

Essays on Political Economy

Diego Canales

School of Economics
University of Nottingham

A thesis submitted for the degree of

Doctor of Philosophy

Nottingham, December 2020

Acknowledgements

First and foremost, I would like to thank my supervisors: Prof. Vincent Anesi and Prof. Daniel Seidmann. Without their truly infinite patience and never-ending support, I would not find myself writing the acknowledgements of this present doctoral dissertation.

I also would like to thank my family and more importantly, my mother, Josefina. She had been the only co-conspirator of my many academic endeavours throughout my life. I shall not forget the many sacrifices you had to make, and I will carry in my heart all your priceless and invaluable teachings. You were the most prominent beacon of light in this almost insurmountable experience.

Additionally, I infinitely thank and acknowledge my partner Yarely. These last thirty months have proven to be the most spirit-lifting in my life. I could not have expected for a better companion every single day. Your presence invigorates my existence and encourages me to become a better person. Nottingham became a different city after I met you. Thank you for sharing your life with me.

Lastly, I substantially appreciate and recognise my colleagues from the Economics department at the University of Nottingham. All of them contributed to my academic and personal growth. I would like to specially commemorate the many endless (but quite nourishing) discussions, along with many cups of coffee, I had the pleasure to have with Vasileios Kotsidis and Despoina Alempaki.

Furthermore, I would always be thankful to Prof. David Vázquez-Guzmán, who I considered my academic foster parent. During my PhD, I experienced several near life-wreaking challenges and I would not have been able to survive without David. He patiently extended his full unconditional support at every single moment.

Contents

Acknowledgements i

List of Figures iv

List of Tables v

1 General Introduction **1**

2 A model of k-winners **5**

2.1 Introduction 6

2.2 Literature Review 7

2.3 Model 13

2.4 Main Results 17

2.4.1 Cases where the median alternative is the k-winner 17

2.4.2 Cases where a non-median alternative is the k-winner 21

2.4.3 Alternatives exterior to N 24

2.5 Examples 27

2.5.1 First Case 27

2.5.2 Second Case 28

2.5.3 Third Case 28

2.5.4 Fourth Case 29

2.6 Concluding remarks 29

2.7 References 31

Appendix A **32**

3 A Model of Sequential Elections: The Case of American Primaries **35**

3.1 Introduction 36

3.2 Literature Review 39

3.2.1 Sequential Elections 39

3.2.2 Momentum 42

3.2.3 Duverger's Law 47

3.2.4 Additional Literature on Primaries 50

3.3 Model 51

3.3.1 List of Conditions 54

3.4 Results ($r = 1$ case) 57

3.4.1 Dynamics 58

3.4.2 Condorcet cycles 61

3.5 Results (r=2 case) 63

3.6 Momentum ($r = 2, n = 3$ case) 65

3.7 Modified Duverger's Law ($r = 2, n = 3$ case) 67

3.8 Spatial Setting 70

3.8.1 Results with $r = 1, n = 3$ 71

3.8.2 Results with $r = 2, n = 3$ 74

3.9 Concluding remarks 76

3.10 References 78

Appendix 1: Numerical Verification **82**

Appendix 2: Primary Election Rules **85**

Appendix 3: Tables **89**

List of Figures

2.1 Graphical example of Conditions (C.0) (Given the space restriction, the axes of this graph are shown uneven/unequal). 18

2.2 Graphical example of Proposition 6. (Given the space restriction, the axes of these graphs are shown uneven/unequal.) 21

2.3 An alternative exterior to N is k-winner for every k. 28

2.4 The 2-winner and the 3-winner both exist but differ. 28

2.5 The 2-winner and the n-winner coincide but no k-winner exists for every $2 < k < n$. 28

2.6 The 2-winner and the n-winner differ but no k-winner exists for every $2 < k < n$. 29

List of Tables

3.1 Distribution of Republican Delegates for the Presidential Primary of 2016 89

3.2 Distribution of Democratic Delegates for the Presidential Primary of 2016 91

3.3 Schedule of Republican Presidential Primaries in 2016 92

3.4 Schedule of Democratic Presidential Primaries in 2016 93

Chapter 1

General Introduction

Typically, political phenomena have been the main focus of study for philosophers and political scientists. However, with the advent of mathematicians (e.g., Hotelling, Debreu, and Arrow) to the Economics discipline, there was an increased inclination to utilise advanced mathematical technique to solve complex problems around political science. This dissertation attempts to follow that tradition. In the next two chapters, we analyse political phenomena using economic tools.

In the second chapter, we start by looking at the concept of a Condorcet winner; which is defined as the alternative that is preferred by a plurality in every pair-wise contest. Then, we put forward a generalisation, the k-winner; which we define as the alternative that is preferred by a plurality in contests consisting of k alternatives. In the literature, Saari (1989) and Meyers et al. (2014) had formulated a notion of generalised Condorcet winners. However, we provide a more general definition of k-winners, and we apply it to a spatial setting where the alternatives are ordered on the unit interval. In that setting, there must be a Condorcet-winner, which is a 2-winner. However, this is not true for a k-winner when $k > 2$. Furthermore, we show that if a k-winner exists for some $k > 2$ then the same alternative must be the k'-winner for any $k < k' \leqslant N$. Moreover, we provide further sufficient and necessary conditions for the existence of a k-winner for some $k > 2$.

