}{P}} > P\sqrt{\frac{SSD_2}{P}}$$

$$PWC_1 > PWC_2$$

Thus, the maximum value a policy-weight coefficient of variation estimate can take on is in the case of strictly-dominating preferences, where $\max(\alpha_{j,p}) = 1$ and $PWC_{SDP} = \sqrt{P-1}$. The minimum estimate occurs where $\alpha_{j,p} = {}^{1}\!/_{P}$ for all values of $\alpha_{j,p}$. Since the definition of fixed-weight preferences states $\alpha_p \neq 1$ for any p or $\alpha_p \neq \frac{1}{P}$ for all p, the policy-weight coefficient of variation value will fall between the minimum and maximum values, such that:

$$PWC_{FWP} = (0, \sqrt{P-1})$$

CHAPTER 3: EVALUATING THE PERFORMANCE OF TWO-STAGE MULTIVOTING ELECTIONS: EXPERIMENTS IN THE PREFACE OF THE 2020 PRIMARIES

The preceding chapter described the individual's choice problem under a two-stage multivoting framework, defined types of voting behavior likely to emerge, and described a method of classifying types of behavior based on the variation within each individual's policy-weighting matrix. Having examined the methodology behind individual and aggregate choices within the 2SMV system, the study now evaluates the performance of the system using experimental data collected from college students at Middle Tennessee State University.[1]

Participants were asked to vote for the 2020 U.S. Democratic Party presidential candidate using both the 1P1V and 2SMV format. When voting in the 2SMV format, participants make two choices: first, how many votes from the endowment to assign to each of the viable candidates, and second to vote "for" or "against" each of the candidates. Allowing participants to vote both "for" and "against" candidates allows participants of all political party affiliations to vote in the study, and turns the primary into a multi-staged choice problem.

While the two-stage multivoting system is defined under the assumption of directly voting on policies, the mechanism itself can be applied to representative elections provided they meet the conditions outlined in the rules stated in the guidelines.

[1] Due to the sensitive nature of the questions asked in this study, survey respondents were assured raw data would remain confidential and would not be shared, in accordance with the institution's IRB protocols.

Specifically, there must be both at least one global level choice and local level choices. Within the context of this experiment, the global level choice is how to distribute votes to each of the candidates, and the local level choices are in regards to whether to vote "for" or "against" each candidate. The "collective decision" for a candidate is calculated as the difference in total "for" and "against" votes, or a 2SMV candidate score. Thus, the 2SMV candidate score for candidate c is:

$$2SMV\ CS_c = \sum_{j=1}^{J}(v_{c,"for"}) - \sum_{j=1}^{J}(v_{c,"against"})$$

The winner of the nomination is the candidate that secures the highest 2SMV candidate score.

Characteristics of the 2020 Democratic primaries at the time of the survey fit well within the guidelines of the 2SMV system, allowing it to take full advantage of the system. In 2SMV, the free endowment of votes allows voters to choose how much weight they want to give each policy, creating a tradeoff between policy decisions. These decisions are reflective of the types of policies that appear in the docket. If policies offer relatively similar outcomes, voters are likely to exhibit indifferent preferences, since the personal stakes are the same regardless of choice. In this experiment, the first-stage choice of policy consists of how to distribute votes to each of the candidates. At the time of the study, twenty-one candidates had officially declared their intent to run in the Democratic primaries, producing a crowded field of candidates of different personal characteristics and political positions. This diversity of candidates, in terms of both personal characteristics and political positions, creates stronger tradeoffs in the "policy" level choice. By exclusively voting for one candidate over another, one is giving up a different set of characteristics unique to that candidate (Hotelling, 1929; Downs, 1957). For example, voting for Joe Biden implies a different set of policies than does voting for Bernie Sanders; the two candidates are not identical. As such, individual rankings of candidates are likely to significantly vary between participants. While the 1P1V system can only detect preferences for each participant's first choice candidate, the 2SMV system can capture information related to participant's secondary choices.

The chapter proceeds as follows: the first section provides a description of the experiment and relevant summary statistics pertaining to the sample, the second section identifies and analyzes intensity of preferences within the sample and subsamples, the third section compares mock election results in the two voting formats, and the fourth examines relationships between behavior types and election results.

Description of Experiment and Summary Statistics

While the 2SMV system is treated as a preference allocation mechanism in a direct democracy styled election, the mechanism can be applied to alternative uses as well. The two essential characteristics that an election must have to use the 2SMV system are: (1) multiple, competing policies in the docket, and (2) the presence of at least two alternatives unique to a policy. In this experiment, respondents were asked to vote in a mock version of the 2020 Democratic Presidential primaries using both the 1P1V and 2SMV mechanisms. The global choice, or "policy" choice, in this instance is the distribution of votes towards each Democratic candidate. The local, or "alternative" level, choices are to vote "for" or "against" each candidate.

The additional ability to vote "against" candidates allows respondents of all political preferences to participate in the mock primaries, regardless of if they would normally vote for a Democratic candidate. Thus, each candidate can potentially receive positive, negative, or no votes from respondents. A candidate score can be constructed by subtracting the number of “against” votes from the number “for” votes a candidate received, showing the net support for each candidate. In a 1P1V system a candidate is nominated if they secure the largest percentage of votes from voters. In this construction of a 2SMV system it is the candidate that has the highest candidate score, or net difference in positive and negative votes, that secures the nomination.

Responses were collected during the period of June 2019 through October 2019. At the start of the second experiment, respondents were given a list of the potential Democratic nominees. At the time the survey received IRB approval, twenty-one individuals were registered as potential nominees. Some of the potential nominees entered and dropped out of the primary race during the duration of the study; however, only the candidates listed in the initial IRB application were presented to survey respondents. Given that the purpose of this experiment is to provide feedback on the 2SMV mechanism in a contemporary setting rather than to provide insight into the 2020 Democratic Primaries, the entry and exit of nominees from the field does not pose any threats to the interpretation of experimental results.

In this survey, respondents voted on candidates using both the 1P1V system and 2SMV system respectively. Respondents first selected the candidate from the list of potential nominees they would vote for if the election were held that day, representing their choice in a 1P1V style election. After making their selection using the 1P1V system, respondents were asked to vote on candidates using the 2SMV system. During this part, respondents were given an endowment of 42 votes (twice the number of candidates) to distribute across the twenty-one candidates. Respondents first distributed votes to each of the candidates, and then selected whether they would vote “for” or “against” each of the candidates. Following the close of the survey, responses that were either incomplete or did not follow the directions correctly were removed from the sample.

At the start of the survey, participants were asked to report their political party affiliation, fiscal preferences, and social preferences. Table 3 shows the frequency and percentage of the sample corresponding to the reported political preferences of participants in the survey.

The study includes sixty-two observations that were both fully complete and used the methodology correctly. Participants were asked to report their political preferences for three categories: political party affiliation, fiscal policy preferences, and social policy preferences.

Table 3

Descriptive Statistics of the Sample

Category	Frequency and Percentage of Sample
Political Party Affiliation	
Democratic	20 (32.3%)
Republican	24 (38.7%)
Libertarian	2 (3.2%)
Other	16 (25.8%)
Fiscal Preferences	
Very Conservative	13 (20.9%)
Somewhat Conservative	20 (32.3%)
Neutral	19 (30.7%)

Table 4

Category	Frequency and Percentage of Sample
Somewhat Liberal	6 (9.7%)
Very Liberal	4 (6.5%)
Social Preferences	
Very Conservative	10 (16.1%)
Somewhat Conservative	14 (22.6%)
Neutral	9 (14.5%)
Somewhat Liberal	10 (16.1%)
Very Liberal	19 (30.7%)
Number of Observations: 62	

In the study, participants were asked to state their political party affiliation given five options: Democratic, Republican, Libertarian, Green, and Other. Rows numbers 1 through 4 of Table 3 show the frequencies of each political party type within the sample. Twenty participants identified as Democrats, twenty-four as Republicans, two as Libertarians, and sixteen as other. None of the participants identified as Green Party members. About an equal number of participants identified as Democrats (32.3%) and Republicans (38.7%).

Participant knowledge and preferences for candidates are likely to differ significantly depending on the participant's party affiliation. For example, candidates who identify as Democrats are more likely to have favorable views of their top ranked candidates than Republican participants. Additionally, they are more likely to have better information regarding less well-known candidates on the ballot. In a conventional, 1P1V format, where participants vote in favor of candidates, Democratic participants have the greatest incentive to vote in the 2020 Primaries, making their choices in the study most realistic since they represent future choices. However, since the 2SMV setting described in the experiment allows participants to vote “for” and “against” candidates, participants of all party affiliations can participate under realistic conditions. While other types of participants may presumably not have an incentive to vote in favor of a candidate, they may have an incentive to vote against certain candidates to make their own party choice most likely to be elected in the general elections.

Fiscal and social preferences were measured on a Likert-scale with options ranging from “Very Conservative” to “Very Liberal”. At the start of the survey, participants were provided with study definitions of fiscal and social conservatism and liberalism. They were then asked to rank their preferences according to their self-perception. The questions and definitions provided to participants can be found in Appendix A.

Fiscal preferences in terms of the study refer to the participant’s beliefs regarding the scope of government intervention in economic activity. Rows 5 through 9 show the fiscal preferences of participants in the sample. According to the survey thirteen respondents identify as being “Very Conservative” in the fiscal preferences, twenty as “Somewhat Conservative”, nineteen as “Neutral”, six as “Somewhat Liberal”, and four as being “Very Liberal”. Fiscal preferences in the sample are generally skewed towards “liberal” preferences. Overall, a majority of the sample considers themselves “conservative” in their preferences, preferring limited government intervention in economic matters.

Social preferences refer to participant beliefs regarding the scope of government intervention in personal and social decision making. Rows 10 through 14 document the frequencies and sample percentages of social preferences ranging from “Very Conservative” to “Very Liberal”. Ten respondents identified themselves as having “Very Conservative” social preferences, fourteen as being “Somewhat Conservative”, nine as being “Neutral”, ten as being “Somewhat Liberal”, and nineteen as being “Very Liberal”. Compared to fiscal preferences, social preferences tend to be more polarized around extremes, with relative equal numbers of “conservative” and “liberal” respondents.

Preference Expression using 2SMV

Two-stage multivoting adds two important sources of information regarding voter choices not discernable in the 1P1V system: information about secondary choices outside the voter’s top choice and their directional perceptions of candidates. Using the methodology described in Chapter 2, different types of voting behavior can be identified from respondents’ candidate-weight vectors.

All else equal, we would expect to observe more instances of strictly dominating preferences in representative elections than in direct elections. The global choice in the context of primary elections is how to distribute votes across candidates rather than policies. For many voters, candidates, especially within the same party, may be perceived as close substitutes. Candidates able to distinguish themselves from others in the field will likely have larger average candidate weight values, primarily driven by respondents with more dominating preferences. This is important within a crowded field of candidates as less distinct candidates are more likely to be treated as irrelevant alternatives. Additionally, voting in favor or against a candidate in a primary election impacts the choice space in the general election. The best strategy of voters is then to select choices in the primaries that increase the likelihood of their preferred candidate winning the general election. As a result, the behavior type frequencies observed in the previous section may not hold when respondents face choices in the mock primaries.

This section identifies voter behavior by calculating the policy-weight coefficient of variation discussed previously for each respondent. Using the respondent's choice of weights for each of the candidates, a candidate-weight coefficient of variation is constructed and categorized based on its value. Coefficient estimates range from 0 to 4.47. A coefficient of 0 corresponds with a respondent being indifferent between candidates, giving the same number of votes to everyone on the ballot. In contrast, a coefficient estimate of 4.47 indicates that the respondent has strictly dominating preferences, using their entire endowment of votes on a single candidate. Values between these two extremes show fixed-weight preferences, with larger values indicating an uneven spread of votes over the ballot.

Figure 1 shows a histogram of the distribution of candidate-weight coefficient of variation estimates of the respondents. Estimates range from 0 to 4.47, and the length of each bin is 0.10. Most of the observed coefficient estimates fall to the right of the midpoint. With the exception of one observation between 0 and 0.1 (indifferent preferences), no observations were found left of 1.3. Compared to the first experiment, where intensities of preference tended to be evenly distributed around the mean, estimates from the mock primaries tend to be skewed left. The most frequent range of coefficient of variation scores are between 4.4 and 4.5, indicative of strictly dominating preferences. These findings suggest that in general respondents tend to prioritize their votes, producing an uneven distribution of votes across candidates, resulting in higher coefficient estimates.

Figure 1

Sample Coefficient of Variation Histogram Plot

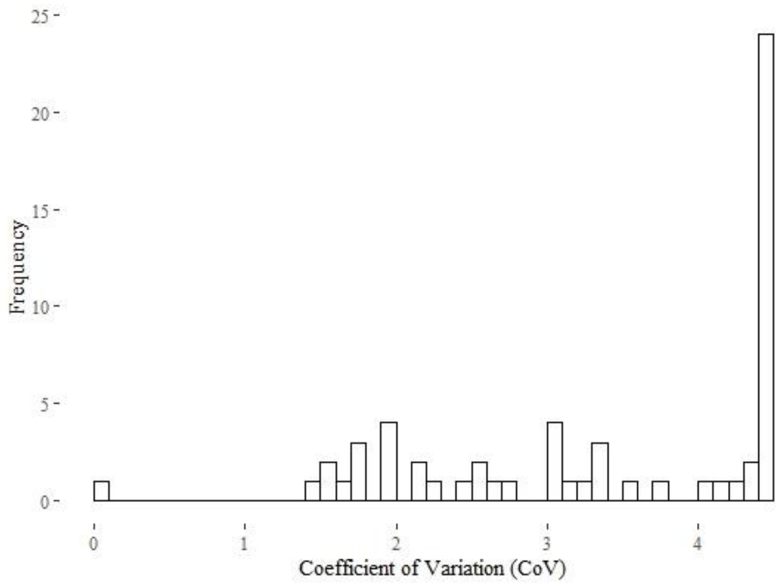

Table 4 shows the frequencies of each type of voting behavior and the associated percent of the sample. Three categories of voting behavior are defined: indifferent preferences (CoV = 0), fixed-weight preferences (0 < CoV < 4.47), and strictly dominating preferences (CoV = 4.47). Two subcategories of fixed-weight preferences are defined: low-variance fixed-weight preferences (0 < CoV ≤ 1), and high-variance fixed-weight preferences (1 < CoV< 4.47).

Table 5

Frequency and Percentage of Voting Behavior Classifications

Voting Behavior	Frequency
Indifferent Preferences (CoV = 0)	1 (1.6%)
Fixed-Weight Preferences (0 < CoV < 4.47)	35 (56.5%)
Low-Variance (0 < CoV ≤ 1)	*0 (0%)*
High-Variance (1 < CoV <4.47)	*35 (56.5%)*
Strictly Dominating Preferences (CoV = 4.47)	26 (41.9%)

Note. First number shows the count of each behavior type. Parentheses show the percent of the sample displaying the behavior.

In representative primary elections with a crowded field of candidates, it would be expected that respondents would prioritize their votes toward certain candidates over an even spread of votes. As shown in Table 4, the most common type of behavior exhibited by respondents was high-variance, fixed-weight behavior. The second most common type of behavior was for respondents to have strictly dominating preferences. These two types of behavior accounted for about 98.4% of the survey. These findings suggest that most respondents tend to prioritize most of their votes towards a relatively few number of candidates.

In the single vote system, only the first choice of the respondent is observed, excluding information about subsequent choices. If respondents utilize the 2SMV mechanism, information regarding inferior options is generated. Excluding the negative/positive component of the 2SMV format, if all voters exhibit strictly dominating preferences, we would obtain the same election results as in 1P1V. If all voters used their entire endowment on a single candidate, this would produce no additional information about preferences. When voting over policies, widespread indifferent preferences returned the 1P1V outcomes. In terms of information generation, the 2SMV format only provides more information about preferences if we observe behavior other than strictly dominating preferences. About 42% of respondents exhibit strictly dominating preferences in the sample, indicating the selection of a single candidate. Nonetheless, more than half of the respondents showed some type of fixed-weight preferences in candidate selection. The prevalence of high-variance candidate-weight coefficients suggested most respondents focused their votes on a select few candidates, showing that respondents are willing to sacrifice part of their endowment to invest in secondary candidate choices. If given the opportunity to use the 2SMV mechanism, most respondents do not exclusively vote for their 1P1V pick, resulting in the system producing more information regarding the rank of individual respondent preferences.

Political Preference Subgroup Analysis of Voting Behaviors

As stated in the description of the experiment, respondents are likely to exhibit different types of behavior based on their political preferences. Respondents intending or most likely to vote in favor of a candidate in both the Democratic primaries and general elections would have different strategies than those intending to vote for another party candidate in the general elections. Additionally, Democratic candidates may perceive the candidates in the field as more diverse in political positions than respondents of other parties. This section provides a subgroup analysis of differences in respondent candidate-weight CoV estimates and voting behaviors.

Table 5 shows summary statistics for the sample subgroups based on political preferences. Column 1 shows the average respondent candidate-weight CoV estimate for the given subgroup. Column 2 shows the subgroup median candidate-weight CoV score, used in identifying subgroup behavior. Column 3 shows the voting behavior classification of the median respondent within that subgroup. Since the definitions of indifferent preferences and strictly dominating preferences are based on exact values of the CoV estimates, the median is used to identify the central tendencies in voting behaviors within the subsample.

Table 6

Summary Statistics of Candidate-Weight CoV Estimates by Political Preference Subgroup

Category	Average Candidate-Weight CoV	Median Candidate-Weight CoV	Behavior Classification
All	3.42	3.89	High-Variance Fixed Weight
	Political Party Affiliation		
Democrats	3.26	3.28	High-Variance Fixed Weight
Republicans	3.92	4.47	Strictly Dominating Preferences
Libertarians	3.14	3.14	High-Variance Fixed Weight
Other	2.93	2.53	High-Variance Fixed Weight
	Fiscal Preferences		
Highly Conservative	3.60	4.47	Strictly Dominating Preferences
Somewhat Conservative	3.62	4.47	Strictly Dominating Preferences

Table 5

Category	Average Candidate-Weight CoV	Median Candidate-Weight CoV	Behavior Classification
Neutral	3.35	3.38	High-Variance Fixed Weight
Somewhat Liberal	2.69	2.55	High-Variance Fixed Weight
Highly Liberal	3.18	3.38	High-Variance Fixed Weight
	Social Preferences		
Highly Conservative	3.39	4.47	Strictly Dominating Preferences
Somewhat Conservative	3.98	4.47	Strictly Dominating Preferences
Neutral	3.53	4.25	High-Variance Fixed Weight
Somewhat Liberal	3.22	3.08	High-Variance Fixed Weight
Highly Liberal	3.09	3.17	High-Variance Fixed Weight

Row number 1 shows the average, median, and median voting behavior classification for the full sample. Using the full sample of observations, the average candidate-weight CoV estimate was 3.42 and the median estimate 3.89. This median estimate corresponds with high-variance, fixed weight behavior. Three categories of political preferences are considered in this study: political party affiliation, fiscal preferences, and social preferences. Across the subsamples, two types of behavior identified by subgroup medians: strictly dominating preferences and high-variance, fixed weight preferences.

Rows 2 through 5 show summary statistics for four political party classifications: Democrat, Republican, Libertarian, and Other. Republican respondents had higher average and median candidate-weight CoV estimates than others. The median Republican respondent exhibited strictly dominating preferences, indicating that they used their full endowment of votes on a single candidate. Republicans were the only group whose median voter exhibited strictly dominating preferences, with all other groups exhibiting median high-variance, fixed weight preferences. Of the other three groups, the average and median CoV estimates were lowest for the "Other" party group, with an average CoV estimate of 2.93 and median of 2.53.

Rows 6 through 10 show the summary statistics for each group within the fiscal preferences category. Respondents identifying as either "very conservative" or "somewhat conservative" produced larger CoV estimates than other groups, with the median respondent in both conservative groups exhibiting strictly dominating preferences. Neutral and liberal respondents reported median CoV estimates within the high-variance, fixed weight preferences category. The observation holds for conservative respondents within the social preferences category as well. Rows 11 through 15 show the summary statistics for groups according to their social preferences. As in the fiscal preferences category, the median CoV estimate for both conservative groups was associated with strictly dominating preferences. Socially liberal respondents had the lowest CoV estimates by both subsample average and median, both classified as having high-variance fixed weight preferences.

The summary statistics in Table 5 suggest that there may be differences in preference intensities between fiscally and socially conservatives and Republican respondents and other associated political preference groups. To determine if there are statistically significant differences within candidate-weight CoV estimates, hypothesis tests for independent samples are conducted.

Assuming that CoV estimates are normally distributed, an unpaired t-test for two samples could be used to assess for differences in average CoV values. However, in the context of this study the distribution of CoV values is unlikely to be normally distributed. As shown by Figure 1, candidate-weight CoV scores tend to skew left, with most of the respondent's estimates occurring to the right of the theoretical median value. CoV estimates are unlikely to be distributed normally. Two tests are used to determine if the CoV estimates for each subgroup samples are normally distributed: the Shapiro-Wilks test and the Anderson-Darling test. Appendix Table B summarizes normality tests of dependent variables used in the study. The Shapiro-Wilks test and Anderson-Darling test show that CoV estimates are not normally distributed for the full sample at the 1% confidence level. Within the party affiliation category, both normality tests conclude that CoV estimates are not normality distributed for the Republican and other party respondents at the 1% confidence level. For the Democratic subsample, the Shapiro-Wilks test suggested non-normality in CoV estimates at the 10% confidence level; however, the Anderson-Darling test suggested normality.

Pett (1997) outlines that nonparametric (or "distribution-free") tests may be more powerful than their parametric counterparts when study samples are drawn from non-normal or unusual distributions, or in cases where group samples are smaller than required by most parametric tests ($n<30$). In general, nonparametrics tests make fewer assumptions regarding a population's distribution, accommodate different types of data, and require less stringent sample size requirements than parametric tests. When assumptions of parametric tests are violated, nonparametric tests may be more powerful as their validity is not affected by whether the variable in question is drawn from a normal population. Others (Rasmussen, 1986; Vickers, 2005) find that under certain conditions, parametric tests are just as powerful as nonparametric tests, even for non-normal populations and small sample sizes.

While the study full sample size is sufficiently large for parametric tests (n=62), several of the subsamples of interest are insufficiently large for parametric subgroup comparisons. Given the results of the normality tests in Appendix Table B, the full and subsample results cannot be assumed to come from normal populations. In order to account for these violations, the study employs both parametric and nonparametric procedures to test for significant differences in candidate-weight CoV estimates. To assess differences between political preference groups, the Kruskal-Wallis One-Way Ranks test, the nonparametric equivalent of ANOVA, is first employed to test if there are differences between any of the groups within a category. In order to identify which, if any, groups are significantly different, both a parametric independent t-test and the nonparametric Dunn test and Mann Whitney U-Test are performed on the relevant groups.

Table 6 shows the results from tests for significant differences in Candidate-Weight CoV estimates between political preference subgroups. The first row of each political preference category shows the Kruskal-Wallis Chi-Squared, a test statistic used to determine if there are any significant differences in that category's subgroup. If there are significant differences between groups, the unpaired two-sample t-test, Dunn test, and Mann Whitney U-Test are performed.

Table 7

Results from Tests for Significant Differences in Candidate-Weight CoV Estimates between Subgroups

Group Comparisons	Unpaired T-Test for Two Samples	Dunn Test	Mann Whitney U-Test
	Political Party Affiliation		
Kruskal-Wallis χ^2= 6.735, df = 2, p-value = 0.03			
Republican / Democrat	2.229** (0.0314)	2.117* (0.0213)	340** (0.0137)
Republican / Other	2.483** (0.0201)	2.259** (0.0358)	256* (0.0580)
Democrat / Other	0.8065 (0.427)	0.262 (0.5970)	182 (0.4874)
	Fiscal Preferences		
Kruskal-Wallis χ^2 = 4.075, df = 4, p-value = 0.43			
	Social Preferences		
Kruskal-Wallis χ^2= 6.0563, df = 4, p-value = 0.19			

Significance Codes: "***" 0.01 "**" 0.05 "*" 0.1

Rows 1 through 4 show hypothesis test results for the political party affiliation subgroup analysis. The first row shows results for the Kruskal-Wallis One-Way Ranks test, a nonparametric hypothesis test that tests for significant differences between the different political parties. The results of the Kruskal-Wallis test indicate that there were significant differences between the CoV estimates of the three political parties (χ^2 = 6.735, p = 0.03). Post hoc analysis using the parametric t-test identifies significant differences in the means of Republican and Democrat estimates (p = 0.03) and between

Republican and Other parties means ($p = 0.02$). Results from the Dunn test show differences in medians between Republican and Democrat estimates ($p = 0.02$) and Republican and Other parties ($p = 0.03$). The Mann Whitney U-Test shows differences in medians between Republican and Democrat estimates ($p = 0.01$) and Republican and Other parties ($p = 0.05$). No significant differences in the CoV estimates of Democratic respondents and other party respondents were detected by any of the three tests.

Rows 5 and 6 shows the results of the Kruskal-Wallis test for differences between fiscal and social preference CoV estimates. Results from the Kruskal-Wallis identify no significant differences between groups for neither fiscal preferences ($\chi^2 = 4.075$, $p = 0.43$) nor social preferences ($\chi^2 = 6.056$, $p = 0.19$).

Based on these summary statistics and tests, the preference intensities of Republican voters measured by candidate-weight coefficient of variation estimates are significantly more dominant than those of Democrats and other political parties. It is not clear if non-Democrat respondents are more likely to prioritize their votes towards individual candidates without more information regarding the preferences of candidates of other parties. While Republican responses are more likely to exhibit dominating preferences based on the results of the independent samples t-test and the Dunn test, there were no significant differences identified between Democrats and other parties.

Voting Outcomes in 1P1V versus 2SMV

The primary purpose of this subsection is to explore the use of the 2SMV format in the context of candidate selection, and how election results potentially differ between the 1P1V and the 2SMV system. In the 1P1V system, candidates win elections by securing support from the greatest number of voters. However, the 2SMV adds a greater quality of information by incorporating the direction of voter preferences into account. Since 2SMV candidate scores use the net value of votes in candidate selection, successful candidates must obtain a sufficiently high number of positive votes without drawing in large amounts of negative votes. Thus, while the winning candidate in a 1P1V election must secure support from the greatest number of people, in a 2SMV election the winning candidate must secure the greatest net votes.

The incorporation of candidate scoring can change election results through the ranking of candidates. Since candidate scoring penalizes candidates that receive large numbers of negative votes, well received, but less recognized candidates potentially can outperform stronger candidates. Candidate rankings, and thus election results, can also differ since 2SMV does not limit respondents to only select one candidate. For an individual voter, multiple candidates can receive votes of varying magnitudes. Assuming the voting base exhibits some degree of fixed-weight preferences, using a free distribution of votes reduces the likelihood of ties in the candidate ranking. Increasing the size of this endowment beyond the candidate space (in this experiment, twice the number of candidates), further reduces likelihood of ties. Since there are more votes than voters, and both positive and negative votes are incorporated into candidate scores, the range of values that candidate scores can take on is much greater than in the 1P1V format. Using information from the survey, election results and candidate rankings are constructed for both the 1P1V and 2SMV versions of the Democratic primaries.

Table 7 shows the mean, standard deviation, and median values of 1P1V and 2SMV candidate scores under the 1P1V system, as well as other relevant variables. Column 1 shows the average and median votes candidates received in the 1P1V system.

Column 2 shows the average and median 2SMV candidate score. Columns 4 and 5 shows summary statistics for candidate “for” and “against” votes used to compute 2SMV candidate scores. Column 6 shows the total votes a candidate received in the 2SMV system. Column 7 shows the ratio of positive to total votes, a measure of the popular perception of candidates.

Table 8

Summary Statistics of 1P1V Votes and 2SMV Candidate Scores

Candidate	Votes (1P1V)	CS (2SMV)	“For” Votes	“Against” Votes	“Total” Votes	Ratio of Positive / Total Votes
Mean	3	76.5	100.2	23.8	124	0.70
Std. Deviation	4	137.2	151.9	30	170.7	0.29
Median	1	20	30	6	69	0.82

Candidates in the 1P1V system received on average 3 votes with a standard deviation of 4 votes. The median value of votes candidates received was 1 vote. In the 2SMV system candidates received an average of 124 total votes with a standard deviation of 170.7 votes. When incorporating both positive and negative votes, the average 2SMV candidate score was 76.5 net votes and a standard deviation of 137.2 votes. The median 2SMV candidate score was 20 net votes. In terms of positive perception, on average about 70% of the total vote received by candidates consisted of positive votes, while the median percentage was 82% of the total vote.

Table 8 shows the candidate scores and election results under 1P1V and 2SMV mechanisms. Column 1 shows the number of votes and rank each candidate received under the 1P1V system. Column 2 shows the 2SMVcandidate score (CS), or the difference in positive and negative votes, and associated rank computed from the 2SMV results. Column 3 shows the change in rank for each candidate between the two systems. A positive value indicates that the candidate moved up in the rankings going from the 1P1V system to the 2SMV system. Columns 4, 5 and 6 respectively show the number of "for", "against", and "total" votes each candidate received. Column 7 shows the ratio of positive to total votes.

Table 9

Candidate Scores and Election Results under 1P1V and 2SMV Systems

Candidate	Votes (1P1V)	CS (2SMV)	Rank Change from 1P1V	"For" Votes	"Against" Votes	"Total" Votes	Ratio of Positive / Total Votes
Joe Biden	14	548	0.5	656	108	764	0.86
Cory Booker	3	118	0	124	6	130	0.95
Pete Buttigieg	1	19	0.5	23	4	27	0.85
Julian Castro	1	61	5.5	65	4	69	0.94
John Delaney	2	87	1	95	8	103	0.92
Tulsi Gabbard	2	1	-8	54	53	107	0.50
Kirsten Gillibrand	1	20	2	24	4	28	0.86
Mike Gravel	1	-87	-7.5	2	89	91	0.02
Kamala Harris	4	52	-4	62	10	72	0.86
John Hickenlooper	1	-2	-5	2	4	6	0.33
Jay Inslee	0	-2	1	2	4	6	0.33
Amy Klobuchar	0	7	4.5	13	6	19	0.68
Wayne Messam	1	14	-0.5	18	4	22	0.82
Seth Moulton	1	20	2	26	6	32	0.81

Table 8

Candidate	Votes (1P1V)	CS (2SMV)	Rank Change from 1P1V	"For" Votes	"Against" Votes	"Total" Votes	Ratio of Positive / Total Votes
Beto O'Rourke	2	164	5	178	14	192	0.93
Tim Ryan	1	-44	-6.5	2	46	48	0.04
Bernie Sanders	14	342	-0.5	359	17	376	0.95
Eric Swalwell	0	0	2.5	6	6	12	0.50
Elizabeth Warren	6	123	-1	181	58	239	0.76
Marianne Williamson	0	26	9.5	30	4	34	0.88
Andrew Yang	7	139	-1	183	44	227	0.81

In both systems, the winner of the Democratic nomination is the candidate that receives the most number of votes. In the 1P1V version of the primaries, this is the candidate most frequently chosen by respondents; whereas, in the 2SMV system this is the candidate with the largest candidate score. The survey data produces different predictions for the nominee using the two systems. Under the 1P1V system, there is a tie for the nomination between Joe Biden and Bernie Sanders, who both receive 14 votes. In the 2SMV system, Biden receives the largest 2SMV candidate score of 542 net votes, followed by Bernie Sanders with 320 votes. The two systems pick the same top two candidates; however, their ranks differ between the two systems.

Under the 1P1V version, the least supported candidate is the one who receives the least number of votes, with a minimum of 0 votes. Four voters received no votes from the respondents in the 1P1V primary: Jay Inslee, Amy Klobuchar, Eric Sawlwell, and Marianne Williamson. Since 2SMV candidate scores incorporate both negative and positive votes, the lowest ranked candidate is the one with the minimum candidate score. Some of these candidate scores may be negative if a candidate is more disliked than liked. Four candidates received negative candidate scores: Mike Gravel, John Hickenlooper, Jay Inslee, and Time Ryan. None of the four candidates who received no votes in 1P1V were among those who received negative votes in 2SMV. Under the 2SMV format, Mike Gravel received the lowest candidate score (-87).

Subsequent rankings of the candidates diverged between the systems. Appendix Table D shows a summary of ranked values of 1P1V and 2SMV candidate scores and other variables. Column 3 in Table 8 shows the change in each candidate’s rank from the 1P1V system to the 2SMV system. Eleven of the candidates experienced an improvement in their rank using 2SMV, and nine fell in the rankings. In general, the 2SMV format tended to improve the rankings of candidates who received small proportions of the total vote, but were otherwise “uncontroversial”, such as Marianne Williamson and Julian Castro. The 2SMV system penalizes candidates with a sufficiently high number of negative votes. These candidates may have secured votes in the 1P1V system, but were largely disliked in the 2SMV system, such as Mike Gravel and Tim Ryan.

When examining the ranked ordering of candidates between the two systems, the 2SMV system produces more unique ranks than in the 1P1V system. In general, ties were most prevalent among the low ranked candidates, as many of the candidates received few or no votes. In the 1P1V system, seventeen of the twenty-one candidates were tied with another candidate at four different vote counts. Four candidates received no votes, eight received one vote, and three received two votes. The highest vote count where ties persist is between Joe Biden and Bernie Sanders, who received ranks of 1.5 in the ordering. Alternatively, in the 2SMV format four of the twenty-one candidates faced ties at two different vote counts (-2 and 20 votes). By expanding the endowment size and allowing for the free distribution of votes, the 2SMV system provides greater variance in the ordered ranking than 1P1V.

Aside from differences in the ranking of each candidate, there are substantial differences in the distance between how many votes each candidate has relative to the other candidates between the two systems. In elections that only require a simple majority to be met to secure nomination, the difference in votes between the top candidates may be very narrow. If there is a narrow difference between two candidates, it is unclear whether this represents the dominance of the top candidate or if the rank could change with the inclusion of more voters. Ideally, there would be a substantial difference between in votes secured by each candidate so it is clear which candidate secures the majority.

In general, the difference in votes between ranks were very small in the 1P1V system since each respondent only has one vote to use. As stated previously, the average candidate received 3 votes with a standard deviation of 4 votes. The average candidate score was 76.5 with a standard deviation of 137.2 net votes, resulting in a coefficient of variation of 1.79. In terms of the difference of votes between each ranking (for example, the votes received by the second and third ranked candidates), on average each rank was separated by 2 votes in the 1P1V system and 35 votes in the 2SMV system. In addition to the greater number of unique ranks, the 2SMV system also produces ranks with a greater degree of separation between each rank.

This finding is most obvious in the difference between the top candidates in both systems. In the 1P1V system, there is a tie for the nomination between Bernie Sanders and Joe Biden (14 votes). Within the 1P1V system it is unclear which candidate secures the nomination. With no or small differences in votes between candidates, the inclusion of additional respondents or additional votes could easily change the rank orderings of candidates. This is demonstrated in the election results in the 2SMV system. Biden receives more votes than Sanders with the difference between their candidate scores being much greater than in the 1P1V system. Biden received a candidate score of 548 while Sanders received a score of 342, resulting in a difference of 206 net votes. In other terms, the difference between the two scores would be the equivalent of about 5 more voters using their full endowment of votes to select Biden. With the expansion of votes, Biden unambiguously wins the nomination.

Political Preference Subgroup Analysis of Candidate Choices

The preceding discussion showed the two voting systems predicted different presidential nominees. While the 1P1V system asks voters who is most preferred, the 2SMV system gauges how much a candidate is preferred based on intensities of preferences. This section provides a subgroup analysis of differences in respondent choices based on political preferences.

Table 9 shows a summary of the 1P1V and 2SMV nominee choices using only responses from each respective subgroup. In parentheses next to each name is the percent of total votes the candidate received. In the case of 2SMV nominee choices, this is the candidate score divided by the size of the total endowment relative to the number of observations within the respective group.

Table 10

Subgroup Summary of 1P1V and 2SMV Nominee Choices

Category	1P1V Choice	2SMV Choice
All	Bernie Sanders/Joe Biden (23.3%)	Joe Biden (21.5%)
	Political Party Affiliation	
Democrats	Bernie Sanders (45%)	Bernie Sanders (19.9%)
Republicans	Joe Biden (37.5%)	Joe Biden (28.2%)
Libertarians	Beto O'Rourke/Andrew Yang (50%)	Andrew Yang (39.3%)
Other	Bernie Sanders/Andrew Yang/Kamala Harris (18.75%)	Bernie Sanders (12.8%)
	Fiscal Preferences	
Very Conservative	Joe Biden/Andrew Yang (23.1%)	John Delaney (15%)
Somewhat Conservative	Bernie Sanders (27.3%)	Joe Biden (28%)
Neutral	Joe Biden (26.3%)	Joe Biden (27.9%)
Somewhat Liberal	Bernie Sanders/Elizabeth Warren (33.3%)	Bernie Sanders (28%)
Very Liberal	Bernie Sanders (66.7%)	Bernie Sanders (56.3%)
	Social Preferences	
Very Conservative	Joe Biden (30%)	Joe Biden (33.3%)
Somewhat Conservative	Joe Biden (21.4%)	Joe Biden (21%)
Neutral	Joe Biden (44.4%)	Joe Biden (35.7%)
Somewhat Liberal	Andrew Yang/Bernie Sanders (30%)	Joe Biden (37.6%)
Very Liberal	Bernie Sanders (38.8%)	Bernie Sanders (26.3%)

Notes. Column 1 shows the 1P1V choice and share of the total vote each candidate received. Column 2 shows the 2SMV choice of the subgroup and the share of the total vote received.

Row numbers 2 through 5 show the nominee picks of the policy party affiliation subgroups. Democratic respondents unambiguously select Bernie Sanders as the nominee in both voting systems; however, the percent of the total vote Sanders receives fall from 45% in 1P1V to 19.9% in the 2SMV system, a difference of 25.1 percentage points. Republican respondents unambiguously select Joe Biden as the nominee in both systems. While there is a decline in total vote share between the 1P1V and 2SMV formats, the decline for Biden is only 9.3 percentage points. As Republican voters on average, exhibited more dominating preferences than Democratic voters, suggesting that while Democratic voters spread their votes across multiple candidates, Republican voters were more likely to spend most of their endowment on a single candidate. Respondents who identified with other political parties selected different candidates between the two formats. In the 1P1V election there was a three-way tie between Bernie Sanders, Andrew Yang, and Kamala Harris for the nominee. In the 2SMV format, Bernie Sanders received the highest candidate score from those within the group.

Rows 6 through 10 show the nominee picks of each fiscal preference subgroup. Very fiscally conservative voters tied in the 1P1V decisions, selecting both Joe Biden and Andrew Yang as the top candidate within the group, but selected John Delaney as the nominee in the 2SMV system. Somewhat conservative voters selected Bernie Sanders as the nominee in the 1P1V system, but gave Joe Biden the highest candidate score in the 2SMV system. Fiscally neutral respondents unambiguously chose Joe Biden as the nominee. Fiscally liberal respondents generally chose Bernie Sanders as the nominee across systems. Somewhat fiscally liberal voters chose Bernie Sanders as the nominee in the 2SMV system, but tied in their nominee in the 1P1V system, selecting both Bernie

Sanders and Elizabeth Warren. Very fiscally liberal respondents unambiguously selected Bernie Sanders as the nominee for both voting systems.

The nomination choices of the social preferences group are shown by rows 11 through 15. In the 1P1V system, conservative and neutral respondents unambiguously selected Joe Biden as the nominee across systems. Nominee selections were more nuanced for liberal voters. Somewhat liberal respondents selected Andrew Yang and Bernie Sanders as their pick for the 1P1V nominee but Joe Biden received the highest 2SMV candidate score in the other system. Bernie Sanders was unambiguously selected as both the 1P1V and 2SMV nominee for very liberal respondents.

Several observations can be made regarding subgroup choices. Aggregate choices of the established parties (Republican and Democrat) remained the same regardless of which system was used. As would be expected, respondents of other political parties were divided in their nomination decisions between the two systems. In terms of fiscal and social preferences, in general conservative respondents tended to have more variation in their nominee selections between systems in the fiscal preferences category, while liberal respondents tended to have more diversity in their social preferences. This would suggest that conservative voters emphasize differences in fiscal positions in candidates more when making their selections than liberal respondents. For example, fiscally conservative respondents may identify a greater difference between the fiscal policies of Joe Biden and Bernie Sanders than would a liberal respondent. In contrast, socially liberal respondents are more varied in their selections of nominees than socially conservative respondents. Both of these observations, as well as the party-affiliation subgroup selections, suggest that respondents value their selections differently based on their political preferences.

Voting Behavior and Election Outcomes

In earlier sections statistical differences in preference intensities were found between Republican respondents and other parties affiliations. Overall, Republican respondents exhibited stronger preferences intensities, and were more likely to distribute large percentages of their endowments toward a few candidates. As a result, the difference between election outcomes may be driven by the preference intensities found within political preferences subgroups. As shown previously, political preference subgroups select different nominees based on their party affiliation and political positions. Given that respondents vary in their preference intensities given their political affiliation, in particular that Republican candidates have more intense preferences than others, it would be expected that election outcomes would be significantly influenced by whichever group population exhibits the most dominating preferences.

Previous results suggest that the average and median CoV estimates are larger for Republican respondents than for non-Republican ones. Thus, Republican candidates will distribute larger vote amounts to fewer respondents than non-Republican respondents. While endowment sizes are the same across individuals, a person who devotes more votes to an individual policy will have more direct influence over that policy's outcome than those with more balanced preferences. If preferences for individual candidates are the same between Republicans and non-Republicans, the difference in preference intensities would not have a significant effect on election results. Given the relationship between political preferences and preference intensities, one question to be answered is what happens to the candidate scores as the sample is divided based on voting behavior? That is, does the exclusion of "high intensity" preferences have a significant effect on the outcomes of elections arising from differences in group candidate rankings?

To determine if there are any differences between group candidate vote distributions, the sample is divided into three comparison groups, and six subsamples based on preference intensity classifications. The first comparison group includes all observations excluding strictly dominating preferences (N = 36), only observations with strictly dominating preferences (N = 26). In this comparison group, the classification of observations into subgroups is based upon the voting behavior rules outlined in Chapter 2 One drawback of defining the subsample groups in this manner is that 2SMV results for the "strictly-dominating" subsample will consist entirely of respondents allocating votes to a single candidate. As a result, some candidates are likely to receive no votes in this subsample.

To address this concern, two other comparison groups are formed that divide the sample by frequency ranges of preference CoV estimates. Instead of relying on the behavioral definitions above, the other comparison groups divide the data by percentiles. In each comparison group, the bottom half of the data represents the candidate preferences of those respondents with lower CoV estimates, while the top half of the data represents "high intensity" preferences. The second comparison group divides the full sample in half by the median candidate-weight CoV, with each subsample using only the bottom and top half of the data respectively (N = 31). The third comparison group divides the data by the 30th percentile (the approximate midpoint of the range of possible

CoV values), with two subsamples using observations from bottom (N = 20) and top (N = 42) halves of the data.

Using only observations from these subsamples, new candidates scores are calculated for each of the candidates within the field. Thus, for each subsample, twenty-one new 2SMV candidate scores are generated to be compared across samples. Since observations vary across subsamples, the calculated 2SMV candidate scores will be incorrectly scaled for comparison. For example, assuming that respondents only vote in favor of candidates the expected 2SMV candidate score for each candidate for the strictly dominating preferences subsample is 48 votes ((24 × 42) / 21) but 120 in the full sample. To adjust for the scaling differences, subsample 2SMV scores are transformed by dividing by the group's total endowment (N × 42) and multiplying the results by the total endowment for the full sample (62 × 42 = 2604). This shows the predicted 2SMV score for the subsample had it contained the sixty-two observations as in the full sample, holding the preferences of the subsample constant.

Table 10 shows descriptive statistics for each subsample. Within each comparison group, the first column represents the subsample categorizing "low intensity" preferences while the second column represents those of "high intensity" preferences. Column 1 shows the mean, standard deviation, median, and number of observations for the full sample of observations. Column 2 and 3 show summary statistics for the "Non-Strictly Dominating" and "Strictly Dominating" subsamples. Columns 4 and 5 show the summary statistics for all observations below and above the 50th percentile respectively. Columns 6 and 7 show statistics for the "Below 30th Percentile" and "Above 30th Percentile" subsamples.

Table 11

Behavior Type Subgroup Descriptive Statistics of Transformed 2SMV Candidate Scores

	Full Sample	Non-SDP	SDP Only	Below 50th	Above 50th	Below 30th	Above 30th
		Group 1		Group 1		Group 3	
Mean	77	167	59	138	89	82	145
Std. Deviation	137	250	247	201	315	107	348
Median	20	60	0	60	0	51	6
Obs.	62	36	26	31	31	20	42

In the full sample, the average 2SMV was 77 votes with a standard deviation of 137 votes. In terms of average transformed 2SMV scores, the “Below 30th Percentile” subsample 2SMV score is the closest in value to the full sample estimate, with 82 votes. By contrast, the non-dominating preference subsample is the farthest from the full sample estimate, with 167 predicted votes. In terms of medians, the “Above 30th Percentile” median was the closest in value to the full sample. Between the subsamples within each comparison group, average and median 2SMV candidate scores estimates tend to be higher within the “low intensity” groups than the “high intensity”. The exception to this occurs between the subsamples in comparison group 3, which contains more observations for the “high intensity” preference category.

Table 11 shows the actual and transformed 2SMV candidate score values calculated from the full sample and behavior type subsamples. Column 1 shows each candidate's 2SMV score using the full sample of data. Column 2 through 7 show the transformed 2SMV scores for subsamples within each comparison group.

Table 12

1P1V Votes and 2SMV Candidate Scores in the 2020 Democratic Primaries

Candidate	Full Sample	Non-SDP	Only SDP	Below 50th	Above 50th	Below 30th	Above 30th
		Group 1		Group 2		Group 3	
Joe Biden	548	626	992	278	1341	184	1435
Cory Booker	118	101	248	86	263	80	269
Pete Buttigieg	19	57	0	51	6	51	6
Julian Castro	61	181	0	181	0	92	89
John Delaney	87	9	248	3	254	3	254
Tulsi Gabbard	1	3	0	3	0	3	0
Kirsten Gillibrand	20	60	0	60	0	60	0
Mike Gravel	-87	-8	-248	-8	-248	-8	-248
Kamala Harris	52	154	0	151	3	107	48
John Hickenlooper	-2	-5	0	-5	0	-5	0
Jay Inslee	-2	-5	0	-5	0	-5	0
Amy Klobuchar	7	21	0	21	0	21	0
Wayne Messam	14	42	0	42	0	42	0
Seth Moulton	20	60	0	60	0	-2	62
Beto O’Rourke	164	237	248	237	248	124	361
Tim Ryan	-44	-5	-124	-5	-124	-5	-124

Table 11

Candidate	Full Sample	Non-SDP	Only SDP	Below 50th	Above 50th	Below 30th	Above 30th
Bernie Sanders	342	886	124	765	246	372	638
Eric Swalwell	0	0	0	0	0	0	0
Elizabeth Warren	123	488	-124	488	-124	296	68
Marianne Williamson	26	77	0	77	0	77	0
Andrew Yang	139	535	-124	414	-2	231	181

In terms of 2SMV election results, no candidate unambiguously wins the nomination across subsamples; however, there are definite commonalities based on preference intensity classification. In each of the "low intensity" preference groups, Bernie Sanders is selected as the 2SMV Democratic nominee; however, in the "high intensity" preference subsamples, Joe Biden is selected. This would suggest that there are significant differences between the distributions of votes received by candidates between the two preference intensity groups. As shown in the analysis of CoV estimates by party type, there are significant differences in the behavior types of Republican participants and other voters. Specifically, in this sample Republican voters were more likely to have strictly-dominating preferences than other parties.

In order to identify significant differences between "low" and "high" preference intensity groups, the study performs parametric and non-parametric hypothesis tests. Subsamples are "paired" in the sense that 2SMV candidate scores are matched by candidate. For example, the Biden's 2SMV score in the "low intensity" sample would be compared to the corresponding Biden 2SMV score in the "high intensity" subsample. As with candidate-weight CoV estimates, corresponding candidate 2SMV score values are unlikely to be normally distributed across subsamples. Appendix Table B confirms this through Shapiro-Wilks and Anderson-Darling test for normality. For all the comparison group subsamples of interest, the assumption of normality is rejected at the 1% confidence level.

Table 12 shows the results of paired hypothesis tests for significant differences between preference intensity subgroups. Column 1 shows the test statistic and p-value for the parametric Paired T-Test. Column 2 shows test results for the nonparametric equivalent of the Paired T-Test, the Wilcoxon-Signed ranks test.