There is a natural link between Chapters 2 and 3: primaries. The concept of k-winners found in Chapter 2 can be further extended to analyse the case of primary elections;

more specifically American presidential primaries. We will look thoroughly at this political process which is both unique and complex. In this chapter, we use a non-cooperative model where a continuum of citizens vote, according to their ranked preferences, from a sequence of candidates in a sequential series of primary elections. There are two significant exogenous variables: voting rules and delegation size. In every election, there is a fixed number of delegates given to (i) a single winner in the case of plurality rule and (ii) multiple candidates under a proportional representation rule. Moreover, these candidates take sequential turns to decide if it would be optimal to stay in the race giving there is a cost to run in every primary. Another essential feature of the model is the number of regions or elections; this would enable us to analyse the distinction between simultaneous and sequential elections. To do this, we set-up a benchmark case of a simultaneous election (i.e., only one region) and then we compare the results with the case of sequential elections (i.e., more than one region). More importantly, we use this theoretical framework to study two seemingly unrelated political phenomena: momentum and Duverger's law.

The concept of momentum captures the political wisdom which dictates that "doing well" in the early stages is crucial for any candidate. This wisdom translates into a substantial amount of attention drawn to the Iowa caucus. Iowa is the first electoral contest in the primary schedule. Although it does not award a large number of delegates, the majority of the candidates will focus their initial campaign efforts there. Ultimately, the empirical evidence does not seem to support the existence or the effectiveness of the so-called "momentum effect". In that light, in Subsection 3.6, we use our model to derive conditions under which there is an order of moves in the second region in which the 3-winner (i.e., the front-runner) does not become the nominee even after getting all the delegates from the first region.

The second half of Chapter 3 aims to revise the so-called "Duverger's law" into the context of the primary elections. This sociological "law" was the product of the 20th-century sociologist Maurice Duverger. He noticed a particular pattern across several countries: the voting system (i.e., voting rules) had a particular effect on the number

of parties forming a government. More specifically the law asserts that plurality rules "tend to favour" two-party governments whereas proportional representation rules favour multiple-party governments. In the case of primary elections, the analogous concept will be the number of candidates at a primary contest. Thus the modified Duverger's law version for primary elections would assert that primary candidates tend to quit faster or earlier under plurality rule compared to proportional representation. Our theoretical results will characterise conditions under which this assertion can be challenged. More specifically: we find both sufficient and necessary conditions under which, in equilibrium, the number of candidates remaining in the first stages of a sequential primary election, would be strictly smaller under proportional representation compared to plurality rule.

References

Meyers, A., Orrison, M. E., Townsend, J., Wolff, S., & Wu, A. (2014). Generalized Condorcet winners. *Social Choice and Welfare*, *43*(1), 11–27. doi:https://doi.org/10.1007/s00355-013-0765-8

Saari, D. G. (1989). A Dictionary for Voting Paradoxes. *Journal of Economic Theory*, *48*(2), 443–475. doi:https://doi.org/10.1016/0022-0531(89)90037-9

Chapter 2

A model of k-winners

Abstract

The concept of Condorcet winner has become central to most electoral models in the Political Economy literature. A Condorcet winner is defined as the alternative that is preferred by a plurality in every pairwise competition. The notion of a $k-$winner generalises that of a Condorcet winner. The $k-$winner is the unique alternative that is top-ranked by the plurality in every competition composed exactly of k alternatives, including itself. We use a spatial voting setting to characterise this theoretical concept. More importantly, we show that if a k-winner exists for some $k > 2$ then the same alternative must be the k'-winner for every $k < k' \leqslant n$. Finally, we derive additional results such as sufficient and necessary conditions for the existence of a k-winner for some $k > 2$.

JEL codes: D71, D72

Keywords: elections, k-winners, Condorcet winner, spatial voting model

2.1 Introduction

The design of optimal political mechanisms has been a central problem for politicians and philosophers since the dawn of western civilisation. More specifically, the optimal design of electoral procedures such that aggregated decisions could closely capture the preferences of individuals.

The earliest forms of democratic expressions are found in Athens. Initially, voting was done in public at the assembly with two simple choices: accept or reject the measure at hand. Eventually, Athenians used ballots in order to count more precisely the number of votes (see Blackwell (2003)).

The Romans inherited this democratic tradition from the Greeks. However, the Roman Republic cannot be considered a fully democratic apparatus. According to Vishnia (2012): "(Roman) voters did not choose between one political platform and another, but between candidates who were all drawn from a select group within the citizenry which held the exclusive right to compete for the various positions in line with a set of rules established over the years by both tradition and legislation"

Up to McLean (1990) the consensus in the literature was that Condorcet was the earliest contribution to voting paradoxes, specifically the case with intransitive social preferences. McLean, however, unearthed two earlier academic contributions dating back to the Middle Ages: (i) Ramon Lull (born in 1235 and died in 1315), and (ii) Nicholas of Cusa (born in 1401 and died in 1464). The former was a Catholic missionary from Mallorca (now Palma) who would travel extensively to Southern Spain and Northern Africa with the mission to convert Muslims. His results constitute a series of elementary results around voting processes. Lull (1299) in his "De Arte Eleccionis" he proposed an electoral procedure to determine office-holders within the ranks of the Catholic Church. Lull suggested looking at the winner of every pair of alternatives.