Table 13

Results of Paired Sample Significant Differences Tests in Transformed 2SMV Candidate Score between Behavior Subgroups

Comparison Group	Paired T-Test	Wilcoxon-Signed Ranks Test
Comparison Group 1 (Non-Dominating/Strictly Dominating)	-1.8102* (0.0853)	53* (0.0545)
Comparison Group 2 (Below 50th/Above 50th)	0.6886 (0.4989)	152* (0.0825)
Comparison Group 3 (Below 30th/Above 30th)	-0.9464 (0.3552)	105 (1)

Significance Codes: "***" 0.01 "**" 0.05 "*" 0.1

Column 1 shows the test results of the paired t-test for the three comparison groups. The null hypothesis of the paired t-test is that the mean difference 2SMV scores of the Democratic candidates between the "low preference intensity" and "high preference intensity" subgroups is equal to 0. The alternative hypothesis is that the true difference in 2SMV candidate score means between the two preference intensity groups is not equal to 0. Results for the paired t-test find significant differences in the first comparison group. For comparison group 1, the difference in the means of the “high intensity” preference subgroup ($\mu = 59$) and the “low intensity” subgroup ($\mu = 167$) were statistically significant at the 10% confidence level. No statistically significant differences in mean 2SMV scores were found in the other two comparison groups.

The Wilcoxon-signed ranks test is a nonparametric test that compares the differences between sample medians, that is, do the subsamples come from populations with the same distribution. The null hypothesis of the Wilcoxon-Signed ranks test in the context of this study is that the median 2SMV scores of candidates are the same between the “low preference intensity” and “high preference intensity” subgroups, that is, the median difference in 2SMV candidate scores is equal to 0. The alternative hypothesis is that the median difference in 2SMV scores is not equal to 0.

Results of the Wilcoxon-signed ranks test detect significant differences in two of the three comparison groups. For comparison group 1, the difference in the median 2SMV candidate score in the “low intensity” preference subgroup (Md = 60) and in the “high intensity” preference subgroup (Md = 0) was statistically significant at the 10% confidence level ($p = 0.05$). In comparison group 2, the difference in the 2SMV candidate score medians between the “low intensity” sample (Md = 60) and “high intensity” (Md = 0) was statistically significant at the 10% confidence level ($p = 0.08$). No statistically significant difference was found between the two preference intensity groups for comparison group 3 ($p = 1$).

Overall, hypothesis tests for significant differences show some evidence of differences in 2SMV candidate scores between "high intensity" and "low intensity preferences subgroups. Both the paired t-test and Wilcoxon signed ranks test detect significant differences in the 2SMV candidate scores of the two preference intensity groups. However, overall significant differences tend to diminish for the more "balanced" comparison groups, in terms of vote distributions.

Despite mixed results in tests for differences in central tendencies of 2SMV candidate scores, the dividing of observations based on preference intensity classifications did result in different Democratic nominees. "High intensity" preferences groups unambiguously selected Joe Biden as the nominee across corresponding comparison groups, while "low intensity" preferences group unambiguously selected Bernie Sanders as the Democratic presidential nominee.

Comparison of Alternative Voting Systems

Having discussed the mechanics of two-stage multivoting (2SMV) for both the individual's choice problem and aggregate case, the paper addresses how 2SMV compares to other preference revealing voting mechanisms, namely vote-selling and quadratic voting (QV) systems. For the sake of argument, two forms of quadratic voting are addressed: one that uses real money to purchase votes and one that uses "voice credits". In particular, the discussion examines how different voting systems balance equity and efficiency concerns within their mechanics. Beginning with "one-person, one-vote" (1P1V) as a standard, the paper addresses how each system tries to reveal the preference intensities of voters, if voters are given equal opportunity, and if voters have the equivalent influence in the aggregate election.

In its base construction, 1P1V only reveals the preferences of voters in an election, rather than the intensities of those preferences. As examined in Chapter 2, each individual has both a subjective ranking and a weighted ordering of the policies available. 2SMV approximates the preference intensities by endowing voters with additional votes and asking them to distribute them among elements in the set of policies. The paper attempts to show that provided large enough endowments true intensities of preference, how much more a voter prefers one policy to the others, can be measured when the weightings for all the policies in *P* converge.

Compare this with a vote-selling system as proposed by Buchanan and Lee (1986). Under their framework, voters can express how intense their preferences are for certain policies by purchasing votes from individuals willing to sell their votes. A drawback of this form of vote-selling is that the potential buyer is limited by the number of willing sellers in the market. Quadratic voting overcomes part of this problem by allowing voters to purchase as many votes following a quadratic pricing system. Instead, the voter is limited by their budget constraint of either exogenous currency or "voice credits". Since the price of a vote changes over different quantities of votes, the preference-revealing mechanism found in QV may discourage individuals from voting exclusively on certain policies. This effect would likely be more pronounced under a QV system with an exogenous source of currency, where the individual is not only deciding between different proposals, but also real goods and services.

Moving from the efficiency concerns, the discussion considers how individuals are treated under each voting system. Two types of equity concerns are explored. The first examines the degree to which citizens are afforded equal opportunity within the system. 1P1V satisfies this criteria since each citizen is given exactly one vote per policy. 2SMV, QV with "voice credits", and storable vote systems seem to perform well on this front. Each person is endowed with an initial "budget" of votes or credits and is free to use them in whatever manner they prefer. While the choices they make may result in different degrees of influence over individual policies, everyone is initially provided with the same budget. Provided there are no wasted votes or credits, inter-person equity is preserved under the various model specifications. If voters were not able to fully use the allowance of their budget constraint, perhaps due to differences in choice, some individuals may have purchased more representation than others. Therefore, it is important to select a number of voice credits or votes that ensures the budget constraint is fully used.

The initial equity condition may not hold true under traditional vote-selling and quadratic voting using an exogenous currency, since each voter's budget constraint is based on their personal wealth, rather than the rules of the voting system. In this regard, these two systems may increase the influence of wealthier voters, who have greater resources to expend on vote purchases. Additionally, wealthier voters may have an advantage when it is costly to search and negotiate for additional votes. This is most likely to occur in vote selling models where it is the responsibility of the buyer and seller to broadcast their willingness to participate.

A final equity concern addresses the total representation each individual receives by participating in the election as a whole. Total representation refers to the sum of votes each individual uses in the course of the election. In 1P1V and 2SMV each person is given the same number of votes to use throughout the election, so their total representation is equal to the size of the endowment and the same for all individuals. Under a storable vote system, everyone receives the same number of votes over the span of all elections; however, for a given time period some individuals may be better represented than others due to the choices made in previous periods. Thus, under storable vote systems everyone exerts the same total representation when all elections are considered, but not necessarily in any given election.

Under QV, individuals with strong preferences toward single policies have less total representation than more indifferent voters. Single-policy voters thus face a tradeoff between expressing their true preferences by purchasing more votes or maximizing total representation. To illustrate this, consider the following election using a quadratic voting mechanism. Suppose a voter is given a budget of eight credits to be used in an eight-policy election. If the voter is indifferent between policies, one credit is applied to each of the policies and the individual buys a total of eight votes. If the voter strongly prefers one policy over the others, he can purchase a total of three votes and only vote on his favorite policy. In this case, voters who value policies equally purchase a greater number of votes, and thus total representation, than those with strong preferences. Contrast this with 2SMV where the "price" of votes is constant between different values, thus there is no tradeoff in total representation from purchasing votes for a single policy.

In the case of vote selling, all participants begin under the same framework as 1P1V, where each person is given a single vote for each policy. Voters may either keep or sell their votes and are free to purchase additional votes from other voters. In any given election, some voters may have no representation, effectively opting out of the election. Excluding these voters, total representation among the remaining voters is not necessarily equivalent, with vote buyers exerting greater total influence than those who did not participate in vote buying.

Conclusions

The 2SMV model gives voters an endowment of additional votes exceeding the number of policies under consideration in a direct democracy-style election. Voters may freely allocate this endowment to any of the policies up for election. Using experimental data collected from college students, the study evaluates the performance of the two-stage multivoting system in the context of the 2020 Democratic Presidential Primaries. The results of the experiment can be summarized as follows:

- A majority of respondents (58%) take advantage of the 2SMV system by distributing votes to multiple candidates on the ballot.
- Significant differences were detected in the preference intensities of different political party affiliations. In particular, Republican respondents had stronger preferences intensities than non-Republican voters.
- The choice of nominee varied between voting systems. In general, the 2SMV system produced more unique and distant ranks between candidates, reducing the prevalence of ties found within the 1P1V system. Using the full sample of observations, the 1P1V system resulted in a tie between Joe Biden and Bernie Sanders, while the 2SMV system selected Joe Biden over Bernie Sanders as the nominee by a 206 net vote difference.
- Political preference groups of the same type tended to select the same candidates between voting systems. Notable exceptions include the choices of Democratic

nominees for fiscally conservative and socially liberal respondents, who had mixed results across systems.

- There is mixed evidence suggesting that there are significant differences in the candidate choice vote distributions of “low intensity” and “high intensity” preferences groups. However, the division of observations based on preference intensity classifications did result in different Democratic nominees, with “high intensity” groups selecting Biden as the nominee and “low intensity” groups selecting Sanders as the nominee.

The 2SMV system provides both an alternative to other voting systems, and a methodology of gauging how much more an individual values one policy over another. By changing the size of the initial endowment, 2SMV approximates the intensity of voter preferences and provides a benchmark for identifying different kinds of voter behavior. Since 2SMV requires that each individual be endowed with the same number of votes, equity is preserved in that no individual can obtain more votes due to differences in personal wealth or voting behavior, thus exerting equivalent aggregate influence for all the policies. Through the 2SMV mechanism, the observer not only gains valuable insight into the strength of voters’ preference intensities but generates election outcomes reflective of voters' investment in policies.

References

Arrow, K. J. (1951). *Social Choice and Individual Values*. New Haven: Yale University Press.

Black, D. (1948). On the Rationale of Group Decision-making. *Journal of Political Economy*, 23-34.

Brennan, J. (2012). *The Ethics of Voting*. Princeton: Princeton University Press.

Buchanan, J. M., & Lee, D. R. (1986). Vote buying in a stylized setting. *Public Choice*, 3-15.

Buchanan, J. M., & Tullock, G. (1999). The *Calculus of Consent*. Indianapolis: Liberty Fund, Inc.

Buchanan, J. M. (1954). Individual Choice in Voting and the Market. *Journal of Political Economy*, 334-343.

Caplan, B. D. (2008). *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*. Princeton: Princeton University Press.

Cardenas, J. C., Mantilla, C., & Zarate, R. D. (2014). Purchasing Votes without Cash: Implementing Quadratic Voting Outside the Lab. Working Paper.

Casella, A. (2005). Storable votes. *Games and Economic Behavior*, 391-419.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 135-150.

Dwivedi, A. K., Mallawaarachchi, I., & Alvarado, L. (2017). Analysis of small sample size studies using. *Statistics in Medicine*, 2187–2205.

Hansen, R. L. (2000). Vote Buying. *California Law Review*, 1323-1372.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 41-57.

Howell, D. C. (2010). *Statistical Methods for Psychology*. Belmont: Cengage Wadsworth.

Kochin, M. S., & Kochin, L. A. (1998). When Is Buying Votes Wrong? *Public Choice*, 645-662.

Kraska-MIller, M. (2013). Nonparametric *Statistics for Social and Behavioral Sciences.* Boca Raton: Chapman and Hall/CRC.

Lalley, S., & Weyl, E. G. (2017). Quadratic Voting: How Mechanism Design Can Radicalize Democracy. *American Economic Association Papers and Proceedings*.

Marusteri, M., & Bacarea, V. (2010). Comparing groups for statistical differences: how to choose the right statistical test? *Biochemia Medica*, 15-32.

Pett, M. (1997). Nonparametric *Statistics for Health Care Research.* Thousand Oaks: SAGE Publications.

Posner, E. A., & Weyl, E. G. (2014). Quadratic Voting as Efficient Corporate Governance. *81 University of Chicago Law Review 251*.

Posner, E. A., & Weyl, E. G. (2014). Voting Squared: Quadratic Voting in Democratic Politics. *Coase-Sandor Institute for Law and Economics* Working Paper No. 657.

Rasmussen, J. (1987). Estimating correlation coefficients: Bootstrap and parametric approaches. *Psychological Bulletin*, 136-139.

Schwartz, B. (2004). The *Paradox of Choice: Why Less is More*. New York: Ecco.

Smith, V. L. (1991). *Papers in Experimental Economics*. Cambridge: Cambridge University Press.

Tullock, G. (1959). Problems of Majority Voting. *Journal of Political Economy*, 571-579.

Vickers, A. J. (2005). Parametric versus non-parametric statistics in the analysis of randomized trials with non-normally distributed data. *BMC Medical Research Methodology*, 5-35.

Weyl, E. G. (2013). Quadratic Vote Buying. Working Paper.

Young, P. (1995). Optimal Voting Rules. *Journal of Economic Perspectives*, 51-64.

Appendix A: Survey Experiment Overview and Questions

Appendix Table A

Survey Experiment Questions

Question	Response Choices
Part 1: Political Preferences	
(1) Which political party do you normally support or identify with?	● Democratic Party ● Republican Party ● Libertarian Party ● Green Party ● Other
(2) How fiscally conservative/liberal do you consider yourself? In this question, a fiscal conservative would advocate for low tax rates, reduced government spending, and minimal government interference in economic matters?	● Very Conservative ● Somewhat Conservative ● Neutral ● Somewhat Liberal ● Very Liberal
(3) How socially conservative/liberal do you consider yourself? For example, in the United States a social conservative may not support abortion, same-sex marriage, gun control, or social issues of similar nature.	● Very Conservative ● Somewhat Conservative ● Neutral ● Somewhat Liberal ● Very Liberal
Part 2: 2020 U.S. Democratic Primaries Election Experiment	
Overview: The following section asks you to vote in a mock version of the 2020 U.S. Democratic Party Presidential Primaries under the two voting systems described in Section II. The first question asks which candidate you would vote for if the elections were held today using the one-person one-vote system. The next two questions use the two-stage multivoting system. The second question asks you to vote "for" or "against" each of the candidates. You should only consider the listed candidates when making your decision. In the third question, you will be given 42 votes (or twice the number of candidates) to assign to whichever candidate you prefer. For example, if you strongly dislike a particular candidate, you should vote "against" them in question 2, and assign them most of your votes in question 3. You are free to look up any information you may need on the candidates to help you make your decision.	
(1) If you were required to vote in the 2020 Democratic Primaries today, which candidate would you vote for? Select from the drop down list.	"Biden", "Booker", "Buttigieg", "Castro", "Delaney", "Gabbard", "Gillibrand", "Gravel", "Harris", "Hickenlooper", "Inslee", "Klobuchar", "Messam", "Moulton", "O. Rourke", "Ryan", "Sanders", "Swalwell", "Warren", "Williamson", "Yang"
(2) Please select whether you would vote "for" or "against" each candidate.	● “For” ● “Against”

Appendix Table A

Question	Response Choices
(3) Please enter how many votes you wish to assign each candidate. In this question you will be given 42 votes to use between the policies in the docket. Your responses must total to 42.	"Biden", "Booker", "Buttigieg", "Castro", "Delaney", "Gabbard", "Gillibrand", "Gravel", "Harris", "Hickenlooper", "Inslee", "Klobuchar", "Messam", "Moulton", "O. Rourke", "Ryan", "Sanders", "Swalwell", "Warren", "Williamson", "Yang"

Appendix B: Normality Test for Dependent Variables in the Study

Appendix Table B

Normality Tests for Key Sample Variables

	Shapiro-Wilks Normality Test	**Anderson-Darling Test**
	Political Preferences Subgroup Analysis	
Candidate-Weight Coefficient of Variation (Full Sample)	0.8251*** (0.0001)	4.2449*** (0.0001)
	Political Party Affiliation	
Democrat CoV Estimates	0.9135* (0.0746)	0.5456 (0.1402)
Republican CoV Estimates	0.6223*** (0.0001)	4.0252*** (0.0001)
“Other” CoV Estimates	0.8522** (0.0146)	1.0071*** (0.0086)
	Fiscal Preferences	
Fiscal Preferences -Very Conservative	0.7130*** (0.0007)	1.6801*** (0.0001)
Fiscal Preferences - Somewhat Conservative	0.7276*** (0.0001)	2.4795*** (0.0001)
Fiscal Preferences - Neutral	0.8441*** (0.0053)	0.9629** (0.0118)
	Social Preferences	
Social Preferences -Very Conservative	0.6837*** (0.0005)	1.5116*** (0.0002)
Social Preferences - Somewhat Conservative	0.6581*** (0.0001)	2.0035*** (0.0001)
Social Preferences - Neutral	0.7654*** (0.0081)	0.9441*** (0.0093)
Social Preferences -Somewhat Liberal	0.8906 (0.1417)	0.4407 (0.235)
Social Preferences -Very Liberal	0.9026* (0.0638)	0.5590 (0.1268)

Appendix Table B-2

Normality Tests for Key Sample Variables

	Shapiro-Wilks Normality Test	Anderson-Darling Test
	Candidate Vote Transformed Estimates	
Full Sample	0.7470*** (0.0001)	1.8919*** (0.0001)
Excluding Dominant Preferences	0.7194*** (0.0001)	2.4101*** (0.0001)
Dominating Preferences Only	0.6530*** (0.0001)	2.6504*** (0.0001)
Below 50th Percentile	0.7388*** (0.0001)	1.9972*** (0.0001)
Above 50th Percentile	0.5648*** (0.0001)	3.3476*** (0.0001)
Below 30th Percentile	0.8066*** (0.0008)	1.4317*** (0.0007)
Above 30th Percentile	0.6612*** (0.0001)	2.577*** (0.0001)

Notes: Only results for relevant subsamples with observations greater than 7 included.

Significance Codes: "***" 0.01 "**" 0.05 "*" 0.1

Appendix C: Institutional Review Board Approval

IRB
INSTITUTIONAL REVIEW BOARD
Office of Research Compliance,
010A Sam Ingram Building,
2269 Middle Tennessee Blvd
Murfreesboro, TN 37129

MIDDLE TENNESSEE STATE UNIVERSITY

IRBN007 – EXEMPTION DETERMINATION NOTICE

Thursday, May 16, 2019

Principal Investigator	**Emilia J. Suggs** (Student)
Faculty Advisor	Adam Rennhoff
Co-Investigators	NONE
Investigator Email(s)	*ejs4p@mtmail.mtsu.edu; adam.rennhoff@mtsu.edu*
Department	Economics and Finance
Protocol Title	***Two experiments in two-stage multivoting***
Protocol ID	**19-1244**

Dear Investigator(s),

The above identified research proposal has been reviewed by the MTSU Institutional Review Board (IRB) through the **EXEMPT** review mechanism under 45 CFR 46.101(b)(2) within the research category *(2) Educational Tests* A summary of the IRB action and other particulars in regard to this protocol application is tabulated as shown below:

IRB Action	**EXEMPT from furhter IRB review*****	Date	**5/16/19**
Date of Expiration	NOT APPLICABLE		
Sample Size	100 (ONE HUNDRED)		
Participant Pool	**Healthy adults (18 or older) - MTSU students**		
Exceptions	Online data collection with online informed consent allowed through Qualtrics		
Mandatory Restrictions	1. Participants must be 18 years or older 2. Informed consent must be obtained from the participants 3. Identifying information must not be collected		
Restrictions	**1. All restrictions for exemption apply.** **2. Mandatory informed consent with age-verification prior to the start of Qualtrics survey.**		
Comments	Forced responses in the survey were noted during the review.		

***This exemption determination only allows above defined protocol from further IRB review such as continuing review. However, the following post-approval requirements still apply:

- Addition/removal of subject population should not be implemented without IRB approval
- Change in investigators must be notified and approved
- Modifications to procedures must be clearly articulated in an addendum request and the proposed changes must not be incorporated without an approval
- Be advised that the proposed change must comply within the requirements for exemption

IRBN007 Version 1.3 Revision Date 05.22.2018

- Changes to the research location must be approved – appropriate permission letter(s) from external institutions must accompany the addendum request form
- Changes to funding source must be notified via email (irb_submissions@mtsu.edu)
- The exemption does not expire as long as the protocol is in good standing
- Project completion must be reported via email (irb_submissions@mtsu.edu)
- Research-related injuries to the participants and other events must be reported within 48 hours of such events to compliance@mtsu.edu

Post-approval Protocol Amendments:

The current MTSU IRB policies allow the investigators to make the following types of changes to this protocol without the need to report to the Office of Compliance, as long as the proposed changes do not result in the cancellation of the protocols eligibility for exemption:

- Editorial and minor administrative revisions to the consent form or other study documents
- Increasing/decreasing the participant size

Only THREE procedural amendment requests will be entertained per year. This amendment restriction does not apply to minor changes such as language usage and addition/removal of research personnel.

Date	Amendment(s)	IRB Comments
NONE	NONE.	NONE

The investigator(s) indicated in this notification should read and abide by all applicable post-approval conditions imposed with this approval. Refer to the post-approval guidelines posted in the MTSU IRB's website. Any unanticipated harms to participants or adverse events must be reported to the Office of Compliance at (615) 494-8918 within 48 hours of the incident.

All of the research-related records, which include signed consent forms, current & past investigator information, training certificates, survey instruments and other documents related to the study, must be retained by the PI or the faculty advisor (if the PI is a student) at the sacure location mentioned in the protocol application. The data storage must be maintained for at least three (3) years after study completion. Subsequently, the researcher may destroy the data in a manner that maintains confidentiality and anonymity. IRB reserves the right to modify, change or cancel the terms of this letter without prior notice. Be advised that IRB also reserves the right to inspect or audit your records if needed.

Sincerely,

Institutional Review Board
Middle Tennessee State University

Quick Links:
Click here for a detailed list of the post-approval responsibilities.
More information on exmpt procedures can be found here.

Appendix Table D: 2SMV Candidate Ranks

Appendix Table D

Candidate Ranks in 1P1V and 2SMV Systems and Relevant Subsamples

Candidate	Votes (1P1V)	Candidate Score (2SMV)	Rank Change from 1P1V	"For" Votes	"Against" Votes	"Total" Votes	Positive/Total Votes
Joe Biden	1.5	1	0.5	1	1	1	8
Cory Booker	6	6	0	6	12.5	6	1.5
Pete Buttigieg	13.5	13	0.5	14	18	16	10
Julian Castro	13.5	8	5.5	8	18	11	3
John Delaney	8	7	1	7	10	8	5
Tulsi Gabbard	8	16	-8	10	4	7	16.5
Kirsten Gillibrand	13.5	11.5	2	13	18	15	8
Mike Gravel	13.5	21	-7.5	19.5	2	9	21
Kamala Harris	5	9	-4	9	9	10	8
John Hickenlooper	13.5	18.5	-5	19.5	18	20.5	18.5
Jay Inslee	19.5	18.5	1	19.5	18	20.5	18.5
Amy Klobuchar	19.5	15	4.5	16	12.5	18	15
Wayne Messam	13.5	14	-0.5	15	18	17	11
Seth Moulton	13.5	11.5	2	12	12.5	14	12.5
Beto O'Rourke	8	3	5	5	8	5	4
Tim Ryan	13.5	20	-6.5	19.5	5	12	20
Bernie Sanders	1.5	2	-0.5	2	7	2	1.5
Eric Swalwell	19.5	17	2.5	17	12.5	19	16.5
Elizabeth Warren	4	5	-1	4	3	3	14
Marianne Williamson	19.5	10	9.5	11	18	13	6
Andrew Yang	3	4	-1	3	6	4	12.5

CHAPTER 4: POLITICAL ENTRY IN STATE ASSEMBLY ELECTIONS: EVIDENCE FROM STATE ASSEMBLY ELECTIONS FROM 2016-2020

Political competition is often described as analogous to economic competition in how political candidates and firms make decisions (Stigler (1972)). Firms decide whether it is profitable for them to enter a market, similar to how candidates decide if it is beneficial for them to enter an election. Both act strategically in the presence of their rivals, often by differentiating themselves by appealing to different preferences. However, a fundamental difference exists between political and economic competition. Stigler (1972) argues that while a market can serve consumers with different preferences, representation in elections is inherently constrained due to the discrete nature of group decisions. While individuals perfectly represent themselves through their choices in a market, collective decision-making results in underrepresentation if the members of the group are not unanimous in their positions. An elected official can only adopt one position at a time, leaving some voters potentially underrepresented in a given election.

While the nature of collective decision making is unlikely to result in perfect representation, seemingly unrepresented voters can still affect the degree to which they are underrepresented through the forces of political competition. An important way in which potential candidates appeal to voters is through their political positioning on key issues. Candidates aspiring to win elections must design platforms that appeal to the greatest number of voters. If a candidate is unchallenged by other rivals there is no need to alter their own personal positions. Additional rivals have the potential to reduce a candidate's share of voters. In order to remain competitive, candidates may need to adopt positions different from their own personal beliefs. Downs (1957) argued that in two-party elections candidates tend to adopt the same positions in order to appeal to the "median voter" in elections. However, subsequent empirical work finds that political candidates adopt positions furthest away from one another, becoming polarized in the issues at hand.

This analysis investigates political competition in state assembly elections using two approaches toward modeling political entry. In the first approach, political party entry decisions are influenced by election, demographic, and economic attributes of election districts. Political party-candidates that "fit" the district's preferences are more likely to enter than others.

The second approach models political entry based on the expected payoffs each political party is likely to receive as a result of political party entry. Inspired by entry models found in the empirical industrial organization literature, the expected payoff approach assumes that candidates make their decision to enter elections based on the expected monetary benefits and costs associated with entry. Some of these payoff components are dependent upon success in the general election (for example, representative salary is only received if the candidate holds office) while other components are realized regardless of election outcomes. Thus, probabilistic elements related to election success must be accounted for in the candidate's decision calculus. A candidate enters an election if their expected payoff is greater than the forgone income from outside employment.

Using these two approaches to modeling political entry, the study estimates the marginal effects that certain district attributes and components of expected payoff have on the probability of entry for three political parties in the United States: Democrats, Republicans, and Independents. Predicted probabilities of entry are tested against observed outcomes in order to evaluate the predictive strength of these two approaches.

The paper is organized as follows: the first section reviews the theoretical and empirical literature studying political competition and entrance. The second section introduces the theoretical framework underlying the study's two approaches to political entry. The third section discusses the data and provides descriptive statistics for the dependent and independent variables used in the models. The fourth section shows estimation results for the estimating the probabilities of a political party winning, given district characteristics, as well as descriptive statistics regarding estimated success probabilities. The fifth section shows estimation results of the two models of political entry, and tests the predictive strength of the two models against observed election configurations. The sixth section concludes.

Literature Review

Downs (1957) argues that in a two-party system, rivalry forces candidates to adopt similar positions. This model of political competition is similar to the Hotelling's (1929) model of product differentiation. In the Hotelling-Downs model, the ideal position in a two-party election is to appeal to the median voter in the distribution of voter preferences. For example, the Hotelling-Downs model predicts that in a district that is predominantly conservative, both parties will choose more conservative platforms than they may have otherwise. The stable outcome in this case is for both parties to adopt essentially the same positions. When more than two political parties enter an election, Downs (1957) predicts that the best strategy is to remain as distinct as possible.

The Hotelling-Downs model suggests that political competition in the two-party system results in both candidates adopting similar policies. In this regard, the United States, which has historically been categorized as a two-party system, provides an excellent backdrop to test the predictions of the median voter theorem. Contrary to the predictions of the Hotelling-Downs model, political candidates in the United States seem to appeal to the extremes of their own party rather than to the median voter (Poole and Rosenthal (1984); Fiorina and Abrams (2008); Grosser and Palfrey (2014)). According to Poole and Rosenthal (1984), many congressional districts feature equal proportions of liberal and conservative voters, resulting in near equal chances of winning the district election.

Findings contrary to the Hotelling-Downs predictions suggest that candidates in two-party systems tend to be polarized in their platforms. In an election with political polarization, we would expect candidates to take on more extreme positions than they would have otherwise. The argument for political polarization makes intuitive sense in two ways. First, if the candidates are non-distinct in their own characteristics (political experience, personality, appearance, etc.) and in their platforms, there is no meaningful choice to be made by the voters since the alternatives are essentially the same. Second, political candidates may not be able to credibly convince the median voter of their assumed position if their name is associated with a political party on the other extreme. For example, suppose a liberal candidate enters an election in a predominately conservative election and modifies their platform to appeal to conservative voters. If the voters know that the candidate has held a liberal platform in the past, they may be deemed untrustworthy by the voters. If political competition is characterized by political polarization we would predict, all else equal, that candidates from minority political parties will not enter elections because they are unlikely to win.

In this regard, the nature of political competition affects if and how a potential candidate decides to enter an election. If a candidate cannot appeal to a sufficiently large voting base, whether by appealing to the median voter or most extreme voter, they will not win an election. Theoretical models of political entry describe citizen-candidates as utility-maximizers, facing similar costs and benefits to a firm (Stigler (1972); Osborne & Slivinski (1996)). Candidates run for office if they believe they will receive a positive payoff from holding the position. However, it is not always clear what these costs and benefits are to the candidate. Certain costs and benefits are observable, such as salary packages and campaign expenditures. Other payoffs may be unobservable, such as networking connections made in office, insider information about business performance, and personal gains (or losses) from holding a position of importance. Candidates between different political parties do not necessarily differ in the utility received from political office. If payoffs do not differ between candidates, the only other source of variation in determining whether it is profitable to run for office is the probability that the candidate wins the election. Thus, it is the expected payoff of office that determines whether or not a candidate decides to enter an election.

The likelihood of winning an election is dependent on the number and type of entrant. Lee (2012) estimates the probabilities of party entrance in two-party and multiparty elections using probit estimation; however, his approach does not incorporate strategic effects. Carson (2003) builds upon the work of Signoria (1999), analyzing candidate entry using incorporating strategic interactions into the standard probit models. By neglecting the impact of rivalry and strategic interaction, Carson suggests that the coefficient estimates obtained in the standard probit model are likely to be biased. In addition to pure rivalry, the quality of one's rival is likely to affect the likelihood of success. Incumbent rivals tend to have a strong advantage over other candidates in elections (Gelman and King (1990), Breaux (1990), Cox and Morgenstern (1995)). Potential candidates facing incumbent rivals may not enter on the grounds that they are less likely to win an election. Cox and Morgenstern (1995) find that the incumbency advantage in state house elections has been increasing over time in both single-member and multimember elections. Contribution spending by rivals has also been found to act as a deterrent to entry and reduce party-candidate's probability of an election win (Goldenberg et al. (1986), Hersch and McDougall (1994), Box-Steffensmeier (1996), Epstein and Zemsky (1995)).

As suggested by Downs (1957), candidates segment members of the voting population based on their characteristics. A potential candidate increases the likelihood of winning an election by appealing to the most prevalent economic and demographic characteristics of a district. These could include factors such as a district's population, wealth distribution, poverty rate, ethnic makeup, and occupational makeup. Given the many possible sources of variance, candidates must be selective in how they present themselves in elections.

This paper adds to the existing literature in several ways. First, the study distinguishes between the entry decisions of different political party-candidates (rather than as individuals) using two unique approaches: one using a reduced-form, characteristics based model, and the other using decision rules pertaining to expected payoffs. By estimating entry and success probabilities by political party, the study decomposes the marginal effects that characteristics about a district's election race, demographics, economy, and payoffs have on different political parties' entry decisions. Given that empirical tests of the median voter theorem suggest the presence of political polarization, it is highly likely that there will be inter-party differences in marginal effects. Second, the analysis utilizes recent state assembly election data, an election geography studied less in political competition literature. Using state assembly data (as opposed to congressional data) allows us to observe more individual races over districts with greater heterogeneity, as well as differences between political parties, which is not always possible in Congressional elections.

Theoretical Framework

This paper presents two approaches to modeling political entry decisions: one based on a political party candidate's reaction to changes in the election, demographic, and economic attributes of a district, and the other based on decision rules pertaining to a candidate's expected payoff associated with running for political office.

Political Entry Based on District Characteristics

In this first approach, attributes about a district (for example, racial composition or unemployment) influence political preferences within a district and affect the probability of observing certain types of political parties enter a district election. The study assumes that in the absence of district characteristics, political party-candidates enter at equal rates. For example, all else equal, this assumption would be violated if we observe Democratic candidates entering more frequently than Republican or Independent ones. Under this assumption, attributes about a political party itself do not make an individual candidate predisposed to entry.

Election district geographies within a state are not homogenous. Some districts are more rural, ethnically diverse, have different rates of educational attainment, or specialize in different types of industry trades. For example, some districts have a greater proportion of workers employed in manual labor trades than other industry occupations. Finding from previous political competition literature suggests that political parties divide district voters by these groups and appeal to the extreme ends of each characteristic group (Poole and Rosenthal (1984); Fiorina and Abrams (2008); Grosser and Palfrey (2014)). Designing a political platform that segments a significant group of the voting population is an efficient way to secure votes; as such, political parties develop brands to attract certain types of voters.

Given the heterogeneity of election districts, the amount of voters belonging to a given political party likely varies by district. Some political parties may be more dominant than others because the characteristics of residents make them more likely to have certain political preferences. There are two ways in which this can affect political entry.

First, the political party-candidates that enter elections can be thought of as samples of the districts in which they seek election. For example, if a majority of voters within a district are Republicans and political parties have the same intrinsic rates of entry, it would be expected that a random entrant is more likely to be a Republican than of another political party. The political preferences and party affiliation of potential candidates that come from a given district are influenced by the same characteristics that have shaped the preferences of the majority of voters.

Second, differences in district characteristics are likely to affect the probability that a candidate wins a district election. As district characteristics influence majority preferences and party affiliation, candidates that are similar to those preferences would, all else equal, be more likely to secure votes. Because of this, there may be a strong disincentive to identify as a minority party in a district with a large difference between voters of the majority and minority parties. It may be to the best interest of candidates from rival parties to run under the affiliation of the majority party, even if they personally do not share a similar platform.

In both cases, a candidate's political party affiliation, and ultimately their decision to run as a particular political party, is likely to be dependent on the district characteristics in which an election takes place. Some district characteristics will make it more likely to observe a political party enter than others. Ultimately, this can affect the amount of political competition that is observed in a district. In order to increase political competition, characteristics about the election districts would need to be modified to make entry more likely.

Political Entry Based on Expected Payoffs

The expected payoffs approach follows the intuition of models found in the empirical industrial organization literature to analyze the political entry decisions of candidates of different political parties based on differences in their expected gains from entering an election. In particular, the study's approach is based on the entry models first introduced by Berry (1992) in his analysis of the competition in the airline industry. In his paper, Berry (1992) models the entry decisions of firms as a discrete choice game played over two stages. In this specification, firms choose to "enter" or "stay out" of the market based on their expected profitability. Firms that stay out of the market are assumed to make zero profit. If post-entry profits are greater than those from staying out of the market, the firm chooses to enter.

This analysis considers the entry decisions of candidates based on their political party affiliation: Democrat, Republican, or Independent. The term "party-candidate" is used to specify the political party affiliation of the candidate. Using Berry's (1992) model of competition as a framework, the study models political entry decisions and competition based on the expected monetary payoff to candidates for entering elections. Components of this payoff include expectations about salary, per diem pay, and other benefits, which are only received upon winning an election race. Other components, such as campaign contributions and expenses, are incurred regardless of whether or not the candidate is successful in the general election. Treating candidate entry like firm entry, a candidate enters an election if their expected payoff is positive, or greater than the payoff of non-entry (such as a forgone salary).

It is important to emphasize that the benefit to the candidate of political participation is the *expected* payoff, as the benefits of holding office are only realized if the candidate wins their election. If the candidate enters and is unsuccessful, it would generally be expected that the candidate receives non-positive payoffs; either zero profit if no costs were incurred, or negative if campaign expenses were not offset by external campaign contributions. An exception to this would be in the circumstance where an unsuccessful candidate raises more in contributions than they spend in expenses, in which case they would have a positive payoff without receiving the benefits associated with holding office (salary, per diem, etc.).

Given that there is a risk of not receiving the monetary benefits of office if one is unsuccessful in obtaining office, profit-maximizing candidates would only enter elections from which they expect to perform well enough to receive positive payoffs. Thus, not only should the nominal value of such benefits factor into a candidate decision calculus, but also the probability of obtaining said benefits. The expected payoff to a given candidate of entering an election is then equal to the probability of winning an election times the associated monetary benefits plus benefits obtained regardless of performance minus any costs associated with running for office. Candidates with small chances of winning an election will have smaller expected payoffs than those with high chances of winning. Since expectations are likely to play a significant role in entry decisions, the probabilities of success must be estimated for the three parties, incorporating factors that are likely to impact a given candidate's performance.

In order to model political competition in state assembly elections, two types of equations must be estimated, one estimating the probability that each political party-candidate wins their respective election and subsequently a model of expected payoffs used to predict party-candidate entry. In this quasi-structural model, fitted values obtained from the victory probability model are used to scale salaries, per diem pay, and other benefits to their expected values. For example, representatives in Alaska state assemblies receive an annual salary of $50,400. If a Republican candidate in a certain district election has a 70% probability of success, their expected salary would be $35,200 (0.70 x $50,400).

Equation 1 represents the probability of winning an election for political party-candidate *j* in electoral district *d*. In this model, three types of characteristics are likely to impact the likelihood of a candidate winning a district election: characteristics of the election and political preferences, levels of campaign financing by oneself and rivals, and characteristics about a district's population and economic geography. X_E represents a matrix of covariates characterizing the election, including incumbency status and the past voting history of voters in the district. X_C includes information about one's own level of campaign financings, as well as one's rivals. X_C shows covariates summarizing demographic and economic characteristics of the election district. These include measures of central tendency regarding a district population's age, population size, marital status, educational attainment, and ethnic composition. Economic characteristics include unemployment, district income, health insurance coverage, and industry composition.

$$\Pr\left(Win_{jd}\right) = \beta_j X_E + \beta_j X_C + \beta_j X_D + \varepsilon_{jd}$$

An in-depth discussion of the estimation of Equation 1 is provided further in the study. In summary, the study uses several specifications of probit models to estimate equation 1, adopting a random effects probit panel model as the final specification for success probability estimates.

Following the estimation of the parameters of victory probability model, the study uses the fitted values of this model to estimate the probabilities of success in a given district election for each of the three political party-candidates. Components of a candidate's expected payoff that are only incurred upon winning a district seat are scaled by that party's estimated probability. Additionally, some components of expected payoff are incurred regardless of a candidate's success, such campaign contributions and expenditures.

Equation 2 represents the expected payoff of participation $_{jd}$ for party-candidate *j* in electoral district *d*. X_w represents monetary benefits that are only received if party-candidate *j* is successful in winning a district seat. X_b represents benefits that are received regardless of one's performance, namely campaign contributions.

Rival entry is assumed to have a negative monetary impact on a party-candidate's expected payoff and subsequent entry decisions. In an unopposed race, it is expected that candidates incur a base level of expenditures in order to finance their campaign, perceived in this model as the costs of participation. If a rival political party-candidate enters, it would be expected that candidates would need to increase spending on marketing, strategy, and other campaign expenditures to compete against the new entrant. The necessary additional expenditure required for a candidate to win a district election is likely to increase for candidates with lower probabilities of success. In the model, y_{jd} represents the rank of party-candidate *j* based on their probability of winning district election *d*, with a rank of 1 indicating the candidate most likely to win. The coefficient δj represents the additional expenditures associated with having a lower probability of success.

$$\Pi_{jd} = \Pr\left(Win_{jd}\right) X_w + X_b + \delta_j y_{jd}$$

In the expected payoffs model, a candidate enters if their expected payoff is greater than a baseline level of forgone expenditure. Using this approach, more competition can be induced by raising benefits or lowering the costs associated with participating in an election.

Data

Election Data Sources

There are several reasons why state assembly data is preferable to congressional data. First, extremely high expected payoffs in congressional elections may produce entry regardless of the probability of being successful. Winning a seat in the U.S. Congress is associated with large economic payoffs (both internal and external) such that expected potential benefits overcome even high probabilities of loss. In contrast, the benefits in state and local elections are much lower than in Congressional elections, enough so that potential candidates may not enter. For example, in 2018 the annual base salary for Congressional representatives was $174,000 compared to the average of $40,000 for state representatives as found in our sample (Brudnick (2018)). If the potential benefits are sufficiently large, party-candidates will enter regardless such that we never identify other election configurations.

Second, state and local elections comprise smaller geographies than congressional elections, making them more homogenous in economic and demographic characteristics. For example, if the characteristics of a district make the median voter extremely liberal, it may be difficult for an established Republican to credibly appeal to the median voter. Poole and Rosenthal (1984) observe that congressional election districts have a fairly even split of Democrat and Republican voters. In Congressional elections we may not observe many instances of major parties running unopposed given an even divide of political preferences.

Finally, estimates obtained from using state assembly elections will have more statistical power than congressional elections because we observe significantly more state races. For example, in the United States during the 2018 election cycle, 435 seats in the U.S. House of Representatives were up for election. Combining all state lower chamber elections approximately 4000 state lower chamber seats were contested in 2018, almost ten times the number of Congressional seats. Using state assembly data allows us to observe more elections with greater heterogeneity in entry configurations.

There are two primary goals of this study: first to identify factors that are likely to affect a given party-candidate's ability to win an election, and second to test the performance of a model of political entry based on a candidate's expected payoff. This study uses district-level election data from thirty-six state assemblies from 2016 to 2020. Overall, the dataset includes observations from 8,893 separate races across the United States. Appendix Table A summarizes key election information for the three years.

Election data is obtained from Ballotpedia, an online political encyclopedia that documents election results and procedures for federal, state, and local elections. Several types of information are collected from this source for the study: the configuration of each election, which candidate won the election, and if a candidate was an incumbent. The election configuration defines which political parties enter a given election. Since the paper examines entry behavior of three political parties, there are seven possible election configurations, excluding the possibility that no candidate entered. As a proxy measure of existing political preferences within a district, the study uses vote count percentages for each political party from the 2014 elections, prior to the first year in the survey.

District characteristic data comes from the 2016-2019 American Community Survey’s (ACS2016-2019) five-year estimates. These characteristics highlight the unique demographic and economic makeup of each district, and influence each candidate's probability of winning their district’s election. The demographic variables used in the analysis are population, median age, ethnicity, percent with a high school degree or lower, and percent never married. The economic variables considered in the study include per capita income, unemployment, poverty rate, health insurance coverage, and the industry makeup of the district. Excluding population, income, and age, all other variables are expressed as district proportions.

Information about the benefits and costs associated with candidate political participation comes from the National Conference of State Legislatures (NCSL). This includes the data regarding each state assembly’s salary package, per diem pay, filing fees, and other benefits offered by that state. State assemblies establish their own salary packages, which apply equally across districts. In addition, states differ in whether they define legislative service as full-time or part-time work. The study distinguishes between part-time and full-time work in three ways. First a binary variable is created, indicating a state’s work status that can be interacted with other variables. Second, salary and per diem pay comparisons are calculated based on state guidelines, such that they represent the aggregate salary payments received within the year. Generally, full time legislatures offer annual salaries whereas part-time legislature salaries are calculated on a daily rate tied to the number of legislative session days during the year. Appendix Table B shows a summary of this information for states in the sample.

Finally, full time and part time legislatures likely face different opportunity costs by running for office, as part time legislators are expected to maintain outside employment. When testing the election configurations predicted by the expected payoff political entry model against actual results, some specifications differentiate between full time and part time outside income. Contributions received by candidates during the study comes from the National Institute on Money in Politics' Follow the Money API.

Descriptive Statistics

Table 13 shows the frequencies of each of the seven possible election entrance configurations for the full sample and the three election years. The most frequently observed district configuration is where both Democratic and Republican candidates enter, occurring about 52% of the time. This is consistent with the observation of two-party strongholds in congressional elections. Among uncontested elections, Democrat only elections consistently account for about 20% of observations in the full sample and across the three year subsets. Republican only elections account for about 15% of observations, but were unusually high during 2016. This year had a notable decrease in the frequency of two-party Republican and Democrat elections. Elections with all three party-candidates entering account for about 7% of observations.

Table 14

Entry Configurations in State Lower Chamber Legislative Elections

Configuration	Full Sample	2016	2018	2020
Democrat Only	1784	642	633	509
	(20%)	(20%)	(19%)	(20%)
Republican Only	1391	663	383	345
	(15%)	(21%)	(12%)	(13%)
Independent Only	2	1	0	1
Democrat and Republican	4681	1479	1849	1353
	(52%)	(47%)	(57%)	(53%)
Democrat and Independent	248	84	96	68
	(3%)	(3%)	(3%)	(3%)
Republican and Independent	211	90	50	71
	(2%)	(3%)	(2%)	(3%)
All Enter	666	201	253	212
	(7%)	(6%)	(8%)	(8%)
Observations	8,983	3,160	3,264	2,559

Notes. Data on election configurations comes from Ballotpedia.

Table 14 shows the descriptive statistics for the variable in the analysis. The second column shows the sample mean for all districts in the sample. In terms of entrance, Democratic candidates enter about 82% of elections in the sample, followed by Republicans at 77%, and Independents at 13%. The average district representative serves approximately 72,000 residents, with an average income of $81,000 and unemployment rate of 4.1%.

The third, fourth, and fifth columns subsample the data to only include observations where a candidate from the respective political party entered. For several of the explanatory variables, one political party tends to be significantly greater or less than the sample average. For example, the percentage of residents that have never been married is about 33% in districts where Democratic candidates entered but 29.3% in those with Republican entrants. This would be consistent with the rationale that candidates believe that certain characteristics about a district's population affect their likelihood of success, influencing their entry decisions. Democratic and Republican party-candidates tend to have the greatest distance between their subgroup averages, while the Independent subsample average tends to fall between the Democratic and Republican subsample averages.

Table 15

Characteristics of District with Democratic, Republican, and Independent Entrants

	All District	Democratic Entrants	Republican Entrants	Independent Entrants
		Entrants		
Democrat	82.1	100	76.9	81.1
Republican	77.4	72.5	100	77.8
Independent	12.5	12.4	12.6	100
		Incumbency		
Democrat	39.7	48.2	25.4	37.6
Republican	45.9	36.9	59.3	40.9
		Election		
District Democrats	46.5	53.5	35.4	47.1
District Republicans	51.6	44.6	62.9	49.3
District Independents	1.81	1.81	1.67	3.51
Democrat Contributions	$104,349	$126,912	$96,999	$98,507
Republican Contributions	$98,841	$95,471	$127,642	$101,445
Independent Contributions	$875	$559	$955	$5,772
Salary	$36,641	$37,890	$35,721	$34,542

Table 14

	All District	Democratic Entrants	Republican Entrants	Independent Entrants
Per Diem	$15,327	$15,256	$15,251	$13,884
		Demographic		
Population	72,642	75,045	73,373	71,358
Median Age	38.8	38.6	39.4	38.4
Unmarried	31.9	33	29.3	32.2
High School or Less	40.6	39.9	40.2	38.2
African American	11.8	12.7	8.2	9.8
Hispanic	11.5	12.3	9.9	11.8
Pacific Islander or Native American	1.82	2	1.7	2.5
		Economic		
Unemployment	4.1	4.2	3.9	4.1
Military	0.42	0.39	0.46	0.38
District Income	$81,489	$82,314	$82,925	$83,220
No Health Ins.	9.7	9.6	9.4	9.6
		Industry		
Natural Resources	2.03	1.8	2.3	1.9
Manual Labor	22.9	22.4	23.7	22.4
Professional Labor	10.1	18.8	17.9	18.9
Observations	8,983	7,379	6,949	1,127

Differences between the party-subsample and sample means guide the study's priors about entry. In terms of representative salary, Democratic candidates tend to enter elections that offer larger salary packages compared to the other parties. Compared to the full-sample, districts with Democratic entrants tend to have higher proportions of African American, Native American, Asian, and Hispanic or Latino residents. Districts with Republican entrants tend to have older residents, more military members and manual laborers, and higher proportions of residents with a high school diploma or less.

Estimation of Success Probabilities for Political Party-Candidates

The purpose of the following section is twofold: first to identify the causal effect of district and election characteristics on each political party's success probabilities and evaluate key differences between these estimates, and second to estimate election success probabilities to be used in the expected payoffs model. In this section, we use several specifications of probit models to estimate the following equation:

$$\Pr\left(Win_{djt}\right) = \Phi(\alpha_j + \beta_j X_{djt} + \gamma_j State_t + \eta_j Year_t)$$

The dependent variable, Win_{djt}, represents a binary variable taking on the value of 1 if a candidate from party *j* wins their respective district election *d* during election year *t* (*t* = 2016, 2018, 2020). X_{djt} is a vector of district characteristics that could affect party victory, separated out into five categories: candidate characteristics (incumbency status and campaign contributions), district preferences, demographics, economic characteristics, and employment by industry. $State_t$ are binary variables to take into account the state in which the election took place. $Year_t$ are binary variables for the year in which the election was held. The study controls for, but does not report, estimates for state and year effects. The term Φ denotes the standard normal cumulative distribution function.