The outline of this paper is the following: in Section 2.2 we will discuss a literature review around the concept of Condorcet winners. Next, in section 2.3 we introduce our model and we also provide a complete characterisation of k-winners in a spatial model.

Next, Section 2.4 contains the main results of this paper split into two subsections: (i) in Subsection 2.4.1, we present sufficient and necessary conditions for the median alternative to be a k-winner for every $k > 2$ and (ii) in Subsection 2.4.2 and 2.4.3, we extend the previous results to non-median alternatives and alternatives exterior to N, respectively. In Section 2.5, we present four numerical examples which illustrate the most relevant results. Finally, in Section 2.6 we present our concluding remarks; and in Appendix A, we put forward an alternative formulation for the preliminaries of the model without the assumption of uniformity.

2.2 Literature Review

The first and most cardinal discussion surrounding electoral phenomena was formalised by Marie Jean Antoine Nicolas de Caritat, Marquis of Condorcet. A product of the French Enlightenment, Condorcet was a pioneer of applying formal mathematical methods to the "moral" sciences (i.e., social sciences). In 1785, the Marquis of Condorcet published his manuscript on voting procedures: "Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix". Because of both its relevance and the complete absence of a full English translation, Condorcet (1785) can be considered paradoxically one of the most referenced and the least read manuscript in the literature of Social Choice.

Condorcet's central contribution to voting theory is the concept of a Condorcet winner, which is defined as an alternative with the plurality of votes in a pairwise comparison against every other alternative. In this paper, we present a generalisation of the Condorcet winner: k-winners. A k-winner is an alternative with the plurality of the votes in every k-wise comparison against every other alternative.

Condorcet started looking at the fundamental problem faced by a jury: making a correct decision collectively. His main arguments were based on the early notions of computation of binomial probabilities (given there are two choices: truth and mistake in the original terms of the *Essai*). His intuition is now formalised as Condorcet's Jury Theorem.

This result would constitute a gateway to a more general problem in the social sciences: voting. Condorcet's objective was to devise an optimal mathematical procedure to ag-

gregate individual voting choices to select the optimal choice for the electorate.

The idea was that each voter strictly ranks all of the alternatives. The "winner" in a contest between some alternatives is the alternative which is top-ranked in the contest by most voters. Condorcet considered a population of these alternatives and contests between every pair of alternatives. The Marquis constructed the well-known "Condorcet criterion" by looking at the alternative with the plurality of the votes in every head-to-head comparison. According to the Marquis, this plurality alternative was the "best" choice for society and ought to be elected. This will become the definition of a "Condorcet winner."

However, the Marquis discovered specific voting scenarios where voting yielded what he called "inconsistent propositions": that is, scenarios where the criterion could not compute a plurality winner. Moreover, even worse, as the number of alternatives increased, the number of inconsistent propositions increased quite dramatically. Condorcet had just discovered voting paradoxes. However, his methods were not entirely welcome for a considerable amount of time; according to Black (1958) the German mathematician Eugen Netto was quite critical of Condorcet's approaches:

> "... death prevented that man of genius [Daniel Bernoulli] from applying the theory of probability to civil, moral and economic conditions, as he had intended. Condorcet is said to have carried through Bernoulli's intentions after seventy-two years. Other critics consider Condorcet fantastic rather than rigorous, enthusiastic rather than scientifically exact."

Black in his 1958 book "The Theory of Committees and Elections" reintroduced, translated and reinterpreted the central results from the Condorcet's *Essai*. Then, Black provided a remarkable workaround to Condorcet's paradox by restricting the class of voters' preferences and the population of alternatives. Black restricted the preferences of voters to have one single ideal alternative across a uni-dimensional space. This restriction will be the crucial element in the now-famous Median Voter Theorem. As the alternatives are scalar, they can be ordered (i.e., $a_1 < \ldots < a_n$.) We can therefore order the voters by

the alternatives which they top-rank. Define the median as the voter who top-ranks an alternative (say m) such that at least half of the voters top-rank some alternative that does not exceed m and half of the voters top-rank an alternative that weakly exceeds m.

Theorem 1 (Black (1958))**.** *When there are n odd voters, all of whose utility functions are single-peaked on the real line, the ideal alternative of the median voter gets more votes in a contest against every other alternative, and it is the only alternative which can do so.*

Later on, Newing and Black (1951) explored the (non)existence of a Condorcet winner where the policy space is two-dimensional. This sequence of major breakthroughs was fundamental in the development of Social Choice Theory. Moreover, Black also explored the ideas and results of another forgotten mathematician: Charles Lutwidge Dodgson or more commonly known by his pen name Lewis Carroll. In a paper titled "A method of taking votes on more than two issues", Dodgson (1876) put forward an alternative method to determine a winner whenever the Condorcet winner does not exist. This procedure has been known as the Dodgson criterion, and its main attractiveness is that it always select a winner: the Dodgson Winner. Firstly, if there is a Condorcet winner she is also the Dodgson Winner. However, if there is a Condorcet cycle we will look for a alternative that could become the Condorcet winner if the least number of voters could switch their ranking preference in her favour. The next major breakthrough in Social Choice Theory was provided by Arrow (1951): Arrow's impossibility theorem. This result establishes that no aggregation method (i.e., social welfare function) can aggregate individual ranked preferences without violating at least one of the now canonical Arrovian criteria: (i) Unrestricted Domain, (ii) Pareto Principle, (iii) Independence of Irrelevant Alternatives, and (iv) Non-dictatorship. Arrow's negative result provides evidence that the Condorcet paradox is not a singularity in voting methods but is instead a symptom of a more consistent deficiency of all aggregation methods.