Probit Marginal Effects Estimates

The marginal effects at sample means of probit model estimates for Democratic and Republican party-candidates are presented in Tables 3 and 4 respectively. Due to the infrequency of election wins for independent candidates (only 21 of 8,983 observations), marginal effect estimates obtained from these specifications have little statistical or economic significance. As such, marginal effects for the independent model are not presented, but probit estimates are presented in Appendix Table E. In all the specifications, standard errors are robust to heteroscedasticity and for panel specifications, are clustered to allow for within-district correlation of the error term (with the exception of a pooled probit comparison model).

Column 1 of Tables 15 and 16 presents the results of a pooled probit model with no state or year controls. In this model, district variables are divided into five categories. Of these, characteristics about the candidate (prior incumbency and contributions received), district preferences, and demographic characteristics of the district have significant impacts on the probability that a given political party wins their district's election, and marginal effects are statistically significant at the 1% confidence level.

Table 16

Impact of District Characteristics on Democrat Victory Probability (Probit Estimates: Marginal Effects at Means)

	(1)	(2)	(3)
	Candidate Characteristics		
Democrat Incumbent	0.4954*** (0.0189)	0.5153*** (0.0242)	0.5393*** (0.0265)
Republican Incumbent	-0.4315*** (0.0189)	-0.4554*** (0.0239)	-0.4297*** (0.0255)
Democrat Contributions	0.00433*** (0.00085)	0.00544*** (0.00109)	0.00478*** (0.00103)
Republican Contributions	-0.00403*** (0.00122)	-0.00505*** (0.00151)	-0.00496*** (0.00135)
Independent Contributions	-0.0392*** (0.0143)	-0.0425** (0.0177)	-0.043** (0.0214)
	District Preferences		
District Democrats	0.0086*** (0.0005)	0.0119*** (0.0009)	0.0109*** (0.001)
District Independents	0.0068*** (0.0014)	0.0088*** (0.0022)	0.007*** (0.0022)
	Demographic		
Population (x10,000)	-0.0129*** (0.0022)	-0.0164*** (0.00295)	0.0448 (0.0348)
Median Age	0.0177*** (0.0029)	0.0237*** (0.0043)	0.0212*** (0.0058)
Unmarried	0.0304*** (0.0029)	0.0383*** (0.0047)	0.0334*** (0.0052)
High School or Less	-0.0105*** (0.0016)	-0.0154*** (0.0022)	-0.0178*** (0.0025)
African American	0.0067*** (0.0013)	0.0094*** (0.0018)	0.0263*** (0.0035)
Hispanic	0.011*** (0.0015)	0.0145*** (0.0023)	0.0166*** (0.0023)

Table 15

	(1)	(2)	(3)
Pacific Islander or Native American	0.0137***	0.0195***	0.0159***
	(0.0037)	(0.0048)	(0.0049)
	Economic		
Unemployment	0.0092	0.0038	0.0003
	(0.0087)	(0.0118)	(0.0136)
Military	-0.0056	-0.0065	-0.0113
	(0.0058)	(0.0087)	(0.0104)
District Income (x10,000)	-0.0189**	-0.0272***	-0.00991
	(0.0083)	(0.00744)	(0.00918)
No Health Ins.	-0.0035	-0.0069	0.0235***
	(0.0033)	(0.0043)	(0.0065)
	Industry		
Natural Resources	-0.0093*	-0.0104	-0.0147**
	(0.0051)	(0.0075)	(0.0071)
Manual Labor	-0.001	-0.0004	-0.0004
	(0.0026)	(0.0034)	(0.004)
Professional Labor	0.0128***	0.0215***	0.0223***
	(0.0026)	(0.0041)	(0.0045)
Panel	No	Yes	Yes
Year and State Controls	No	No	Yes
Observations	8,983	8,983	8,983
Pseudo R^2	0.7718	0.7765	0.7996

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Table 17

Impact of District Characteristics on Republican Victory Probability (Probit Estimates: Marginal Effects at Means)

	(1)	(2)	(3)
Candidate Characteristics			
Democrat	-0.4772***	-0.4983***	-0.5081***
	(0.0185)	(0.024)	(0.0262)
Republican	0.4429***	0.4699***	0.4508***
	(0.0188)	(0.0241)	(0.026)
Democrat Contributions	-0.00479***	-0.00611***	-0.00532***
	(0.00095)	(0.00118)	(0.00108)
Republican Contributions	0.00463***	0.0059***	0.00579***
	(0.00146)	(0.0018)	(0.0016)
Independent Contributions	-0.00897**	-0.0115**	-0.0137*
	(0.00381)	(0.00456)	(0.00765)
District Preferences			
District Democrats	-0.0086***	-0.0122***	-0.0109***
	(0.0005)	(0.001)	(0.0009)
District Independents	-0.0099***	-0.0126***	-0.0105***
	(0.0016)	(0.0025)	(0.002)
Demographic			
Population (x10,000)	0.0109***	0.0139***	-0.0635*
	(0.00209)	(0.00284)	(0.0339)
Median Age	-0.0203***	-0.0276***	-0.0261***
	(0.0029)	(0.0044)	(0.0057)
Unmarried	-0.0331***	-0.0434***	-0.0382***
	(0.0029)	(0.0047)	(0.005)
High School or Less	0.0097***	0.0142***	0.0162***
	(0.0016)	(0.0023)	(0.0026)
African American	-0.0059***	-0.0083***	-0.0249***
	(0.0013)	(0.0019)	(0.0035)
Hispanic	-0.0098***	-0.0133***	-0.0156***
	(0.0016)	(0.0024)	(0.0023)
Pacific Islander or Native American	-0.0188***	-0.0267***	-0.0229***
	(0.0034)	(0.0045)	(0.0052)

Table 16

	(1)	(2)	(3)
	Economic		
Unemployment	-0.0049 (0.0086)	-0.0007 (0.0116)	0.009 (0.0133)
Military	0.0049 (0.0059)	0.0057 (0.0091)	0.0091 (0.0106)
District Income	0.0139* (0.0083)	0.022*** (0.00743)	0.00915 (0.00947)
No Health Ins.	0.0014 (0.0033)	0.0041 (0.0044)	-0.0296*** (0.0062)
	Industry		
Natural Resources	0.0081* (0.0048)	0.0092 (0.0073)	0.0166** (0.0071)
Manual Labor	0.0001 (0.0025)	-0.0006 (0.0034)	0.0001 (0.0041)
Professional Labor	-0.0137*** (0.0027)	-0.0217*** (0.0042)	-0.0244*** (0.0047)
District Random Effect	No	Yes	Yes
State and Year Controls	No	No	Yes
Observations	8,983	8,983	8,983
Pseudo R^2	0.7713	0.7766	0.8006

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Candidate characteristics refer to variables unique to an individual candidate within a district, including whether a candidate is an incumbent, and the contributions received during the election year. If a candidate is an incumbent, it would be expected that they would be more likely to win election in subsequent terms, and discourage rivals from other political parties. Consistent with the previous literature, prior incumbency has significant effects on the probability of victory for both incumbents and their rivals. For incumbent candidates running for re-election, the probability of victory is higher than non-incumbents (about .49 percentage points higher for Democratic candidates and .45 percentage points higher for Republican candidates). Challenging an incumbent on the other hand has a negative impact on the probability of victory. Democratic candidates challenging Republican candidates face a reduction in their probability of victory by about .43 percentage points. Republican candidates challenging Democratic incumbents face a penalty of .47 percentage points. These results are consistent in sign across year and part time/full time subsamples (see Appendix Tables F and G), and fairly consistent in magnitude (marginal effects tended to be more subdued in full time legislatures).

Generally it would be expected that candidates who spend more on campaigning are more likely to win elections. Candidate level expenditure data is only available for a few states; however, contribution data to candidates is well recorded even at state and local levels. For this part of the study, it is assumed that the monetary value of campaign expenditures equals that of the amount received in campaign contributions. As expected, self-contributions positively affect the probability of election success and contributions received by rivals lower the probability. A $10,000 increase in campaign contributions received by a candidate increases the probability of success by about .0043 percentage points for Democratic candidates and .0046 percentage points for Republicans. Across samples, rival contributions negatively affect the probability of success, but vary by political party. For Democratic candidates, a $10,000 increase in contributions reduces the probability of success by about .004 percentage points if the contributions were received by a Republican candidate and .039 percentage points if the recipient was an Independent candidate. The probability of a Republican success is reduced by about

.0047 percentage points for a proportional increase in Democratic contributions and .0089 percentage points for an increase in Independent contributions.

Existing political preferences for political parties are likely to significantly contribute to the probability of certain political-party candidates being elected. This may be particularly relevant in state assembly elections compared to larger election geographies, if district voters are more likely to vote on the basis of political party affiliation rather than by individual candidates. Ideally, political party voter registration data would provide a good measure of district political preferences; however, only a few states provide party registration data at the state legislative level. Since this level of data is unattainable, the study uses historic vote shares received by political parties during the 2014 state assembly elections as a proxy variable for district preferences. The primary drawback of this approach is that the vote share for a given political party is dependent upon a candidate from that party entering that district's election. The non-entry of a political party results in a zero vote share for the candidate, and will tend to overstate the proportion of voters of entering parties and understate that of non-entering parties. Regardless, non-entry does convey information about preferences, suggesting that candidates have abstained from entering on the grounds that they are unlikely to win under a certain party affiliation.

The study includes vote shares received by a district's Democratic and Independent candidates from the 2014 state assembly elections as proxy variables for district preferences. For Democratic candidates, a 1% increase in the vote share of district Democrats increases the probability of winning by .0086 percentage points. For Republican candidates, an equivalent increase results in a decline of .0086 percentage points. Surprisingly, an increase in the vote share received by Independent candidates' yields opposite estimates for Republican and Democratic candidates. Results for a 1% increase in the Independent vote share results in a .0068 percentage point increase in the probability of victory for Democrats, but a .0099 percentage point decrease for Republican candidates. These results are largely consistent across year and work time subsamples.

Seven demographic characteristics are considered in this study: population, median age, percent of population never married, percent with a high school diploma or less, percent African American, percent Hispanic or Latino, and percent Pacific Islander or Native American. Districts with older citizens, more unmarried residents, and more diverse ethnicities increase the probability of election success for Democratic Party candidates, while decreasing the probability of success for Republican candidates. Districts with larger populations and more residents with a high school degree or less are more likely to elect Republican candidates over Democratic ones. Estimates for marital status, education, and racial makeup are consistent across subsamples and with previous findings. Estimates for population are not consistent across specifications and subsamples. All estimates are statistically significant in this specification at the 1% confidence level.

Of the five categories of explanatory variables, those pertaining to economic and industry characteristics of the district tend to have less statistical and economic significance. District average income, percent of workers employed in natural resource industries, and percent of workers employed in professional trades are among those with statistical significance. For Democratic candidates, a $10,000 increase in a district average income results in a .018 percentage point reduction in the probability of success, statistically significant at the 5% confidence level. An equivalent increase in average income predicts .014 percentage point increase in the probability of a Republican win, significant at the 10% confidence level, suggesting that districts with higher incomes are more likely to vote for Republican candidates. A 1% increase in the proportion of workers employed in natural resource occupations decreases the probability of a Democratic win by .0093 percentage points but increases the probability of a Republican win by .0081 percentage points. Both marginal effect estimates are statistically significant at the 10% confidence level. Finally, estimates suggest that workers employed in professional trades (financial services, information, managerial, and office workers) are more likely to elect Democratic candidates. A 1% increase in the percent of professional workers within a district predicts a .0128 percentage point increase in the probability of a Democratic win but a .0137 percentage point decrease in the probability of a Republican win, with both effects statistically significant at the 1% confidence level.

The remaining two specifications use random effects probit models to estimate the causal effects of district characteristics on success probabilities. In the sample, the same districts are observed over the course of three election years. While the identities of candidates may vary across elections, it would be reasonable to think that a district that elects a given-party candidate in one year may be predisposed to electing a candidate of the same party in subsequent election years. District heterogeneity may make some districts more likely to elect a given party than others, and adopting a panel approach accounts for this unobserved heterogeneity. A panel probit specification is preferred over a logit specification as it allows for the inclusion of time-invariant characteristics (such as state).

Column 2 through 3 show the results of random effects probit model estimations, to control for within-district heterogeneity. Column 2 shows estimates for a simple panel probit model, without state and year controls (identical to the pooled probit model). Column 3 includes state and year controls, in addition to district covariates. Given the detail of discussion in the previous paragraphs, only meaningful differences between specifications will be mentioned here.

Coefficient estimates for candidate characteristics and district preferences are very similar across specifications, all consistent in sign with only small differences in magnitude. The first major difference between estimates occurs for district population. Specification 1 and 2 predict that a 10,000 person increase in population reduces the probability of a Democrat winning between .0129 to .0164 percentage points. Controlling state and year in the third specification predicts a .044 percentage point *increase* in the probability of a Democratic win, albeit not with statistical significance. Results for the probability of Republican wins are similar; an equivalent increase in population predicts an increase in the probability of a Republican winning between .011 and .013 percentage points in the first to specifications, but a .064 percentage point decline in the third specification (statistically significant at the 10% confidence level). Given that states vary in district sizes (for example, Pennsylvania has more populous districts than Alaska), it makes intuitive sense that controlling for state accounts for these baseline differences, resulting in different, more precise marginal effects.

While consistent in sign, marginal effects pertaining to a district's racial composition tend to estimate stronger impacts on success probabilities in panel specifications. For example, an increase in the proportion of African American residents predicts an increase in the probability of a Democratic candidate winning. In the pooled probit model, the predicted increase in probability is estimated as .0067 percentage points. In the panel models, the predicted increase is between .0097 and .026 percentage points respectively. This observation is consistent across the three racial variables in the study, as well as for Republican candidates.

One variable that was not significantly significant in the pooled probit model, percentage of the district without health insurance coverage, is statistically significant in the third specification, with a change in sign for the two political parties. Districts without coverage are more likely to elect Democratic candidates, with a 1% increase in the proportion of uncovered residents increasing the probability of a Democratic win by .023 percentage points (statistically significant at the 1% confidence level). For Republicans, an equivalent percentage increase results in a 2.9 percentage point decline in the probability of winning.

As predicted probabilities obtained from these estimates are to be used in subsequent parts of the study, it is especially important for chosen specifications to fit the observed data well. In order to compare goodness-of-fit across specifications, pseudo R^2 statistics are calculated for each of the specifications for the two samples. In OLS models, coefficients are estimated in order to minimize the sum of the squared residuals, while probit and logit model coefficients are estimated by maximum likelihood estimation. Since the dependent variables in probit models are binary and parameters are estimated by MLE, the OLS approach to R^2 is not applicable. Instead, pseudo R^2 values are calculated based on McFadden's approach to treat the log likelihood of the full model of predictors as the sum of squared errors and the log likelihood of the intercept only model as the total sum of squares. McFadden's Pseudo R^2 is calculated as follows:

$$R^2 = 1 - \frac{\ln(L(M_{full}))}{\ln(L(M_{intercept}))}$$

Since each specification uses the same dataset and dependent variable, the study can reasonably compare pseudo R^2 statistics across specifications to determine goodness-of-fit. For both the Democrat and Republican models, the third specification, the random effects probit with state and year interactions, has the greatest pseudo R^2 value. Additionally, parameter estimates obtained in this model tend to make intuitive sense within the context of existing findings. As such, specification three is considered the preferred model when constructing predicted probabilities of victory for the three political parties for each district.

Predicted Probabilities

One of the primary goals of this study is to examine the role of monetary payoffs in the party-candidates entry decision. In this study, entry decisions are assumed to be based on payoffs to party candidates. If the payoff of entry is greater than a threshold level (for example, \$0 or the forgone salary from assuming office), the candidate enters the election. The payoff of political participation consists of several components, some of which are always incurred regardless of performance, and others that are only realized with an election win.

There are two ways in which victory-dependent components can be treated. The first is to assume that candidates ignore the probabilistic nature of such benefits, and run for election with the assumption that a win is guaranteed. With these candidate beliefs, nominal values of salary, per diem pay, and other benefits should be used. If it is to be assumed that candidates incorporate probabilistic beliefs into their decision making, victory-dependent components must be adjusted accordingly.

The study obtains predicted probabilities of an election win for each of the political parties using each district's covariates. When estimating predicted probabilities the third specification (random effects probit with state and year controls) is used for the construction of Democratic and Republican probabilities. The pooled probit model estimates are used for the construction of independent probabilities. After obtaining fitted values from the probit model, the standard normal cumulative distribution function is used to obtain predicted probabilities.

Table 17 summarizes descriptive statistics comparing observed election outcomes for the three political parties versus the predicted probabilities estimated by the random effects probit model. The first three rows show the mean, standard deviation, and minimum and maximum values for observed wins for each of the three political parties. On average in the sample, Democratic candidates win about 45.6% of elections, Republican candidates win about 54.04%, and Independents win about 0.24%. Rows four through six show descriptive statistics for the predicted probabilities of victory for the three parties. Estimates are remarkably close to the true observed outcomes: on average, the probability that a Democratic candidate wins election is 45.3%, 54.2% for Republican candidates, and 0.1% for independent candidates. This suggests that the predicted probabilities obtained from the probit model fit the observed data relatively well, and can be used in further analysis requiring success probability estimates.

Table 18

Descriptive Statistics for Observed Political Party Outcomes and Predicted Probabilities

	Mean	Std. Deviation	Minimum	Maximum
		Observed Outcome Values		
Democrat Win	45.6%	49.8	0	1
Republican Win	54.04%	49.8	0	1
Independent Win	0.24%	4.9	0	1
		Predicted Probabilities		
Democrat Win	45.3%	46.1	0	1
Republican Win	54.2%	46.2	0	1
Independent Win	0.1%	2.3	0	1

Monetary Payoffs and Candidate Entry

In the previous section, the study estimated victory probabilities for each of the political parties under consideration. Several specifications of the success model were estimated and marginal effects were presented. In addition to a discussion of the causal effects of district characteristics on a given political party's success probability, the study uses the probit estimates obtained from the previous section to obtain predicted success probabilities for each district and political party.

Using the predicted probabilities of success for the observations in the sample, the study estimates the expected payoff for each party-candidate of political participation. The goals of this section are two-fold: first to estimate the causal effects that individual components of a party's expected payoffs have on party entry, and second to compare entrance predicted by the expected payoffs approach against the actual data.

Constructing Expected Payoffs

The full model of the expected payoff to a candidate from party *j* entering a given district election *d* is calculated as follows:

$$\Pi_{jd} = \widehat{Pr}\left(Win_{jd}\right)(Salary_d + PerDiem_d + Benefits_d) + Contributions_{jd} - Expenses_{jd} - FilingFees_d$$

Where $\widehat{Pr}\left(Win_{jd}\right)$ is the estimated probability that party-candidate *j* wins district election *d*. Three components of expected payoff are only received upon successfully winning election: salary, Per Diem, and any external benefits (such as expense allowances), and are scaled by the predicted probability. Campaign contributions, campaign expenses, and filing fees are received and incurred regardless one's performance in an election, thus are not multiplied by the probability. Π_{jd} represents the monetary value of a candidate's expected payoff.

Table 18 shows sample averages for the components nominal and expected payoffs for the full sample and by political party-entrants. Each row shows the total payoff component for a two-year term. For example, in Alaska state representatives receive a base salary of $50,400 making the term salary $100,800. Rows 1 through 5 show the average nominal term salary, per diem pay, benefits, campaign contributions, and filing fees at their stated values. Columns 2 through 4 show the averages for subsamples based on which political parties enter. At nominal values, Democratic entrants tend to enter elections with higher average salaries and per diem pay. Independent entrants face higher filing fees and receive less in campaign contributions than other parties. Republican entrants on average received the most in campaign contributions.

Table 19

Components of Expected Payoff

	All Political Parties	Democrat	Republican	Independent
	Nominal Values			
Term Salary	$74,363.82	$75,780	$71,422	$69,084
Term Per Diem	$30,654.89	$30,512	$30,502	$27,768
Other Benefits	$1,148.46	$1310.40	$924.67	$1463.38
Campaign Contributions	$68,022.05	$126,912.82	$127,642.33	$5,772.46
Filing Fees	$255.88	$198.86	$242.16	$259.61
	Sample: Expected Payoff Components			
Expected Term Salary	$24,688.11	$36,538.30	$37,438.01	$88.03
Expected Term Per Diem	$10,166.68	$14,022.25	$16,424.42	$53.36
Expected Other Benefits	$380.78	$732.55	$405.56	$4.23
	Entrants: Expected Payoff Components			
Expected Term Salary	$48,120	$44,223.85	$48,003.14	$641.70
Expected Term Per Diem	$10,514	$16,989.14	$21,075.92	$394.79
Expected Other Benefits	$1,816	$884.92	$518.18	$33.21
	Total Nominal and Expected Payoffs			
Total Nominal Payoff	$105,910	$107, 404	$102,606	$98,574
Total Expected Payoff	$34,979	$61,908	$69,354	$809

Rows 6 through 8 show the average expected payoffs from salary, per diem pay, and benefits pay. Since campaign contributions and filing fees are incurred regardless of actual performance, there is no need to adjust these values. Because expected payoffs are based on success probabilities, political party subgroup averages can be calculated regardless of whether or not a candidate entered in a given election. There are several differences between sample and subsample averages when probabilistic components are incorporated into the payoff structure. Unsurprisingly, the expected salary, per diem, and benefits are smaller than their stated counterparts, as there is a chance such benefits will not be received in the case of failure. Nominal values for these payoff components are about three times that of the expected values. Another change is that Republican candidates enter elections with greater expected payoffs, while Democratic candidates enter elections with higher nominal benefits.

Rows 9 through 11 shows sample averages for only party-candidates that choose to enter elections. Using our theoretical approach, which states that candidates enter when expected payoff are greater than some baseline level, it would be expected that expected payoffs would be higher in the entrants-only party subsamples over the full model. For Democratic and Republicans, expected salary, per diem pay, and benefits are about 1.2 times greater for entering candidates than the full sample averages. For independent candidates, these components are about 7 times greater.

Finally, rows 12 and 13 show the total values of nominal and expected payoffs for the full sample and entrants from the three political parties. Given limitations in obtaining campaign expenditure data, it is assumed that campaign expenditures equal campaign contributions, eliminating these two components from total payoff. The study relaxes this assumption later in the section. Using the full sample of data, on average the total nominal payoff to candidates is about 3 times greater than the expected payoff. Among entrants, sample Democrats have the greatest average nominal payoffs, while Republican entrants have the greatest expected payoffs. Between the two measures, Independent candidates consistently have lower estimated payoffs than the two major parties, about .94 times the estimated nominal payoff and .01 times the expected payoff of entrants.

Probit Marginal Effects Estimations of Party-Candidate Entry Decisions

Having estimated expected payoffs and individual payoff components of political participation for each political party over all the districts in the sample, the study proceeds to estimate probit models estimating the causal effect different variables have on the probability of a given political party-candidate entering a district election. In this section, the study adopts two different approaches to the estimation of entry probabilities. In the first approach, political entry is presented as a reduced-form model, with a party-candidate making their entry decisions based on district characteristics. The second approach, in contrast, assumes that the most relevant factors in a candidate's decision to enter are based on their expectations of the payoffs they are likely to receive. Increasing expected benefits is predicted to positively affect a candidate's decision to enter while decreases in cost predict a decline in the probability of entry. For estimation purposes, a "quasi-structural" approach is adopted to estimate the marginal effects payoff components have on entry.

Political Entry Based on District Characteristics. The first approach to modeling political entry is based on the assumption that party-candidates make their decision to enter based on different characteristics pertaining to an election district. These include characteristics about rivals, monetary incentives, demographics, and economic attributes. In this subsection, the study uses several specifications of probit models to estimate the following equation:

$$\Pr(Enter_{djt}) = \Phi(\alpha_j + \beta_j X_{djt} + \gamma_j State_t + \eta_j Year_t)$$

The dependent variable, $Enter_{djt}$, represents a binary variable taking on the value of 1 if a candidate from party *j* enters their respective district election *d* during election year *t* (*t* = 2016, 2018, 2020). X_{djt} is a vector of district characteristics that could affect party entry, including rival entry, district preferences, salary and per diem pay, as well as previously stated demographic and economic characteristics. $State_t$ are binary variables to take into account the state in which the election took place. $Year_t$ are binary variables for the year in which the election was held. The study controls for, but does not report, estimates for state and year effects. The term Φ denotes the standard normal cumulative distribution function.