After the publication of Arrow's book in 1951, there was a major breakthrough that was published independently by two researchers: Gibbard (1973) and Satterthwaite (1975).

These two results are now jointly called "Gibbard-Satterthwaite theorem" and it constitutes a generalisation of Arrow's theorem by instead considering the more general case of deterministic ordinal voting.

Given the general conditions and the extensiveness of these two major theorems (Arrow's and Gibbard-Satterthwaite's) most of the literature concentrated on (i) alternative approaches to prove the theorems (see Fishburn (1970), Blau (1972), Barberá (1983), Blackorby et al. (1990), Geanakoplos (2005), Reny (2001), Feldman and Serrano (2008), and Sen (2017)) and (ii) according to Maskin's foreword to Arrow's 1951 book "Much of the literature stemming from Social Choice amounts to trying to avoid the Impossibility theorem by relaxing one or more of the Arrow conditions". Indeed there is a vast literature in this direction (see Blau (1971), Mas-Colell and Sonnenschein (1972), and Wilson (1972)).

However, it was a mathematician who turned this complex aggregation problem inside out; Don Saari in a series of papers and books reformulated the voting paradoxes using a simpler mathematical toolbox: geometry. A Purdue graduate, Saari (1976) completed his doctoral degree with a dissertation titled "Singularities of the n-Body Problem of Celestial Mechanics". After decades working on Celestial Mechanics, Saari started to become interested with mathematical applications to Economics. Firstly, he was interested in price mechanisms in the general equilibrium model context. He would later argue that Social Choice and Celestial Mechanics faced similar dilemmas on how to resolve seemingly complex problems. Saari (2016) preaches a rather reductionist approach composed of the following steps: "(i) Divide a complex problem into tractable components, (ii) Analyse each part, and (iii) Assemble the answers to discover resolutions."

Following his research programme, Saari (1989) provided a theoretical framework where elections outcomes could be categorised (and ordered) into a "dictionary". His approach starts with a general setting for positional voting methods. Firstly a finite number of voters rank (i.e., order) the n alternatives available. The distribution of votes will then be captured by the voting vector $W = (w_1, \ldots, w_n)$; where each weight w_j is the number of voters who rank some alternative (say, a) in the j^{th} position when faced with the set

of alternatives K which contains a. Using this notation we have that Plurality voting and Borda's Count can respectively be identified with the voting vectors $(1, \ldots, 0)$ and $(n-1, n-2, \ldots, 1, 0)$. Now, he labels each of the non-empty subsets[i] (whose cardinality is at least two) as $S_1, \ldots, S_{2^n-(n+1)}$. Next, for each j of these subsets assign a voting vector W_j to count the weights of the rankings. Let $W^n = (W_1 \ldots, W_{2^n-(n+1)})$ contain all the possible voting vectors. Now, Saari constructed a "space of all possible election outcomes": for every subset S_j there exists a listing R_j which contains "the listing of all possible rankings of the alternatives that are generated by complete, transitive, symmetric binary relationships" on S_j. Then, we have two important objects: a word and a profile. A word contains a list of $2^n - (n+1)$ rankings, and each word is an element of the set U^n, called the universal set, which is defined as the Cartesian product of every listing. A profile contains the rankings of voters over the n alternatives and P^n is the space of all possible profiles. Then, we arrive at a definition of a "word" from Saari's dictionary: "a list of election rankings that is attained with a profile of voters".

With this taxonomical description of voting systems, Saari then focuses on the distinction where paradoxes might exist. For that, he uses the concept of "binarily consistent" defined as the case when "... the ranking for each subset of candidates is generated by the same complete, transitive, binary relationship." Therefore, any intransitivity such as the Condorcet paradox case will be classified as "binarily inconsistent". Finally, one of the main results found in Saari (1989) is that the set of "binarily inconsistent" words is a proper subset of the dictionary (i.e., U^n) when using the Borda Count. That is, according to Saari "only Borda avoids many paradoxes".

A related notion of k-winners is found in Meyers et al. (2014). The authors follow Saari's (1989) framework of the "dictionary" of voting outcomes, and furthermore characterise "Condorcet words". The authors use this refinement of Saari's word to contextualise their concept of k-winners.

Meyers et al. (2014) proposed a model with n alternatives and label the subset of k alternatives as K. Given a voting vector $W = (w_1, \ldots, w_n)$, the score of any given

[i] For n alternatives there are $2^n - (n+1)$ of these subsets

alternative is defined as the weighted sum of the voters who rank a in each position. Then the authors say that a dominates K when a has the highest score of the alternatives in K for every voting vector and if, for any other alternative (say, a') in K, there is a voting vector such that a has a strictly higher score than a'. Finally, a k-winner is defined as the alternative that "dominates every subset of k alternatives that contains her". Our definition of k-winner is much less restrictive than the definition found in Meyers et al. because we look at plurality voting vectors (i.e., $(1, 0, \ldots, 0)$).