Probit marginal effects estimates at sample means for Democratic, Republican, and Independent party-candidates are presented in Table 19. Two specifications are compared: a pooled probit model (Columns 1:3) and a random effects probit model (Columns 4 through 6). In all the specifications, standard errors are robust to heteroscedasticity and are clustered to allow for within-district correlation of the error term (with the exception of a pooled probit comparison model).

Table 20

Characteristics-Based Political Entry Probabilities (Probit Estimates: Marginal Effects at Means)

	(1)	(2)	(3)	(4)	(5)	(6)
			Election District Characteristics			
Democrat Incumbent		-0.1568*** (0.0109)	-0.0399*** (0.00887)		-0.1193*** (0.01217)	-0.0350*** (0.00781)
Republican Incumbent	-0.0990*** (0.00783)		-0.0359*** (0.00858)	-0.0790*** (0.00827)		0.03142*** (0.00762)
District Democrats	0.00181*** (0.00012)	-0.0020*** (0.00014)	0.00011 (0.00015)	0.00153*** (0.00013)	-0.0015*** (0.00016)	0.00008 (0.00014)
District Independents	-0.0005 (0.00036)	-0.0017*** (0.00033)	0.00298*** (0.00042)	-0.00038 (0.0004)	-0.0012*** (0.00028)	0.00264*** (0.00047)
Salary	0.00049 (0.00189)	0.00279 (0.00251)	-0.0107** (0.00425)	0.00039 (0.00162)	0.00187 (0.00181)	-0.0092** (0.00368)
Per Diem	0.00132 (0.00399)	-0.00516 (0.00349)	0.0039 (0.00484)	0.00078 (0.00349)	-0.00412* (0.00245)	0.00381 (0.0037)
Filing Fees	-0.05918 (0.04086)	-0.00145 (0.02628)	-0.024 (0.03864)	-0.04625 (0.03327)	-0.00167 (0.01968)	-0.0206 (0.03269)
			Demographic Characteristics			
Population	0.0287*** (0.00538)	0.0213*** (0.00561)	0.0128** (0.00579)	0.0255*** (0.00464)	0.016*** (0.00487)	0.011* (0.00588)
Median Age	0.00072 (0.00068)	0.00196*** (0.00066)	-0.00032 (0.00083)	0.00064 (0.00057)	0.00153*** (0.00057)	-0.0003 (0.00079)
Unmarried	0.00331*** (0.00036)	-0.0031*** (0.00028)	-0.0006* (0.00033)	0.00282*** (0.00034)	-0.0023*** (0.00028)	-0.00059* (0.00031)
High School or Less	0.00236*** (0.0004)	-0.0014*** (0.00026)	-0.0011*** (0.00039)	0.00192*** (0.00035)	-0.0010*** (0.00024)	-0.00096** (0.00038)
African American	0.00301** (0.00148)	-0.0018*** (0.00061)	0.00044 (0.00077)	0.00246** (0.00123)	-0.0012*** (0.00039)	0.0003 (0.00062)
Hispanic	0.00049 (0.00189)	0.0053** (0.00212)	0.00276 (0.00278)	-0.00072 (0.00211)	0.00405** (0.00168)	0.00365 (0.00246)
Pacific Islander or Native American	0.00132 (0.00399)	0.0056*** (0.00166)	-0.00267 (0.00186)	-0.00227* (0.00118)	0.00432*** (0.00127)	-0.00225 (0.00182)

Table 19

	(1)	(2)	(3)	(4)	(5)	(6)
			Economic Characteristics			
Unemployment	-0.00036 (0.0025)	0.00012 (0.00166)	-0.0088*** (0.00211)	0.00031 (0.00171)	-0.00027 (0.00143)	-0.0076*** (0.00196)
Military	-0.00277** (0.00115)	-0.0023** (0.0009)	-0.00131 (0.00124)	-0.00041 (0.00091)	-0.00172** (0.00077)	-0.00127 (0.00118)
District Income	-0.00028 (0.00194)	0.0017 (0.00137)	-0.00251 (0.0016)	-0.0045*** (0.00087)	0.00111 (0.00117)	-0.00204 (0.00152)
No Health Ins.	-0.00096 (0.001)	0.00387*** (0.00058)	0.00051 (0.00072)	-0.0012*** (0.00048)	0.00287*** (0.00052)	0.00038 (0.00069)
Natural Resources	-0.0054*** (0.00101)	-0.00033 (0.0007)	0.00395*** (0.00094)	0.00117 (0.00074)	-0.00008 (0.00059)	0.0035*** (0.0009)
Manual Labor	-0.0015*** (0.00056)	-0.0020*** (0.00014)	-0.0399*** (0.00887)	-0.0790*** (0.00827)	-0.0015*** (0.00016)	-0.0020 (0.00152)
Professional Labor	0.00186** (0.00085)	-0.0017*** (0.00033)	-0.0359*** (0.00858)	0.00153*** (0.00013)	-0.0012*** (0.00028)	0.00038 (0.00069)
Panel	No	No	No	Yes	Yes	Yes
Observations	8,653	8,983	8,983	8,653	8,983	8,983
Pseudo R^2	0.3753	0.5507	0.0885	0.3903	0.5584	0.0999

Notes: Salary, Per Diem, and Expense Allowance are scaled by 10,000. Filing fees are scaled by 100.

Significance Codes: "***" 0.01 "**" 0.05 "*" 0.1

Rows 1 through 7 of Table 19 show the marginal effects election characteristics have on the probability of entry for Democratic, Republican, and Independent candidates, respectively. The first three columns show marginal effects estimates for the three political party types from the pooled probit model. The study examines six variables within this category: rival incumbency, district preferences, salary, per diem pay, and filing fees.

The presence of a rival incumbent is likely to discourage the entry of other political party candidates, given that they are more likely to win elections than challengers, as shown previously. Based on our definition, incumbency is defined by the individual rather than the party. All sample incumbents are, by definition, entrants, so the inclusion of self-incumbent variables is irrelevant. This was not a problem in estimating success probabilities, as not all incumbents win reelection. The political affiliation of the rival incumbent has different effects on challenging political parties. The presence of a Democratic incumbent reduces the probability of Republican entry by 0.15 percentage points in the pooled probit model and 0.11 percentage points in the random effects model. For Independent challengers, Democratic incumbency reduces the probability of entry by 0.039 percentage points in the pooled probit model and 0.035 percentage points in the random effects model. A Republican incumbent reduces the probability of Democratic entry in the pooled probit by 0.09 percentage points and 0.07 percentage points in the random effects model. Independents experience a reduction in the probability of entry by 0.035 and 0.031 percentage points, in the pooled probit and random effects models respectively. For Independent candidates the identity of the Incumbent party has little difference on the marginal effect estimate of entry; however, Republican candidates are less likely to enter with the presence of a Democratic rival. All estimates are statistically significant at the 1% confidence level.

As with rival incumbency, district preferences can affect the probability of a certain candidate winning and the likelihood of certain types of candidates entering a district election. The political party affiliation of potential candidates can also be thought of as a product of the district in which they seek election. If there are more Democrats than other political parties in a district, then of the people who do run for office, it would be expected that they are more likely to be a Democrat. Likewise, from the intuition of the median voter theorem, potential candidates belonging to minor parties have an incentive to run under the party affiliation of the majority party, as mimicking the majority party is associated with a larger vote share.

Vote shares received by Democratic and Independent candidates for the 2014 state assembly elections are used as proxy variables for district preferences. The vote share of Democratic candidates has positive marginal effects on the probability of entry for Democratic and Independent candidates, but a negative impact on the entry of Republican candidates. A 1% increase in the Democratic vote share increases the probability of Democratic entry by .0018 percentage points in the pooled probit model and .0015 in the random effects model. A similar increase reduces the probability of Republican entry by .0021 percentage points in the pooled probit model and .0015 percentage points in the random effects model. Both estimates are statistically significant at the 1% confidence level, while the estimate for Independent candidates is not statistically significant.

An increase in the Independent vote share reduces the probability of Democratic and Republican entry but increases the probability of Independent entry. A 1% increase in the Independent vote share decreases the probability of Republican entry by .0017 to .0012 percentages points in the pooled probit and random effects models respectively, significant at the 1% confidence level. An equivalent increase in the vote share increases the probability of Independent entry by .0029 percentage points in the pooled probit model and .0026 percentage points in the random effects models, with both estimates statistically significant at the 1% confidence level. The marginal effect estimates for Democratic candidates are not statistically significant in either specification, but are negative.

Each state legislature sets its own standard for salary pay, per diem pay, and filing fees. Generally, there is not much (if any) variation between districts within a state in regards to these payments and fees. Marginal effect estimates for the impact these variables have on entry probabilities are unimpressive. Increases in salary and per diem pay generally has a positive impact on entry, while filing fees are unambiguously negative. However, there are notable exceptions. Independent candidates tend to enter elections with lower salary pay, such that a $10,000 increase in salary pay is predicted to reduce the probability Independent entry by 0.01 to 0.009 percentage points in the pooled probit and random effects models, with both estimates statistically significant at the 5% confidence level. Likewise, Republican candidates enter elections with lower per diem pay. However, given that there is little variation in state assembly benefits, it is more likely that lower pay is correlated with the increased entry of Republican and Independent candidates, rather than by a causal relationship.

Differences in district demographic characteristics are likely to result in differences in political party entry. The study considers seven demographic characteristics: population, median age, percent of the district never married, percent with a high school diploma or less, percent of African American, percent Hispanic, and percent Pacific Islander or Native American. Within this section, one point to distinguish is that while in the election success probability model only one political-party wins election, multiple candidates can enter a district election. Coefficient estimates in the success probabilities models are expected to be positive for the party benefiting from that attribute and negative for the other parties, as election victory is a zero sum game. However, in terms of entry, some characteristics may benefit multiple political parties. For example, salary is a characteristic expected to increase the probability of entry for all political-party candidates, one party does not exclusively benefit.

The probability of entry for the three political parties is estimated to increase with an increase in a district's population. The marginal effect is strongest for Democratic candidates. A 10,000 person increase in a district's population increases the probability of seeing a Democratic entrant by 0.028 to 0.025 percentage points in the pooled probit and random effects specifications. Both estimates are statistically significant at the 1% confidence level. A similar increase in population increases the probability of Republican entry by .021 to 0.016 percentage points, with both estimates statistically significant at the 1% confidence level. For Independent candidates, the marginal effect of an equivalent population increase is estimated to be 0.012 percentage points in the pooled model (statistically significant at the 5% confidence level) and 0.011 percentage points in the random effects model (statistically significant at the 10% level).

Increases in a district's median age is estimated to have positive marginal effects on the probability of Democratic and Republican entry but negative effects on the probability of Independent entry; however, these estimates are only statistically significant for Republican entry. A 1 year increase in a district's median age is estimated to increase the probability of Republican entry by 0.0019 percentage points in the pooled model and 0.0015 percentage points in the random effects model. Both estimates are statistically significant at the 1% confidence level.

Marginal effect estimates suggest that increases in the proportion of residents who have never been married increases the probability of observing Democratic entrants and decreases the probability of observing Republican and Independent entrants. A 1% increase in the percentage of unmarried residents increases the probability of Democratic entry by 0.033 to 0.028 percentages points, with both estimates statistically significant at the 1% confidence level. A similar increase reduces the probability of Republican entry by 0.031 to 0.023 percentage points and Independent entry by 0.0006 to 0.0005 percentage points. Marginal effect estimates for Republican entry are statistically significant at the 1% confidence level and significant at the 10% level for Independent candidates.

Somewhat surprisingly, Democratic candidates are estimated to be more likely to enter an election when the percent of residents with a high school degree or less increases, despite being less likely to win such elections. A 1% increase in this demographic increases the probability of Democratic entry by 0.002 percentage points in the pooled model and 0.01 percentage points in the random effects model. Both estimates are statistically significant at the 1% confidence level. The marginal effect on Republican entry is estimated to reduce the probability of observing a Republican entrant by 0.0014 to 0.0011 percentage points and is statistically significant at the 1% confidence level for both estimates. The Independent probability of entry is estimated to be reduced by 0.001 to 0.0009 percentage points, with the estimate for the pooled probit model being statistically significant at the 1% confidence level and the estimate for the random effects model being statistically significant at the 5% level.

The study looks at the marginal effects associated with changes in the proportions of residents of three racial groups: African Americans, Hispanics, and Pacific Islanders and Native Americans. Marginal effect estimates for changes in the percentage of African American residents are statistically significant for Democratic and Republican entry. Democratic and Independent candidates are more likely to enter elections with an increase in African American residents, while Republican candidates are less likely to enter. The marginal effect associated with Democratic entry following a 1% increase in the proportion of African American residents is estimated to be between 0.003 and 0.0024 percentage points. The probability of Republican entry is reduced by 0.0018 to 0.0012 percentage points. Changes in the proportion of Hispanic residents positively affects the probability of Republican entry. A 1% increase in the percent of Hispanic residents increases the probability of Republican entry by 0.005 percentages points in the pooled probit model and 0.004 percentage points in the random effects model. Both estimates are statistically significant at the 5% confidence level. Finally, a 1% increase in the proportion of the sum of Pacific Islander and Native American residents increases the probability of observing Republican entry from 0.005 to 0.004 percentage points, but decreases the Democratic entry probability by 0.002 percentage points in the random effects model.

Interestingly, demographic characteristics marginal effect results suggest a discontinuity between entrance and success. A certain political party candidate may be less likely to enter an election given its demographic characteristics, but would have relatively high probabilities of success if they did enter. One explanation is that political-party candidates may choose to enter in response to changes in certain characteristics, while neglecting existing district preferences. For example, a candidate may enter an election in a district with lower educational attainment with the purpose of making educational improvements. However, this approach ignores that voters have self-selected certain outcomes; voters may not believe that educational reform is important. As a result, a candidate may enter but may have a low probability of success because there is a mismatch between the district's and candidate's political preferences.

The final category of explanatory variables consists of economic and industry employment characteristics including unemployment, proportion of military members, average district income, percent without health insurance coverage, percent employed in natural resource trades, proportion of manual labor workers, and proportion of professional workers. Unemployment is estimated to have a negative marginal effect on the probability of Independent entry, reducing the probability of entrance by 0.008 to 0.007 percentage points. Entrance for all three political parties is less likely as the proportion of people serving in the military increases. Increases in district income is estimated to have a negative marginal effect on the probability of Democratic and Independent entry, but a positive, insignificant marginal effect on Republican entry. A $10,000 increase in the average income of a district decreases the probability of Democratic entry by 0.002 to 0.0045 percentage points. The estimate for the random effects model is statistically significant at the 1% confidence level.

Changes in the proportion of residents employed in natural resource trade increases the probability of entrance for Independent candidates, but reduces the probability for Democratic ones. The estimated marginal effect is an increase in the probability of entrance between 0.0035 and 0.0039 percentage points for Independent candidates. The reduction in probability to Democratic candidates is between 0.005 and 0.001 percentage points. Increases in the proportion of workers employed in manual labor trades is estimated to have negative marginal effects for all three political parties.

Finally, professional labor positively affects the probability of entrance for Democratic candidates, but reduces the probability for Republican and Independent candidates. A 1% increase in the proportion of professional workers increases the probability of Democratic entry between 0.0018 and 0.0015 percentage points. For Republicans the estimated reduction is between 0.0017 and 0.0012 percentage points. Independent candidates experience a reduction in the probability of entrance between 0.035 and 0.0003 percentage points.

Overall, coefficient and standard error estimates between the pooled probit and random effects probit model are largely equivalent. In terms of goodness-of-fit, Pseudo R^2 estimates are slightly higher for panel specifications.

Political Entry Based on Expected Payoffs. There are two main drawbacks to using the first approach towards the estimation of entry probabilities. First, modeling entry probabilities as a function of a large number of variables is unlikely to accurately reflect a candidate's decision calculus in practice. Individual variables, such as the median age, may be correlated with the entry of a specific party, but it would likely be inappropriate to claim candidates enter because of a district's characteristic composition. The basis of much decision making is made under the assumption that financial incentives matter, and a strictly reduced-form approach is unlikely to be a good reflection of the practical, heuristic decision making of political candidates.

A second concern with the characteristics-based approach is that while model estimates can be used to estimate the probability that a party enters, no clear decision rule exists regarding whether a candidate *would* enter given certain characteristics. Predicting entry configuration from data requires the researcher to make an assumption regarding entry from probability. For example, one such decision rule could be that a candidate chooses to enter if their predicted probability of entry is greater than 50%. In contrast, the study's model of expected payoff derives a decision rule based on the theory that a candidate chooses to enter if their expected payoff is greater than a baseline value (either $0 or some measure of forgone income). As such, the criteria of predicting the entry of a party-candidate, and election configurations is clearer with a quasi-structural model.

In this subsection, the study uses several specifications of probit models to estimate the following equation:

$$\Pr\left(Enter_{djt}\right) = \Phi(\alpha_j + \beta_j \Pi_{djt} + \gamma_j State_t + \eta_j Year_t)$$

As was in the previous subsection, the dependent variable, $Enter_{djt}$ represents a binary variable taking on the value of 1 if a candidate from party *j* enters their respective district election *d* during election year *t* (*t* = 2016, 2018, 2020). Π_{djt} is a vector of the estimated components of party-candidate *j*'s payoff including expected salary, per diem pay, filing fees, and campaign costs. $State_t$ binary variables to take into account the state in which the election took place. $Year_t$ are binary variables for the year in which the election was held. The study controls for, but does not report, estimates for state and year effects. The term Φ denotes the standard normal cumulative distribution function.

Tables 20, 21, and 22 show estimates for the marginal effects that different components of expected payoff have on the probabilities of entry for Democratic, Republican, and Independent party candidates respectively. Marginal effects are calculated at sample means, and standard errors are robust to heteroscedasticity and clustered at the district level in panel specifications. The first specification of the expected payoffs models shows the marginal effect that expected salary has on the probability of entry. A $10,000 increase in the salary that a candidate expects to receives increases the probability Democrat entry by 0.008 percentage points (statistically significant at the 10% confidence level), Republican entry by 0.009 percentage points (significant at the 5% confidence level), and Independent entry by 0.23 percentage points (significant at the 5% confidence level). Given that Independent candidates have lower expected payoff (and are more willing to enter at negative expected payoffs), it makes intuitive sense that an equivalent increase in salary would have stronger impacts on Independent entry.

Table 21

Probability of Democrat Candidate Entry (Probit Estimates: Marginal Effects at Means)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Expected Salary	0.0081* (0.004)	0.0035 (0.0023)	0.0022 (0.0019)	0.0032 (0.0023)	0.0032 (0.0023)	0.00094 (0.0011)	0.00143 (0.0015)
Expected Per Diem		0.0070* (0.0038)	0.00524 (0.00364)	0.00676* (0.00374)	0.00676* (0.00374)	0.00154 (0.00154)	0.00335 (0.0031)
Filing Fees		-0.0003 (0.0001)	-0.0002 (0.00015)	-0.00025 (0.00019)	-0.00025 (0.00019)	-0.00007 (0.00009)	-0.00226 (0.00285)
Republican Entrant			-0.0037 (0.0032)	-0.005671 (0.003908)	-0.00567 (0.003909)	-0.0013 (0.001450)	-0.0027 (0.002686)
Independent Entrant			-0.00120 (0.00126)	-0.0017 (0.00159)	-0.0017 (0.0015965)	-0.0007 (0.0009)	-0.0027 (0.0028)
Rival Penalties							
Costs (Linear)				0.00001 (0.00001)			
Costs (Log Costs)					0.00001 (0.00001)		
Panel	No	No	No	No	No	Yes	Yes
State and Year Controls	No	No	No	No	No	No	Yes
Observations	8,983	8,983	8,983	8,983	8,983	8,983	8,653
Pseudo R^2	0.2504	0.2816	0.2858	0.2868	0.2869	0.3014	0.3749

Notes: Salary and Per Diem are scaled by 10,000. Filing fees and rival penalties are scaled by 100.

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Table 22

Probability of Republican Candidate Entry (Probit Estimates: Marginal Effects at Means)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Expected Salary	0.0090** (0.0041)	0.0059** (0.0026)	0.0059*** (0.0022)	0.0059*** (0.0023)	0.0059*** (0.0023)	0.00059 (0.00059)	0.0025 (0.0022)
Expected Per Diem		0.0091** (0.0039)	0.0098*** (0.0033)	0.0098*** (0.0033)	0.0097*** (0.0033)	0.00094 (0.0009)	0.0041 (0.00303)
Filing Fees		-0.0003* (0.0002)	-0.0003** (0.0002)	-0.0003** (0.0002)	-0.0003** (0.00017)	-0.00004 (0.00005)	-0.00078 (0.142)
Democrat Entrant			-.01139** (0.0053)	-0.0113** (0.00532)	-0.0113** (0.0053)	-0.0011 (0.0012)	-0.0052 (0.0043)
Independent Entrant			0.0011 (0.00089)	0.0011 (0.0008)	0.00106 (0.00088)	0.00003 (0.0001)	-0.0003 (0.0006)
Rival Penalties							
Costs (Linear)				0.00001 (0.00001)			
Costs (Half)					0.00001 (0.00001)		
Panel	No	No	No	No	No	Yes	Yes
State and Year Controls	No	No	No	No	No	No	Yes
Observations	8,983	8,983	8,983	8,983	8,983	8,983	8,983
Pseudo R^2	0.3631	0.4000	0.4047	0.4050	0.4052	0.4481	0.5167

Notes: Salary and Per Diem are scaled by 10,000. Filing fees and rival penalties are scaled by 100.

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Table 23

Probability of Independent Candidate Entry (Probit Estimates: Marginal Effects at Means)

	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Expected Salary	0.239** (0.105)	0.0883 (0.218)	0.086 (0.214)	0.086 (0.214)	0.086 (0.214)	0.0523 (0.114)	0.0293 (0.073)
Expected Per Diem		0.254 (0.341)	0.255 (0.337)	0.255 (0.337)	0.255 (0.337)	0.197 (0.225)	0.226 (0.208)
Filing Fees		-0.0015** (0.00068)	-0.0015** (0.00068)	-0.0015** (0.0006)	-0.0015** (0.0006)	-0.0013** (0.0006)	-0.03736 (0.03624)
Democrat Entrant			-0.0072 (0.0097)	-0.0072 (0.0098)	-0.0072 (0.0098)	-0.0141 (0.0092)	-.038*** (0.0108)
Republican Entrant			0.0023 (0.0086)	0.0024 (0.0086)	0.0023 (0.0087)	0.0011 (0.0081)	-0.0165* (0.0087)
Rival Penalties							
Costs (Linear)				0.00001 (0.00001)			
Costs (Half)					0.00001 (0.00001)		
Panel	No	No	No	No	Yes	Yes	Yes
State and Year Controls	No	No	No	No	No	No	Yes
Obs.	8,983	8,983	8,983	8,983	8,983	8,983	8,983
Pseudo R^2	0.0069	0.0078	0.0079	0.0079	0.0079	0.0351	0.0911

Notes: Salary and Per Diem are scaled by 10,000. Filing fees and rival penalties are scaled by 100.

Significance Codes: "***" 0.01 "**" 0.05 "*" 0.1

Specification 2 includes two other components of expected payoff in addition to expected salary: per diem pay and filing fees. Consistent with theory, increases in per diem pay are estimated to have positive effects on the probability of entry while increases in filing fees have negative impacts for the three parties. With the inclusion of these two components, the estimated marginal effect of expected salary decreases for the three party candidates. A $10,000 increase in expected per diem pay increases the probability of Democratic entry by 0.007 percentage points (significant at the 10% confidence level), Republican entry by 0.009 percentage points (significant at the 5%) confidence level, and Independent entry 0.25 percentage points. The estimated marginal effect of filing fees on Democratic entry is not statistically significant; however, is for Republican and Independent candidate entry. A $100 increase in filing fees predicts a 0.0003 percentage point decrease in the probability of Republican entry and a 0.0015 percentage point decrease in Independent entry probabilities.

Specification 3 includes variables indicating the presence of a rival and their party affiliation, in addition to the components of expected payoff. Generally coefficient estimates for these variables are negative, but not necessarily statistically significant. For Democratic candidates, the presence of a Republican rival decreases the probability of entry by 0.003 percentage points and the presence of an Independent rival decreases entry by 0.0012 percentage points. For Republican candidates, a Democratic rival decreases entry by 0.01 percentage points (statistically significant at the 5% confidence level) and an Independent rival is predicted to increase entry by 0.001 percentage points. A Democratic rival decreases Independent entry by 0.007 percentage points but a Republican rival increases entry by 0.002 percentage points.