A Condorcet word in Meyers et al. (2014) is shorter than a word from Saari's dictionary because the former "only records which candidates are the generalised Condorcet winners". Then, an $n-1$ vector w is a Condorcet word if there exists a voting vector such that for every $2 \leqslant k \leqslant n-1$ the k^{th} element identifies the $(k+1)$-winner attained by the profile P or a blank. In the case where the profile P does not have a $(k+1)$-winner then the k^{th} letter will be a blank.

For example, if there are four alternatives then all Condorcet words are 3-element vectors: the first element identifies the 2-winner, the second element identifies the 3-winner, and the third element identifies the 4-winner. Using this notation, for example, the word $(a_1, a_1, *)$ is a Condorcet word because it describes a voting outcome where alternative a_1 is the 2-winner and the 3-winner but no alternative is the 4-winner.

The main contribution in Meyers et al. (2014) is the characterisation of the set of Condorcet words; in particular, they show that some words are not Condorcet words. Then, their main characterisations on existence of k-winners are highlighted in the following results: (i) there are Condorcet words whose first elements are blanks up to the ℓ-th position, and whose remaining elements are the same alternative; (ii) there are Condorcet words whose first elements are the same alternative up to the ℓ-th position, and whose remaining elements are blanks; and (iii) there are Condorcet words whose elements are the same alternative (say a) up to the ℓ-th position, and whose remaining elements are a different alternative (say a').

The present chapter differs in three critical respects: (i) the main contribution is the sufficient and necessary conditions for the existence of a k-winner for every $k > 2$, (ii)

the definition of a k-winner in this chapter is much less restrictive, and (iii) we restrict the set of alternatives to be ordered on the unit interval. In particular, we fix a spatial profile and ask which if any alternatives is a k-winner (in this work's sense). To illustrate the difference in concepts: there must always be an N-winner in our sense, irrespective of the spatial setting; but, as noted above, there may be no N-winner in their sense.

2.3 Model

Call $a \in [0,1]$: an *alternative*. There is a continuum of voters, indexed by j, whose preferences over alternatives is represented by the loss function $-(a-a^j)^2$; a^j is voter j's *ideal alternative in* $[0,1]$.

Distribution of ideal alternatives

In order to simply the exposition of our main results, we assume that a^j is uniformly distributed on $[0,1]$. The *median* voter is then the voter whose ideal alternative on $[0,1]$ is $1/2$. The reader must note that every result up to (excluding) Section 2.4 does not rely on this assumption of uniformity. Moreover, every result in Section 2.4 does rely on this assumption but only to simplify the exposition of results.

In Appendix A, we put forward an alternative formulation for the preliminaries of the model without the assumption of uniformity. We only assume the distribution is continuous and has a unique median. Next, we reformulate Proposition 10 and the necessary and sufficient conditions for an a_m to be a k-winner for any given k.

Preliminaries

We focus on generic[ii] collections of $n > 1$ (finite) alternatives, denoted N, with representative element a_i. Order these alternatives such that

$$0 \leq a_1 < ... < a_n \leq 1.$$

For any collection $K \subseteq N$, we say that the lowest and highest alternatives in K are *exterior to* K, writing the former as $a_{\underline{K}}$ and the latter as $a_{\overline{K}}$; and we say that any other alternatives are *interior to* K.

[ii] We define generic collections as collections with the property that if a_i and $a_{i'}$ are distict elments of N then $a_i + a_{i'} \neq 1$.

Each voter's loss function determines an ordering over the alternatives in any subset $K \subseteq N$. We will denote $a^j(K)$ as the voter j's preferred alternative on K. We will refer to $a^{1/2}(N)$, the median voter's preferred alternative in N, as the *median alternative*. It will prove convenient to write $a^{1/2}(N)$ as a_m. Take any nonempty subset of N, say K. For any $a \in K$, write $\theta(a; K)$ as the proportion of voters j such that $a = a^j(K)$.

The next assumption will be useful to remove the possibility of ties in any contest.

Assumption 2.1. *For every subset $K \in N$, no pair of alternatives can have the same vote share:*

$$\forall K : \theta(a; K) = \theta(a_{i'}; K) \iff a = a_{i'}$$

Now, before putting forward our definition of k-winners we need to firstly define which alternative wins on any given subset K. We say that a is a *winner on* K if

$$a \in \arg\max_{\alpha \in K} \theta(\alpha; K),$$

and that a is a *loser on* K if

$$a \in \arg\min_{\alpha \in K} \theta(\alpha; K).$$

We now expand this notion to every collection of alternatives with k elements.

Definition 1. *We say that alternative a is a k-winner if alternative a is a winner on K for every nonempty $K \subseteq N$ with k elements; we define a k-loser analogously.*

A 2-winner is therefore a Condorcet winner, which coincides with $a^{1/2}(K)$, the median's preferred alternative for every pair of alternatives K.

One of the inherent features of the previous definition is that if a k-winner exists then it must be unique, except for the trivial case of $k = 1$. That is, if a_i were the only alternative available then a_i must win. This is clearly true for every alternative.

Proposition 1 (Uniqueness). *For every $k \in \{2, ..., n\}$, there is at most one k-winner.*

Proof. Fix any k and suppose by contradiction that there are two or more k-winners, say a_i and $a_{i'}$. As N is generic, median voter m strictly prefers one of the pair, say a_i.