In previous parts of the study, the difference between campaign contributions and campaign expenditure was assumed to be equal to zero and candidates break even. Campaign contribution data, but not expenditure data, is available for all the candidates within the study. While expenditure data is not available, specifications 4 and

5 attempt to incorporate campaign costs elements based on rivalry. In these specifications, candidates face different net campaign contributions based on their predicted probabilities of success. For each district, the party candidate with the maximum amount of campaign contributions is used as a benchmark for the minimum amount of expenditure needed to be competitive in the election. Party-candidates with lower probabilities of success are required to spend additional amounts to remain competitive, thus, subsequent candidates face increasing costs. In specification 4, cost structures follow a linear form: the top-ranked candidate spends the base expenditure, the second-ranked candidate spends twice the maximum amount, and the third ranked candidate spends three times the base expenditure. Specification 5 uses a natural logarithmic cost structure equal to the base expenditure times *ln(rank)*, with a rank of 1 indicating the top ranked candidates. After constructing predicted costs for party-candidates in each district, the net amount of campaign financing is calculated by subtracting predicted costs from the base contribution, leaving only the additional costs incurred by lower ranks.

Estimation results for both the linear and logarithmic rivalry expenditure specifications produce unsubstantial results for the three political parties. Coefficient estimates are near zero and none are statistically significant. Cost variables are dropped from subsequent specifications.

In order to account for unobserved heterogeneity between districts, specifications 6 and 7 use panel probit models in order to estimate marginal effects. In specification 6, five variables are included: expected salary, per diem pay, filing fees, and the two appropriate rival entrant variables. Results tend to be subdued; none of the coefficient estimates are statistically significant in the Democrat and Republican samples. However, increases in filing fees continues to reduce the probability of Independent entry by about 0.001 percentage points and is statistically significant at the 5% confidence level.

Specification 7 continues to use the random effect probit model, with the inclusion of state and year controls. When controlling for these variables, coefficient estimates tend to be larger in magnitude but not necessarily statistically significant. The marginal effect on the probability of entry given a $10,000 increase in a candidate's expected salary increases Democratic entry by about 0.001 percentage points, 0.002 percentage points for Republican candidates, and 0.029 percentage points for Independent candidates. Likewise, equivalent increases in per diem pay increase the probability of entry by 0.003 percentage points for Democratic candidates, 0.004 percentage points for Republican candidates, and 0.26 percentage points for Independent candidates. Filing fees are estimated to have a negative impact on the probability of entry for the three parties, ranging from a reduction of 0.0007 percentage points for Republican candidates to 0.03 percentage points for Independents. Finally, the presence of rival candidates predicts negative impacts on entry probabilities. In previous specifications, Republican and Independent entrants had positive predicted effects on one another. When controlling for state and year effects, the presence of a Republican candidate reduces the probability of Independent entry by 0.016 percentage points (statistically significant at the 10% confidence level). An independent entrant reduces the probability of Republican entry by 0.0002 percentage points.

Across specifications, pseudo R^2 values become significantly greater with the inclusion of state and year controls, and respecifying the model as a panel probit model. R^2 estimates are highest for models measuring Republican entry, and lowest for those measuring Democratic entry. While total expected payoff may be relevant to a party-candidate's entry decisions, changes in individual components of payoffs are unlikely to have a strong marginal effect on entry.

Concluding Remarks. The purpose of this subsection is to estimate the effect various classes of explanatory variables have on the probability of a given political party-candidate entering using two approaches. In the first approach, candidate entry decisions are assumed to be made on the basis of district characteristics. Characteristics of districts, such as its demographic and economic attributes, influence what types of political party candidates will emerge. Using this approach suggests that party affiliation and party-candidate entry is a product of the characteristics of a district, the types of people who are willing to run for political office have similar political preferences to that of the districts they represent. Candidates do not choose to run for office because of certain district characteristics, their own preferences are shaped by that of the district, making some political party-candidates more likely to enter than others.

In the second approach, entry decisions are made on the basis of expected payoffs to candidates. Instead of assuming that the candidates who are most likely to run for office have similar preferences to (or mimic those) of the district, candidates of a given political party choose whether or not to enter based on if they will benefit from running for office. The total payoff to candidates is composed of some payoff components that are only realized if the candidates wins their respective general election (salary and per diem pay) and other benefits received (campaign contributions) and costs (filing fees and campaign expenses) incurred regardless of election outcomes. District characteristics influence the probability of a given party-candidate winning, and determine the monetary value of expected salary and per diem pay.

Overall, modeling political entry based on a district characteristics is likely a better approach to identifying determinants of political entry, and relative changes in probability associated with unit-changes in explanatory variables, than the expected payoffs approach. The primary benefit of this approach is that it provides better insight into how individual characteristics affect entry, rather than incorporating them into a single probability component measuring the probability of success. As shown in the analysis, a characteristic can affect the probability of entry and the probability of success differently for the same political party-candidate.

Political Entry Decisions

The purpose of the previous section was to estimate the marginal effects that certain explanatory variables had on the probability of entrance for Democratic, Republican, and Independent candidates. Two approaches to estimating marginal effects were employed. In the first specification, entry decisions were modeled based on the assumption that certain attributes about a district (for example, the presence of rival incumbents, or the unemployment rate) make entry more likely for some political parties over others. Candidates in this approach generally have less autonomy over entry decisions; the observed entry of a political party is determined by the environment in which an election takes place.

The second approach models entry decisions on theoretical grounds. Candidates are assumed to enter an election if the expected payoff of running for office is greater than some baseline level. In this model, candidates can influence their own payoffs through their choice of expenditures. The primary advantage of this approach is that there is a clear decision rule in place that identifies whether or not candidates enter. Entry decisions are defined by a theoretical approach, as opposed to a benchmark chosen by the researcher.

In this section, the study uses the two approaches to modelling political entry to compare predicted entrance against observed entry decisions, for both individual parties and the election configuration. Using the full sample of data, the study compares three categories of entry decisions: the observed entry decision of the three political parties and the election configuration, the predicted entry decisions from the characteristics entry decision model, and the predicted entry decisions from the expected payoffs model.

For the characteristics entry decision model, entry decisions are based on the predicted entry probabilities obtained post-estimation. A political party candidate enters an election if their predicted probability of entry is greater than 50%. For the expected payoffs model, a candidate enters if their expected payoff is greater than the baseline level of forgone income. Several specifications of the expected payoffs model are used, including additional payoff components with each specification.

Upon constructing predicted entry decisions for each party, the predicted political party entry decision is compared to the observed entry decision for that district. If the predicted decision matches the observed decision, the prediction is considered a success. After entry decisions have been compared for the three political parties, and success and failures have been determined, the predicted election configuration is compared to the observed configuration. Within a district, if the predicted model results in success for all three political parties, the predicted model is said to have correctly predicted the observed election configuration. If a single political party test results in failure, the model fails to predict the configuration.

Table 23 shows the number of correctly predicted election configurations and entry decisions for the three political parties. In each cell, the top number shows the count of correctly predicted values and the number in parentheses shows the percentage of correct predictions relative to the sample size (N = 8,983). In this model, the forgone income, relevant to entry decisions based on expected payoff, is assumed to be equal to $0. Column 1 shows the count and percentage of correctly predicted election configurations, Column 2 shows the number of correctly predicted Democrat entries, Column 3 shows Republican entry, and Column 4 shows Independent entry tests.

Table 24

Observed Entry Decisions vs. Predicted Entry Configurations

Model	Correctly Predicted	Democratic Entry	Republican Entry	Independent Entry
Model 1: District Characteristics	5878 (65.4%)	7728 (86%)	7953 (88.5%)	7844 (87.3%)
	Forgone Income = $0			
Model 2: Expected Salary	5328 (59.3%)	7647 (85.1%)	7551 (85.5%)	7682 (85.5%)
Model 3: Salary, Per Diem, Benefits, Filing Fees	5128 (57.1%)	7402 (82.4%)	7586 (84.4%)	7675 (85.4%)
Model 4: Salary, Per Diem, Benefits, Campaign contributions	3898 (43.4%)	6233 (69.4%)	7153 (79.6%)	7885 (87.9%)
Model 5: Salary, Per Diem, Filing Fees, Linear Expenditures	3889 (43.1%)	6232 (69.3%)	7152 (79.6%)	7871 (87.6%)
Model 6: Salary, Per Diem, Filing Fees, Logarithmic Campaign Expenses	4337 (48.1%)	6560 (73.1%)	7335 (81.7%)	7876 (87.7%)

Table 23

Observed Entry Decisions vs. Predicted Entry Configurations

Model	Correctly Predicted	Democratic Entry	Republican Entry	Independent Entry
Forgone Income = District Average				
Model 7: Salary, Per Diem, Filing Fees, Logarithmic Campaign Expenses	1088 (12.1%)	2671 (29.7%)	4281 (47.7%)	7858 (87.5%)
Forgone Income: Part Time = $0, Full Time = District Average				
Model 8: Salary, Per Diem, Filing Fees, Logarithmic Campaign Expenses	3748 (41.7%)	5771 (64.3%)	7030 (78.3%)	7863 (87.5%)

Model 1 in the first row compares the district characteristics based entry model against the observed observations. Under the assumption that a candidate enters if their probability of entry is greater than 0.50, the model correctly predicts 65% of the election configurations. For individual political parties, this model correctly predicts 86% of Democratic entry decisions, 88% of Republican decisions, and 87% of Independent entry decisions.

Model 2 if the first of the specifications that use the expected payoffs approach to entry. In this specification, the only component of total payoffs is the expected salary for each political party candidate. A candidate enters if their expected salary is greater than $0. This specification correctly predicts 59% of the observed district election configurations. In terms of individual political parties, the model correctly predicts 85% of Democratic, Republican, and Independent entry decisions.

Model 3 includes expected per diem pay, benefits, and filing fees into the expected payoffs model. The inclusion of these components causes a slight decline in the number of correctly predicted election configurations and political party decisions, but within 5 percentage points of Model 2.

Models 4 includes campaign contributions as a part of total payoffs without making assumptions about campaign expenditures. Including contribution data results in a significant decrease in accuracy, with the model only correctly predicting 43% of the observed election configurations. Accuracy decreases for Democratic and Republican subgroup tests: predicting 69% of Democratic entry decisions and 79% of Republican decisions.

Models 5 and 6 reintroduce the campaign expenditure component into the model. Model 5 uses a linear expenditure structure such that the candidate with the highest probability of winning has a net difference between contributions and expenditures of $0, the second entrant pays the baseline expenditure, and the third entrant pays twice the baseline. Model 6 uses a logarithmic cost structure: the first entrant breaks even, and subsequent entrants pay the difference between the baseline expenditure and *ln(rank)* times the baseline expenditure. Model 5 does not offer much improvement from Model 4; correctly predicting nearly identical percentages for the four categories. Using a logarithmic cost structure; however, does improve accuracy from Models 4 and 5. Model 6 correctly predicts 48% of election configurations, 73% of Democratic decisions, 81% of Republican decisions, and 87% of Independent entry decisions.

Models 7 and 8 keep the specification of Model 6 but change the baseline forgone income expected payoffs are compared against. Additional specifications with the changes in baseline opportunity costs can be found in Appendix Table H. In model 7, the forgone income is equal to a district's average income over a two-year term. In model 8, the study distinguishes between opportunity costs faced by part-time and full-time legislators. The income forgone by full time legislators is the term district average, while part-time workers have no forgone income. Of these specifications, Model 7 offers the least predictive power, correctly predicting only 12% of election configurations. Model 8 does a significantly better job predicting entry. 41% of observed election configurations were correctly predicted, 64% of Democratic decisions, 78% of Republican decisions, and 87% of Independent decisions.

Overall, the characteristics based model continues to be a better measure of political entry for both individual political party decisions and election configurations in general. Model 2, which only considers expected salary, produces close party entry decisions predictions to Model 1, but does not predict election configurations as well.

Conclusions

Political competition has the power to influence who enters elections, and ultimately what positions receive attention. Political party candidates may choose to enter elections for many different reasons, and a clear understanding of the decision calculi faced by party candidates allows us to make more predictive assessments of competitiveness. This study presents two methods of determining the probability of individual party-candidate entry and predicting political competitiveness.

In the first approach, political party entry decisions are influenced by election, demographic, and economic attributes of election districts. Political party-candidates that “fit” the district’s preferences are more likely to enter than others. The second approach of modeling political entry is inspired by empirical models of firm profit and entry found in the empirical industrial organization. Using this theoretical approach, the study models candidate entry in terms of expected payoffs.

Using state assembly data over three election years (2016, 2018, and 2020) from thirty-two states, the study estimates the marginal effects that changes in district characteristics and payoff components have on the probabilities of political entry and election success for Republican, Democratic, and Independent candidates. The analysis finds substantial evidence that suggests changes in district attributes affect the entry and success probabilities of political parties differently. Unit-changes in some characteristics (such as population) estimate similar marginal effects on the entry probabilities of all political parties; whereas most characteristics (median age, marital status, ethnicity, etc.) have differing effects on entry probabilities. Prior incumbency and previous district preferences have strong marginal effects on entry probabilities.

The study finds that modeling political entry probabilities using the characteristics approach generally provides better estimates of the marginal effects of component changes and predictive estimates of entry than the expected payoffs approach. The characteristics model correctly predicts 65% of the observed election configurations, while the best specification of the expected payoffs model correctly predicts 59% of the actual configurations.

References

Ansolabhere, S., Snyder Jr., J. M., & Stewart III, C. (2001). Candidate Positioning in U.S. House Elections. *American Journal of Political Science*, 136-159.

Bender, B., & Haas, T. C. (1996). Electoral Equilibria with Entry: Some Simulations. *Public Choice* , 303-14.

Berry, S. T. (1992). Estimation of a Model of Entry in the Airline Industry. *Econometrica*, 889-917.

Box-Steffenmeier, J. M. (1996). A Dynamic Analysis of the Role of War Chests in Campign Strategy. *American Journal of Political Science*, 352-371.

Breaux, D. (1990). Specifying the Impact of Incumbency On State Legislative Elections: A District-Level Analysis. *American Politics Quarterly*, 270-286.

Bresnahan, T. F., & Reiss, P. C. (1991). Entry and Competition in Concentrated Markets. *Journal of Political Economy*, 977-1009.

Brudnick, I. A. (2018). *Congressional Salaries and Allowances: In Brief.* Washington, D.C.: Congressional Research Service.

Burden, B. C. (2004). Candidate Positioning in US Congressional Elections. *British Journal of Political Science*, 211-227.

Carson, J. L. (2003). Strategic Interaction and Candidate Competition in U.S. House Elections: Empirical Applications of Probit and Strategic Probit Models. *Political Analysis*, 368-380.

Carson, J. L., Engstrom, E. J., & Roberts, J. M. (2007). Candidate Quality, the Personal Vote, and Incumbency Advantage in Congress. *American Political Science Review*, 289-301.

Ciliberto, F., & Tamer, E. (2009). Market Structure and Multiple Equilibria in Airline Markets. *Econometrica*, 1791-1828.

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical Methods for Comparing Regression Coefficients between Models. *American Journal of Sociology*, 1261-1293.

Cox, G. W., & Morgenstern, S. (1995). The Incumbency Advantage in Multimember Districts: Evidence from the U.S. States. *Legislative Studies Quarterly*, 329-349.

Downs, A. (1957). *An Economic Theory of Democracy.* New York: Harper.

Ellickson, P. B., & Misra, S. (2011). Estimating Discrete Games. *Marketing Science* , 997-1010.

Epstein, D. & Zemsky, P. (1995). Money Talks: Deterring Quality Challengers in Congressional Elections. *American Political Science Review*, 295-308.

Fiorina, M. P., & Abrams, S. J. (2008). Political Polarization in the American Public. *Annual Review of Political Science*, 563–588.

Gelman, A., & King, G. (1990). Estimating Incumbency Advantage without Bias. *American Journal of Political Science*, 1142-1164.

Goldenberg, E. N., Traugott, M. W, & Baumgartner, F. R. (1986). Preemptive and Reactive Spending in U.S. House Races. *Political Behavior*, 3-20.

Grosser, J., & Palfrey, T. (2014). Candidate Entry and Political Polarization: An Antimedian Voter Theorem. *American Journal of Political Science*, 127-143.

Hersch, P. L., & McDougall, G. S. (1994). Campaign War Chests as a Barrier to Entry in Congressional Races. *Economic Inquiry*, 630-641.

Hogan, R. E. (2004). Challenger Emergence, Incumbent Success, and Electoral Accountability in State Legislative Elections. *The Journal of Politics*, 1283–1303.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 41-57.

Jewell, M. E., & Breaux, D. (1988). The Effect of Incumbency On State Legislative Elections. *Legislative Studies Quarterly*, 495-514.

King, G. (1989). Representation through Legislative Redistricting: A Stochastic Model. *American Journal of Political Science*, 787-824.

Lee, D. J. (2012). Anticipating Entry: Major Party Positioning and Third Party Threat. *Political Research Quarterly*, 138–150.

Osborne, M. J., & Slivinski, A. (1996). A Model of Political Competition with Citizen-Candidates. *The Quarterly Journal of Economics*, 65-96.

Palfrey, T. R. (1984). Spatial Equilibrium with Entry. *The Review of Economic Studies*, 139-156.

Pasternoster, R., Brame, R., Mazerolle, P., & Piquero, A. (1998). Using the Correct Statistical Test for the Equality of Regression Coefficents. *Criminology*, 859-866.

Poole, K. T., & Rosenthal, H. (1984). The Polarization of American Politics. *The Journal of Politics*, 1061-1079.

Signorino, C. (2003). Structure and Uncertainty in Discrete Choice Models. *Political Analysis*, 316-344.

Squire, P. (2000). Uncontested Seats in State Legislative Elections. *Legislative Studies Quarterly*, 131-146.

Stigler, G. J. (1972). Economic Competition and Political Competition. *Public Choice*, 91-106.

Stratmann, T. (2005). Ballot Access Restrictions and Candidate Entry in Elections. *European Journal of Political Economy*, 59-71.

Tullock, G. (1965). Entry Barriers in Politics. *The American Economic Review*, 458-466.

Appendix A: Election and District Characteristics across Years

Appendix Table A

Election and District Characteristics Across Years

	All District	2016	2018	2020
	Entrants			
Democrat	0.822 (0.384)	0.762 (0.427)	0.868 (0.34)	0.838 (0.37)
Republican	0.774 (0.419)	0.77 (0.421)	0.777 (0.417)	0.775 (0.419)
Independent	0.126 (0.332)	0.119 (0.324)	0.123 (0.328)	0.138 (0.345)
	Victory			
Democrat	0.457 (0.499)	0.431 (0.496)	0.463 (0.499)	0.481 (0.5)
Republican	0.541 (0.499)	0.568 (0.496)	0.535 (0.499)	0.514 (0.5)
Independent	0.003 (0.05)	0.002 (0.04)	0.003 (0.047)	0.004 (0.063)
	Incumbency			
Democrat	0.397 (0.49)	0.364 (0.482)	0.349 (0.477)	0.499 (0.501)
Republican	0.46 (0.499)	0.459 (0.499)	0.436 (0.496)	0.491 (0.5)
	Election			
District Democrats	46.323 (35.656)	45.921 (35.794)	47.651 (35.69)	46.323 (35.656)
District Republicans	51.863 (35.893)	52.311 (36.053)	50.315 (35.917)	51.863 (35.893)
District Independents	1.765 (6.493)	1.72 (6.408)	1.972 (6.842)	1.765 (6.493)
Democrat Contributions	$104349 (362182)	$86461 (293295)	$102058 (403143)	$129360 (382132)

Appendix Table A

	All District	2016	2018	2020
Republican Contributions	$98841 (344738)	$99783 (451733)	$95235 (277145)	$102278 (259398)
Independent Contributions	$875 (15733)	$598 (6778)	$725 (6074)	$1407 (27658)
Salary	$74363.822 (55998)	$72651.661 (52069)	$73594.185 (56454)	$77459.768 (59859)
Per Diem	$30654 (34226)	$30126 (31410)	$30564 (33620)	$31423 (38106)
Filing Fees	$1148 (3981)	$578 (2639)	$1440 (4608)	$1480 (4397)
	Demographic			
Population	72642 (82693)	70592 (79876)	70461 (79856)	77957 (89205)
Median Age	38.9 (5.551)	38.684 (5.514)	38.883 (5.501)	39.187 (5.65)
Unmarried	31.9 (9.234)	31.629 (9.193)	31.685 (9.226)	32.508 (9.27)
High School or Less	40.619 (13.047)	41.282 (13.024)	40.776 (13.008)	39.598 (13.067)
African American	11.793 (16.644)	11.474 (16.462)	12.016 (16.932)	11.901 (16.497)
Hispanic	11.558 (15.018)	11.055 (14.602)	11.089 (14.559)	12.779 (16.005)
Pacific Islander or Native American	1.83 (4.643)	1.319 (4.641)	3.022 (5.148)	0.94 (3.515)
	Economic			
Unemployment	4.139 (1.803)	4.507 (1.721)	3.886 (1.514)	4.006 (2.139)
Military	0.425 (2.014)	0.432 (2.037)	0.428 (2.082)	0.414 (1.892)
District Income	$81489 (31073)	$74156 (26654)	$77828 (28277)	$95214 (34942)
No Health Ins.	9.765 (5.602)	11.055 (5.756)	9.769 (5.477)	8.166 (5.137)

Appendix Table A

	All District	2016	2018	2020
	Industry			
Natural Resources	2.04 (2.992)	2.176 (3.2)	2.08 (3.034)	1.821 (2.639)
Manual Labor	22.948 (6.643)	22.679 (6.626)	23.103 (6.732)	23.084 (6.543)
Professional Labor	18.314 (6.712)	18.11 (6.579)	18.2 (6.684)	18.712 (6.893)
Observations	8,983	3,162	3,264	2,560

Notes. Standard deviation in parentheses.