If $a_i < a_{i'}$ then there is a voter $j' > 1/2$ such that every voter $j < j'$ strictly prefers a_i over $a_{i'}$; and if $a_i > a_{i'}$ then there is a voter $j' < 1/2$ such that every voter $j > j'$ strictly prefers a_i over $a_{i'}$. Now consider any collection of k alternatives (say, K) which contains a_i and $a_{i'}$. As $a_{i'}$ is a k-winner, it must be a winner on K. However, this is impossible because more than half of voters strictly prefer a_i over $a_{i'}$. Consequently, a_i and $a_{i'}$ cannot both be k-winners. □

In the next proposition, we provide a set of sufficient conditions for the existence of the k-winner for every $k \geqslant 1$. Essentially, the alternative with the majority of the votes (i.e. more than one-half of the total number of votes) will be the strongest alternative to compete with (See the first example in Section 7.1). Furthermore, Proposition 2 exploits the related fact that any voter top-ranks a in a contest K must also top-rank a in every contest $K' \subset K$ such that $a \in K'$.

Proposition 2. *If alternative a has $\theta(a; N) > \frac{1}{2}$ then it is the k-winner for every $k \geqslant 1$.*

Proof. Regardless of any combination of alternatives that would run against alternative a_i, she would have the plurality of the votes, thus she would win every k-wise competition. □

Now, we present the first set of main results. The next two propositions explore the relationship between a k-winner (if it exists) and a $(k+1)$-winner. Firstly, Proposition 3 establishes one sufficient condition for the 2-winner to be not only the 3-winner but also the k-winner for every $k \geqslant 2$. Essentially, the premise of the next proposition is to link the distinction made between alternatives exterior to N and alternatives interior to N, and the notion of a k-winner for every $k \geqslant 2$ in the context of a spatial setting.

Proposition 3. *If $a \in \{a_1, a_n\}$ is the 2-winner then it is the k-winner for every $k \geqslant 2$.*

Proof. Any 2-winner is top-ranked by the median voter. If $a = a_1$ then there is $j^{'} \geqslant 1/2$ such that every voter $j < j^{'}$ strictly prefers a over every $a_i > a_1$. Hence, $\theta(a_1; K) \geqslant 1/2$ for every collection K which contains a_1; and a_1 must therefore be the k-winner. An equivalent argument applies if $a = a_n$ □

Next, we will show a quite striking result: in our spatial context there exists an inherent relationship between the k-winner and the $(k+1)$-winner for $k > 2$. The relevance of this result becomes apparent as soon as we consider its implications: if we can identify a k-winner for any $2 < k < n$ then we know surely the existence and identity of the k'-winner (i.e., for any $k < k' \leqslant n$). The reader must note that the argument for Proposition 4 does not rely on the supposition that voters' ideal alternatives are uniformly distributed.

Proposition 4. *If alternative a is the k-winner for some $3 \leqslant k < n$ then alternative a is also the k'-winner for every $k < k' \leqslant n$.*

Proof. Start by supposing that alternative a is the k-winner for some $3 \leqslant k < n$. It is sufficient to show that a is the $k+1$-winner.

If a is not the $k+1$-winner then some $a' \neq a$ is the winner on some collection, say K' of $k+1$ alternatives which includes a. In this case, we can select some $a'' \neq a'$ in K' such that a' is the winner on $K'' \equiv K' \backslash \{a''\}$: where

$$a'' \in \begin{cases} \{a_{\underline{K'}}, a_{\overline{K'}}\} & \text{if } a \text{ is interior to } K' \\ \{a_{\overline{K'}-1}, a_{\overline{K'}}\} & \text{if } a = a_{\underline{K'}} \\ \{a_{\underline{K'}}, a_{\underline{K'}+1}\} & \text{if } a = a_{\overline{K'}} \end{cases}$$

As $|K"| = k$, a can therefore not be the k-winner, contradicting the premise. □

We can compare the result in Proposition 4 with a sequence of propositions from Meyers et al. (2014) in which they prove the existence of the following three results:

1. There exists a "profile" [iii] where there is no k-winner when $2 \leqslant k \leqslant k'$ but there is a k-winner for any $k' < k \leqslant n$.

2. There exists a "profile" where there is no k-winner for any $k' < k \leqslant n$ but there is a k-winner when $2 \leqslant k \leqslant k'$.

[iii] The formal statement is the following: "...there exist a profile whose associated Condorcet Word..." (see Meyers et al (2014, p.19))

3. There exists a "profile" where an alternative (say a) is a k-winner for any $2 \leqslant k \leqslant k'$ and a different alternative (say a') is the k-winner when $k' < k \leqslant n$.

In their model, they have constructed profiles where k-winners come in "blocks". That is, there is an $2 \leqslant k' \leqslant n$ for which an alternative is the k-winner for any: (i) $k > k'$ or (ii) $k < k'$. In our model, $k' = 3$. That is, if an alternative is the 3-winner it is also the k-winner for every $3 \leqslant k \leqslant n$. However, this alternative might not be the 2-winner as well.

In our model, Propositions 3 and 4 imply that the 2-winner is the k-winner for every $k > 2$ provided she is either (i) exterior to N or (ii) also the 3-winner. Propositions 6 through 11 below will provide necessary and sufficient conditions for existence (and non-existence) of a k-winner.

2.4 Main Results

This section will be split in three subsections. The first one is exclusively focused on conditions under which the median alternative (interior to N) is the k-winner for every k. In the second subsection, we present more general results where any non-median (and interior to N) alternative $a_i \in K$ is the k-winner. In the last subsection, we look at the special cases of the alternatives exterior to N.

2.4.1 Cases where the median alternative is the k-winner

Consider the following condition,

$$\exists a_i \in N : \theta(a_i; N) > \max\left\{\sum_{j<i} \theta(a_j; N), \sum_{j>i} \theta(a_j; N)\right\} \tag{C.0}$$

In the next proposition we present a sufficient (but not necessary) condition under which the median alternative (i.e., a_m) is the k-winner for all $k > 1$.

Condition (C.0) essentially describes a collection of alternatives where there exists a specific alternative with the largest vote share compared to the sum of vote share of all the alternatives to either side.

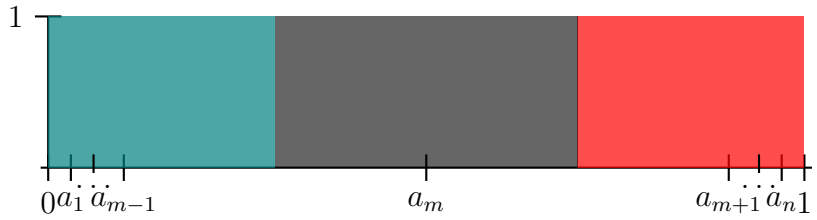


Figure 2.1: Graphical example of Conditions (C.0) (Given the space restriction, the axes of this graph are shown uneven/unequal).

Proposition 5. *If* (C.0) *holds then median alternative* a_m *is the* k*-winner for all* $k > 1$.

Proof. Fix a k, and fix a $K \subseteq N$ containing a_m. Then (C.0) implies that alternative a_m not only has the largest vote share (whenever $K = N$), but she will beat any combination of $k-1$ alternatives, from either her left or right side. □

Propositions 6 through 9 below will provide necessary and sufficient conditions for existence (and non-existence) of a k-winner for four specific cases:

Case #	Condition	Used in:
Case 1	$k \leqslant \min\{m, n-m\}$	Proposition 6
Case 2	$m \leqslant \min\{k, n-k\}$	Proposition 7
Case 3	$m \geqslant \max\{k, n-k\}$	Proposition 8
Case 4	$k \geqslant \max\{m, n-m\}$	Proposition 9

In each case, for each proposition we have a (necessary) condition on both k and m. This set of inequalities defines different thresholds for either: (i) the cardinality of K (i.e, k) or (ii) the ordinal position of the median alternative (i.e, m).

In the first case, in order to prove Proposition 6 we need Lemmata 1 and 2. In conjunction, it provides necessary conditions under which a median alternative (i.e, a_m) is the k winner for every $k > 2$.

Lemma 1. *If* $k \leq m$ *then no alternative* a_i *for* $a_i < a_m$ *is a* k*-winner.*

Proof. Consider any collection K which contains a_m and $k-1$ other alternatives including a_i, all of which are less than a_m. Then, a_m is the winner on K; so a_i cannot be a k-winner. □

The argument in Lemma 1 is the following: as long as k is at most equal to the number of alternatives to the left of a_m (including m) then we can find a subset K composed of exactly these k alternatives. Then, the median alternative has the largest proportion of voters (to the left). For Lemma 2, the argument is analogous. Moreover, the reader should note that the number $n-m$ represent how many alternatives (including m) exist to the right of the median alternative.

Lemma 2. *If $k \leqslant n-m$ then no alternative a_i for $a_i > a_m$ is a k-winner.*

Proof. Consider any collection K which contains a_m and $k-1$ other alternatives including a_i, all of which are greater than a_m. Then, a_m is the winner on K; so a_i cannot be a k-winner □

Next, we put forward sufficient and necessary conditions under which the median alternative is the k-winner for any given k.

Proposition 6. *Let $k \leqslant \min\{m, n-m\}$. For a given $k > 2$, the (interior to N) median alternative a_m is the k-winner if and only if*

$$\tfrac{1}{2}(a_m + a_{m-1}) < \tfrac{1}{2}(a_{m+1} - a_{m-1}) \text{ and}$$
$$1 - \tfrac{1}{2}(a_{m+1} + a_m) < \tfrac{1}{2}(a_{m+1} - a_{m-1})$$

Otherwise, there is no k-winner.

Proof. If $k \leqslant \min\{m, n-m\}$ then no alternative a_i different than a_m can be a k-winner as established in Lemmata 1 and 2. Fix any $K \in N$ such that $\{a_{m-1}, a_m, a_{m+1}\} \in K$. Now, we need to construct the most damaging collection for alternative a_m, which would minimise a_m's vote. This collection would necessarily include her neighbours a_{m-1} and

a_{m+1}, yielding alternative a_m a share of $\theta(a_m; K) = \frac{1}{2}(a_{m+1} - a_{m-1})$. Now, suppose the following two conditions hold,

$$\frac{1}{2}(a_m + a_{m-1}) > \frac{1}{2}(a_{m+1} - a_{m-1}) \text{ and}$$
$$1 - \frac{1}{2}(a_{m+1} + a_m) > \frac{1}{2}(a_{m+1} - a_{m-1})$$

Take the first inequality, which implies that the proportion of voters that top-rank alternative a_{m-1} is strictly greater than the proportion of voters that top-rank alternative a_m, if the subset of alternatives were $K = \{a_{m-1}, a_m, \ldots\}$. Then clearly alternative a_m cannot be a k-winner. Nonetheless since $k \leqslant m \leqslant n - k$ also holds, then there cannot be a k-winner, since the only potential alternative to be a k-winner was alternative a_m. Taking the second inequality, we have the exact same argument as before but for alternative a_{m+1} and it would also imply that there cannot be a k-winner (given the fact that alternative a_m fails to be a k-winner).