Appendix B: Comparison of State Work Expectations, Benefits, and Filing Fees

Appendix Table B

Comparison of State Work Expectations, Benefits, and Filing Fees

State	Part Time	Average Session Length	Term Salary	Term Per Diem	Term Expense Allowance	Filing Fees (Major Party)	Filing Fees (Minor Party)
Alabama	1	30	185028	0	0	925.14	0
Alaska	0	130	100800	67508	9333	30	30
Arkansas	1	40	81344	12170	0	3000	3000
California	0	291	214820	112025	0	1074	1074
Colorado	1	120	60000	56100	0	0	0
Colorado	1	120	70242	38160	0	0	0
Connecticut	1	125	56000	0	0	0	0
Delaware	1	180	92582	0	0	0	945
Delaware	1	181	89082	0	0	0	945
Florida	1	60	59394	18240	0	890	890
Georgia	1	40	34684	13840	0	400	400
Hawaii	0	106	123376	43695	0	250	250
Illinois	0	227	136757	49606	0	0	0
Indiana	1	45	50616	15030	0	0	0
Iowa	1	105	50000	34790	0	0	0
Kansas	1	60	10639	17250	0	120	120
Kentucky	1	45	56452	14340	0	200	200
Maine	1	124	23254	17453	0	0	0
Massachusetts	0	321	125890	11586	11734	0	0
Michigan	0	355	143370	0	10800	0	100
Missouri	1	150	71830	34080	0	0	100
Montana	1	67.5	11876	15338	0	15	15

Appendix Table B

State	Part Time	Average Session Length	Term Salary	Term Per Diem	Term Expense Allowance	Filing Fees (Major Party)	Filing Fees (Minor Party)
North Carolina	1	67	27902	13936	0	139	139
Ohio	0	360	124464	0	0	85	85
Oklahoma	1	147	76800	45717	0	500	500
Oregon	1	81	52656	23298	0	0	25
Pennsylvania	0	351	175236	125520	0	0	100
Rhode Island	1	165	31335	0	0	0	0
South Carolina	1	153	20800	49194.6	0	0	208
Tennessee	1	90	45244	43020	0	0	0
Texas	1	70	14400	28046	0	0	750
Utah	1	45	24954	8430	0	65	65

Appendix C: Democrat Election Victory Probit Estimates

Appendix Table C

Democrat Election Victory Probit Estimates

	Pooled Probit (Simple)	Panel Probit (Simple)	(Full Panel)
	Incumbency		
Democrat	1.3348*** (0.0594)	1.3976*** (0.0772)	1.4816*** (0.0893)
Republican	-1.1551*** (0.0576)	-1.2349*** (0.0756)	-1.1417*** (0.0761)
	Election		
District Democrats	0.0217*** (0.0012)	0.03*** (0.0023)	0.0273*** (0.0023)
District Independents	0.017*** (0.0034)	0.0222*** (0.0056)	0.0174*** (0.0053)
Democrat Contributions	0.0109*** (0.00212)	0.0138*** (0.00265)	0.012*** (0.00247)
Republican Contributions	-0.0102*** (0.00308)	-0.0128*** (0.00381)	-0.0125*** (0.00337)
Independent Contributions	-0.099*** (0.0361)	-0.108** (0.0448)	-0.108** (0.0538)
	Demographic		
Population (x10,000)	-0.0324*** (0.00555)	-0.0415*** (0.00748)	0.113 (0.0875)
Median Age	0.0445*** (0.0073)	0.0599*** (0.0109)	0.0531*** (0.0144)
Unmarried	0.0766*** (0.0073)	0.0968*** (0.0117)	0.0839*** (0.0129)
High School or Less	-0.0265*** (0.0039)	-0.0389*** (0.0055)	-0.0446*** (0.0062)
African American	0.0169*** (0.0032)	0.0237*** (0.0045)	0.0659*** (0.0086)
Hispanic	0.0277*** (0.0038)	0.0366*** (0.0058)	0.0415*** (0.0057)

Appendix Table C

	Pooled Probit (Simple)	Panel Probit (Simple)	Full Panel
Pacific Islander or Native American	0.0345*** (0.0091)	0.0494*** (0.0121)	0.0398*** (0.0123)
	Economic		
Unemployment	0.023 (0.022)	0.009 5 (0.0297)	0.0006 (0.034)
Military	-0.014 (0.0146)	-0.0164 (0.0219)	-0.0282 (0.0261)
District Income	-0.0477** (0.0209)	-0.069*** (0.0189)	-0.0249 (0.0233)
No Health Ins.	-0.0088 (0.0083)	-0.0174 (0.0108)	0.059*** (0.0162)
	Industry		
Natural Resources	-0.0233* (0.0127)	-0.0264 (0.019)	-0.0369** (0.0177)
Manual Labor	-0.0025 (0.0064)	-0.0011 (0.0086)	-0.001 (0.01)
Professional Labor	0.0323*** (0.0066)	0.0543*** (0.0104)	0.0559*** (0.0112)
Constant	-4.6998*** (0.5591)	-6.1205*** (0.8629)	-6.1083*** (1.2223)
Panel	No	Yes	Yes
Year and State Controls	No	No	Yes
Observations	8,983	8,983	8,983
Pseudo R^2	0.7718	0.5	0.8215

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Appendix D: Republican Election Victory Probit Estimates

Appendix Table D

Republican Election Victory Probit Estimates

	Pooled Probit (Simple)	Panel Probit (Simple)	Full Panel
	Incumbency		
Democrat	-1.28*** (0.0575)	-1.3451*** (0.0763)	-1.3885*** (0.0866)
Republican	1.1833*** (0.0577)	1.2687*** (0.0762)	1.1991*** (0.0784)
	Election		
District Democrats	-0.0216*** (0.0012)	-0.0306*** (0.0023)	-0.0273*** (0.0023)
District Independents	-0.0247*** (0.0039)	-0.0317*** (0.0063)	-0.0263*** (0.005)
Democrat Contributions	-0.012*** (0.00237)	-0.0154*** (0.00296)	-0.0133*** (0.00272)
Republican Contributions	0.0116*** (0.00366)	0.0149*** (0.00456)	0.0145*** (0.00402)
Independent Contributions	-0.0226** (0.00957)	-0.029** (0.0115)	-0.0343* (0.0192)
	Demographic		
Population (x10,000)	0.0274*** (0.00523)	0.0351*** (0.00721)	-0.159* (0.0853)
Median Age	-0.051*** (0.0073)	-0.0694*** (0.0111)	-0.0653*** (0.0143)
Unmarried	-0.083*** (0.0071)	-0.1093*** (0.0117)	-0.0957*** (0.0126)
High School or Less	0.0242*** (0.004)	0.0357*** (0.0058)	0.0406*** (0.0065)
African American	-0.0149*** (0.0032)	-0.0209*** (0.0046)	-0.0623*** (0.0087)
Hispanic	-0.0245*** (0.0038)	-0.0334*** (0.0059)	-0.0391*** (0.0057)

Appendix Table D

	Pooled Probit (Simple)	Panel Probit (Simple)	Full Panel
Pacific Islander or Native American	-0.0473***	-0.067***	-0.0574***
	(0.0084)	(0.0114)	(0.0129)
	Economic		
Unemployment	-0.0123	-0.0017	0.0225
	(0.0215)	(0.0292)	(0.0334)
Military	0.0123	0.0144	0.0227
	(0.0149)	(0.0228)	(0.0265)
District Income	0.0349*	0.0554***	0.0229
	(0.0209)	(0.0187)	(0.0238)
No Health Ins.	0.0035	0.0101	-0.0741***
	(0.0083)	(0.011)	(0.0154)
	Industry		
Natural Resources	0.0202*	0.023	0.0415**
	(0.0119)	(0.0183)	(0.0178)
Manual Labor	0.0002	-0.0015	0.0001
	(0.0063)	(0.0086)	(0.0103)
Professional Labor	-0.0343***	-0.0546***	-0.0611***
	(0.0066)	(0.0106)	(0.0116)
Constant	5.403***	7.2093***	7.3057***
	(0.528)	(0.8407)	(1.1869)
Part Time Interactions	No	No	No
District Random Effect	No	Yes	Yes
State and Year Controls	No	No	Yes
Observations	8,983	8,983	8,983
Pseudo R^2	0.7713	0.484	0.8405

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Appendix E: Independent Election Victory Probit Estimates

Appendix Table E

Independent Election Victory Probit Estimates

	Pooled Probit (Simple)	Panel Probit (Simple)	(Full Panel)
	Incumbency		
Democrat	-0.6423**	-0.8797**	-0.912***
	(0.2527)	(0.3581)	(0.3276)
Republican	-0.5729**	-0.5137	-0.4811
	(0.2694)	(0.4313)	(0.4071)
	Election		
District Democrats	0.0018	0.0061	0.0059
	(0.0048)	(0.0097)	(0.0092)
District Independents	0.0265***	0.0466***	0.0445***
	(0.0053)	(0.0141)	(0.0124)
Democrat Contributions	0.00296**	0.00317	0.00315
	(0.0015)	(0.00436)	(0.00309)
Republican Contributions	-0.364***	-0.569**	-0.564***
	(0.105)	(0.241)	(0.201)
Independent Contributions	0.454***	0.708***	0.679***
	(0.106)	(0.22)	(0.189)
	Demographic		
Population (x10,000)	-0.225**	-0.359**	-0.333*
	(0.114)	(0.178)	(0.171)
Median Age	0.0189	0.053	0.0437
	(0.0276)	(0.052)	(0.046)
Unmarried	-0.0036	0.0073	-0.0029
	(0.0261)	(0.051)	(0.045)
High School or Less	0.0069	0.0081	0.0031
	(0.0204)	(0.0349)	(0.033)
African American	-0.1365	-0.2066	-0.2002
	(0.0983)	(0.1949)	(0.1693)
Hispanic	-0.0638**	-0.0653	-0.0628
	(0.0307)	(0.0463)	(0.0428)

Appendix Table E

	Pooled Probit (Simple)	Panel Probit (Simple)	(Full Panel)
Pacific Islander or Native American	0.0103 (0.0168)	0.0215 (0.0282)	0.0205 (0.027)
Economic			
Unemployment	-0.0359 (0.0824)	-0.0184 (0.1536)	0.0029 (0.1321)
Military	-0.0394 (0.1463)	0.0591 (0.1501)	0.0582 (0.134)
District Income	0.185 (0.0706)	0.114 (0.0725)	0.0619 (0.0606)
No Health Ins.	0.0395 (0.0416)	0.0502 (0.0686)	0.0587 (0.0703)
Industry			
Natural Resources	-0.0883* (0.0496)	-0.1515** (0.0684)	-0.1593** (0.0696)
Manual Labor	0.0759*** (0.0242)	0.1163*** (0.0403)	0.1126*** (0.0369)
Professional Labor	0.0519** (0.0245)	0.0462 (0.0655)	0.0638 (0.0587)
Constant	-6.5288*** (2.4358)	-10.2199** (4.7108)	-9.281** (4.1567)
District Random Effect	No	Yes	Yes
State and Year Controls	No	No	Yes (No State)
Observations	8,983	8,983	8,983
Pseudo R^2	0.5875	0.6104	0.6142

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Appendix F: Comparison of Estimates by Year

Appendix Table F

Comparison of Estimates by Year

	2016 Dem	2018 Dem	2020 Dem	2016 Rep	2018 Rep	2020 Rep
			Candidate Characteristics			
Democrat Incumbent	0.55248*** (0.04198)	0.58031*** (0.03765)	0.49472*** (0.04585)	-0.5420*** (0.04235)	-0.5402*** (0.03694)	-0.4507*** (0.04389)
Republican Incumbent	-0.3977*** (0.0404)	-0.3360*** (0.03524)	-0.6040*** (0.03684)	0.40182*** (0.04008)	0.34803*** (0.03698)	0.60879*** (0.03741)
District Democrats	0.01185*** (0.00111)	0.00722*** (0.00085)	0.00889*** (0.00108)	-0.0119*** (0.00111)	-0.0076*** (0.00084)	-0.0077*** (0.00105)
District Independents	-0.00374 (0.00282)	0.00829*** (0.00209)	0.0097*** (0.00257)	-0.00234 (0.00265)	-0.0116*** (0.00247)	-0.0114*** (0.0027)
Democrat Contributions	0.00268** (0.00104)	0.0065*** (0.00186)	0.00442*** (0.00107)	-0.0027*** (0.00106)	-0.0090*** (0.00188)	-0.0043*** (0.00103)
Republican Contributions	-0.0031*** (0.00111)	-0.00264 (0.00206)	-0.0092*** (0.00172)	0.00319*** (0.00114)	0.00635*** (0.00198)	0.00902*** (0.00166)
Independent Contributions	-0.0158 (0.0185)	-0.036* (0.0219)	-0.055 (0.0602)	-0.0293* (0.0166)	0.012 (0.0319)	-0.0119* (0.00696)
Population	-0.0903 (0.0672)	0.112** (0.0456)	0.00384 (0.0337)	0.0915 (0.0682)	-0.135*** (0.046)	-0.0139 (0.0315)
Median Age	0.02374*** (0.00721)	0.01534** (0.00707)	0.01428* (0.00867)	-0.0245*** (0.00719)	-0.0207*** (0.00719)	-0.01857** (0.00871)
Unmarried	0.02957*** (0.00658)	0.02288*** (0.00558)	0.03292*** (0.0072)	-0.0284*** (0.00659)	-0.0280*** (0.00569)	-0.0337*** (0.00683)
High School or Less	-0.01068** (0.0044)	-0.0166*** (0.00371)	-0.0262*** (0.00359)	0.0114** (0.00442)	0.01619*** (0.00385)	0.02257*** (0.00354)
African American	0.0247*** (0.00401)	0.01851*** (0.00407)	0.0256*** (0.00448)	-0.0248*** (0.00404)	-0.0189*** (0.00418)	-0.0196*** (0.00468)
Hispanic	0.01577*** (0.00254)	0.0137*** (0.00308)	0.01633*** (0.00338)	-0.0160*** (0.00255)	-0.0136*** (0.0032)	-0.0126*** (0.00316)
Pacific Islander or Native American	0.01882** (0.00886)	0.01389** (0.00564)	-0.00168 (0.00547)	-0.02079** (0.00905)	-0.01452** (0.00603)	-0.01851 (0.01195)
Unemployment	-0.00537 (0.02386)	0.01434 (0.02201)	0.02584 (0.01882)	0.00528 (0.02381)	-0.00343 (0.02303)	-0.00533 (0.01758)

Appendix Table F

	2016 Dem	2018 Dem	2020 Dem	2016 Rep	2018 Rep	2020 Rep
Military	0.00386 (0.01322)	-0.00934 (0.00859)	-0.02418** (0.01062)	-0.0038 (0.01341)	0.0089 (0.00879)	0.02278* (0.01166)
District Income	-0.0319* (0.0167)	-0.0393*** (0.0131)	-0.0155 (0.0122)	0.0354** (0.0168)	0.0405*** (0.0139)	0.0152 (0.0123)
No Health Ins.	0.0106 (0.00834)	0.01808** (0.00874)	0.02052** (0.01014)	-0.00993 (0.00841)	-0.0244*** (0.00857)	-0.0323*** (0.01141)
Natural Resources	-0.00218 (0.0094)	-0.00975 (0.00853)	-0.00935 (0.01194)	0.00424 (0.00944)	0.01535* (0.00871)	0.01318 (0.01215)
Manual Labor	0.00217 (0.00684)	0.00042 (0.00502)	0.00762 (0.0062)	-0.00305 (0.00681)	-0.00125 (0.00526)	-0.0044 (0.00661)
Professional Labor	0.01902*** (0.00632)	0.03153*** (0.00597)	0.02749*** (0.00647)	-0.0192*** (0.00637)	-0.0343*** (0.00622)	-0.0269*** (0.00717)
Observations	3,160	3,264	2,559	3,160	3,264	2,559
Pseudo R^2	0.8390	0.7884	0.8295	0.8364	0.7991	0.8137

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Appendix G: Political Party Part Time vs. Full Time Election Victory Probit Panel Marginal Effects

Appendix Table G

Political Party Part Time vs. Full Time Election Victory Probit Panel Marginal Effects

	Dem (Full Panel)	Dem (Part Time Only)	Dem (Full Time Only)	Rep (Full Panel)	Rep (Part Time Only)	Rep (Full Time Only)
Democrat Incumbent	0.5393*** (0.0265)	0.51991*** (0.03331)	0.30617*** (0.09362)	-0.5081*** (0.0262)	-0.5044*** (0.0328)	-0.13187** (0.05296)
Republican Incumbent	-0.4297*** (0.0255)	-0.3396*** (0.02743)	-0.2957*** (0.07487)	0.4508*** (0.026)	0.35738*** (0.02721)	0.14861*** (0.05037)
District Democrats	0.0109*** (0.001)	0.00809*** (0.00073)	0.00563*** (0.00191)	0.0109*** (0.001)	-0.0076*** (0.00072)	-0.00238** (0.00114)
District Independents	0.007*** (0.0022)	0.0058*** (0.00181)	0.00136 (0.00198)	0.007*** (0.0022)	-0.0079*** (0.0016)	-0.00313* (0.00179)
Democrat Contributions	0.00478*** (0.00103)	0.0105*** (0.00183)	0.0014 (0.002)	-0.0052*** (0.00108)	-0.0102*** (0.00179)	-0.00061 (0.00269)
Republican Contributions	-0.0049*** (0.00135)	-0.0083*** (0.00197)	-0.00176 (0.00181)	0.00579*** (0.0016)	0.00825*** (0.0019)	0.00082 (0.00208)
Independent Contributions	-0.043** (0.0214)	-0.0424 (0.0284)	-0.0141 (0.014)	-0.0137* (0.00765)	-0.00961 (0.0178)	-0.00218 (0.00233)
Population (x10,000)	0.0448 (0.0348)	0.0833** (0.0391)	-0.0263 (0.0318)	-0.0635* (0.0339)	-0.0808** (0.0369)	-0.00169 (0.011)
Median Age	0.0212*** (0.0058)	0.02034*** (0.00492)	0.00432 (0.00622)	-0.0261*** (0.0057)	-0.0209*** (0.00488)	-0.00613* (0.00355)
Unmarried	0.0334*** (0.0052)	0.02657*** (0.00421)	0.01498*** (0.00574)	-0.0382*** (0.005)	-0.0266*** (0.00413)	-0.01023** (0.00502)
High School or Less	-0.0178*** (0.0025)	-0.0115*** (0.00275)	-0.0126*** (0.0042)	0.0162*** (0.0026)	0.01002*** (0.00273)	0.00548** (0.00267)
African American	0.0263*** (0.0035)	0.01976*** (0.00277)	0.01594*** (0.00519)	-0.0249*** (0.0035)	-0.0180*** (0.00266)	-0.00591** (0.00259)
Hispanic	0.0166*** (0.0023)	0.01268*** (0.00191)	0.0085** (0.00399)	-0.0156*** (0.0023)	-0.0116*** (0.00191)	-0.00399* (0.00206)
Pacific Islander or Native American	0.0159*** (0.0049)	0.02134*** (0.00483)	0.00379 (0.00302)	-0.0229*** (0.0052)	-0.0205*** (0.00475)	-0.00284 (0.00189)

Appendix Table G

	Dem (Full Panel)	Dem (Part Time Only)	Dem (Full Time Only)	Rep (Full Panel)	Rep (Part Time Only)	Rep (Full Time Only)
Unemployment	0.0003 (0.0136)	0.00387 (0.01678)	0.0095 (0.01069)	0.009 (0.0133)	-0.00112 (0.01631)	0.00115 (0.00366)
Military	-0.0113 (0.0104)	-0.00239 (0.01063)	-0.01271* (0.00771)	0.0091 (0.0106)	0.00121 (0.01028)	0.00289 (0.00282)
District Income	-0.00991 (0.00918)	-0.0108 (0.00865)	-0.00941 (0.0103)	0.00915 (0.00947)	0.0117 (0.00855)	0.00344 (0.00451)
No Health Ins.	0.0235*** (0.0065)	0.02359*** (0.00573)	-0.0039 (0.00798)	-0.0296*** (0.0062)	-0.0252*** (0.0058)	-0.00235 (0.00218)
Natural Resources	-0.0147** (0.0071)	-0.01209* (0.00676)	-0.00426 (0.00548)	0.0166** (0.0071)	0.01471** (0.00672)	0.0023 (0.00237)
Manual Labor	-0.0004 (0.004)	-0.00481 (0.00388)	0.00565 (0.00351)	0.0001 (0.0041)	0.00566 (0.00388)	-0.0032 (0.00201)
Professional Labor	0.0223*** (0.0045)	0.01896*** (0.00401)	0.01115* (0.0061)	-0.0244*** (0.0047)	-0.0201*** (0.00411)	-0.00456 (0.00295)
Observations	8,983	6,418	2,564	8,983	6,418	2,564
Pseudo R^2	0.7996	0.8014	0.8368	0.8006	0.7949	0.8182

Significance Codes: “***” 0.01 “**” 0.05 “*” 0.1

Appendix H: Additional Specifications of Observed Entry Decisions vs. Predicted Entry Configurations

Appendix Table H

Additional Specifications of Observed Entry Decisions vs. Predicted Entry Configurations

Model	Correctly Predicted	Correctly Predicted Democratic Entry	Correctly Predicted Republican Entry	Correctly Predicted Independent Entry
Forgone Income = District Average				
Model 1: Expected Salary	413 (4.5%)	2205 (24.5%)	2496 (27.7%)	7855 (87.5%)
Model 2: Salary, Per Diem, Benefits, Filing Fees	679 (7.5%)	2567 (28.6%)	2824 (31.4%)	7858 (87.5%)
Model 3: Salary, Per Diem, Benefits, Filing Fees, Estimated Campaign Expense	676 (7.5%)	2563 (28.5%)	2774 (30.8%)	7858 (87.4%)
Model 4: Salary, Per Diem, Filing Fees, Linear Expenses	676 (7.5%)	2563 (28.5%)	2774 (30.8%)	7858 (87.5%)
Model 5: Salary, Per Diem, Filing Fees, Logarithmic Expenses	1088 (12.1%)	2671 (29.7%)	4281 (47.7%)	7858 (87.5%)
Forgone Income: Part Time = $0, Full Time = District Average				
Model 1: Expected Salary	1687 (18.9%)	5905 (65.7%)	6250 (69.6%)	4713 (52.5%)
Model 2: Salary, Per Diem, Benefits, Filing Fees	3027 (33.7%)	6140 (68.4%)	6690 (74.4%)	6555 (72.9%)
Model 3: Salary, Per Diem, Benefits, Filing Fees, Estimated Campaign Expense	3367 (37.4%)	5448 (60.7%)	6596 (73.4%)	7877 (87.6%)
Model 4: Salary, Per Diem, Filing Fees, Linear Expenses	3353 (37.3%)	5447 (60.6%)	6595 (73.4%)	7859 (87.5%)
Model 5: Salary, Per Diem, Filing Fees, Logarithmic Expenses	3748 (41.7%)	5771 (64.3%)	7030 (78.3%)	7863 (87.5%)

CHAPTER 5: CONCLUSIONS

The purpose of this dissertation is to explore how modifying the structure of voting systems and procedures can identify and change people's political behavior. By changing the voting rules or election structures faced by individual political agents, we ultimately change the incentive structures that underlie their decision making and should expect changes in political outcomes.

The study examines two types of institutional changes and their associated effects on the behavior of political agents. The first part of this study discusses the design of an alternative voting system that accounts for the intensity of individual voters' preferences towards certain policies, while maintaining reasonable equity constraints. The second part of the study examines the political entry decisions of political party candidates in state assembly elections, and what factors are likely to influence their entry decisions, and ultimately political competition.

Chapter 2 presents the rules of a two-stage multivoting system and discusses how the system can be used to measure intensities of preference and classify types of voting behavior. In an election with multiple policies under consideration, voters are given an endowment of votes that can be freely distributed across any of the policies under consideration. This results in two stages of choices: the first on how to distribute votes across policies (global choice), and the second on how to vote on the alternatives within a policy (local choices). The collective decision for a policy is to select the alternative that received the greatest number of aggregate votes (rather than the one with the greatest number of supporters). This allows a minority group of voters with strong preferences to overcome a relatively indifferent majority. In addition, the quality of information generated by the multivoting system is greater than that of a traditional one-person, one-vote system, as it account for strengths of preference.

The two-stage multivoting system can also be used as a measure of the intensity of voter preferences, and as a way of classifying types of voters. Since certain types of voting behavior correspond with exact distributions, behavior types can be classified by looking at the variation within a voter's policy weighting vector. Thus, voting behaviors (such as single-issue voting) that we expect to exist but cannot be revealed within traditional voting, can be identified by the two-stage multivoting system.

Using the theoretical framework established in the second chapter, Chapter 3 evaluates the performance of the two-stage multivoting system using experimental data gathered from college students. Using the 2020 Democratic primaries as a framework for participants, participants are asked to vote for the Democratic nominee in both a "one-person, one-vote" system and a "two-stage multivoting system". The study finds that the predicted top candidate is consistent between the two systems; however, the two-stage multivoting system produces more unique rankings between candidates, increasing the difference in votes between candidate ranks and breaking ties common between less well-known candidates.

Chapter 4 investigates political competition in state assembly elections using two approaches toward modeling political entry. In the first approach, political party entry decisions are influenced by election, demographic, and economic attributes of election districts. Political party-candidates that "fit" the district's preferences are more likely to enter than others. The second approach models political entry based on the expected payoffs each political party is likely to receive as a result of political party entry. Using these two approaches to modeling political entry, the study estimates the marginal effects that certain district attributes and components of expected payoff have on the probability of entry for three political parties in the United States: Democrats, Republicans, and Independents. Predicted probabilities of entry are tested against observed outcomes in order to evaluate the predictive strength of these two approaches. Estimation results suggest that modeling political entry based on district attributes is the preferred approach, having the greater predictive ability and generalizability than the expected payoff approach.

These two studies broadly examine elections using a public choice approach, emphasizing the importance of institutional factors (voting rules, election characteristics, etc.) in the decisions of political agents, whether they are voters or political candidates. Given improvements in voting technology and increasing exploration of alternative voting systems, two-stage multivoting system is a viable option in elections with multiple policies, crowded fields, or circumstances where a deeper understanding of preference intensities may be valuable. Results from the political entry model provide new insight into political competition in state assembly elections. Future research regarding the effect of changes in ballot access procedures (such as filing fees and signature requirements) on political entry at different election geographies would enrich the findings of this study.

References

Arrow, K. J. (1951). *Social Choice and Individual Values*. New Haven: Yale University Press.

Black, D. (1948). On the Rationale of Group Decision-making. *Journal of Political Economy*, 23-34.

Black, D. (1958). *The Theory of Committees and Elections*. Cambridge: Cambridge University Press.

Buchanan, J. M., & Lee, D. R. (1986). Vote buying in a stylized setting. *Public Choice*, 3-15.

Buchanan, J. M., & Tullock, G. (1999). The *Calculus of Consent*. Indianapolis: Liberty Fund, Inc.

Casella, A. (2005). Storable votes. *Games and Economic Behavior*, 391-419.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 135-150.

Hansen, R. L. (2000). Vote Buying. *California Law Review*, 1323-1372.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 41-57.

Tullock, G. (1959). Problems of Majority Voting. *Journal of Political Economy*, 571-579.

Weyl, E. G. (2013). Quadratic Vote Buying. Working Paper.