If neither of these conditions hold then a_m is the k-winner. To show this, take any K which includes a_{m-1}, a_m and a_{m+1} and any $a_j \in K$ such that $j \leqslant m + 1$. Removing alternatives a_j for any $j < m - 1$ and replacing them with some alternatives a_j for any $j > m + 1$, increases the vote share of some remaining alternatives a_j for any $j < m - 1$ without increasing the vote share for alternative a_j for any $j > m$. However, the increased vote shares of a_j for any $j < m$ are bounded above by $\frac{1}{2}(a_m - a_{m-1})$; and alternative a_m then wins against a_{m-1}. The same argument applies to removing any alternative a_j for any $j < m - 1$.

Finally, removing alternative a_{m-1} increases a_m's vote share and cannot increase the vote share of alternatives a_j for any $j < m - 1$ above $\frac{1}{2}(a_m - a_{m-1})$ which, by supposition is less than $\frac{1}{2}(a_{m+1} - a_{m-1})$, which itself is less than a_m's new vote share. An equivalent argument precludes eliminating alternative a_{m+1}. □

Now, let us look more closely at the conditions in the last proposition. If both of these conditions hold then alternative a_m is the k-winner for a given k. For further intuition, Figure 2.2 shows the graphical interpretation of the following set of conditions:

$$\tfrac{1}{2}(a_m + a_{m-1}) < \tfrac{1}{2}(a_{m+1} - a_{m-1}) \text{ and}$$

$$1 - \tfrac{1}{2}(a_{m+1} + a_m) < \tfrac{1}{2}(a_{m+1} - a_{m-1})$$

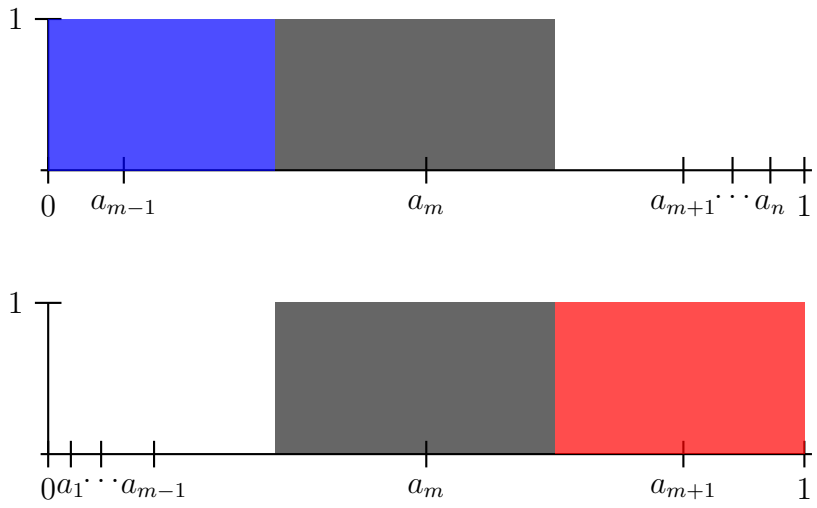


Figure 2.2: Graphical example of Proposition 6. (Given the space restriction, the axes of these graphs are shown uneven/unequal.)

Additionally, by Proposition 4, if the median alternative is a k-winner for a given k then it is also the k-winner for every $k > 2$. Therefore, the same conditions in Proposition 6 are sufficient and necessary for the median alternative, a_m, to be a k-winner for every $k > 2$. The next corollary formalises this result.

Corollary 1. *Let $k \leqslant \min\{m, n-m\}$. The median alternative a_m (which is interior to N) is the k-winner for every $k > 2$ if and only if $\frac{1}{2}(a_m + a_{m-1}) < \frac{1}{2}(a_{m+1} - a_{m-1})$ and $1 - \frac{1}{2}(a_{m+1} + a_m) < \frac{1}{2}(a_{m+1} - a_{m-1})$. Otherwise, there is no k-winner.*

2.4.2 Cases where a non-median alternative is the k-winner

In contrast with the previous subsection, we now turn our focus on results involving non-median alternatives. The next proposition involves the following sufficient condition: $m \leqslant \min\{k, n-k\}$. That is, both (i) the size of the subset K, and (ii) the number of alternatives excluded from K are weakly larger than the ordinal position of the median alternative.

Proposition 7. *Let $m \leqslant \min\{k, n-k\}$. For any $a_1 < a_i \leqslant a_m$ and for any given k alternative a_i is the k-winner if and only if*

$$a_{i-1} \geqslant \tfrac{1}{2}(a_{i+1} - a_i) \text{ and } a_{i+1} - a_{i-1} > \max\{a_{n-k+i+1} - a_i, 2 - a_n - a_{k-1}\}.$$

Otherwise, there is no k-winner.

Proof. Firstly, no alternative $a_i > a_m$ can be the k-winner by Lemma 2. Now take any $a_i \leqslant a_m$. Just like before, we need to construct the most damaging collection of alternatives for alternative a_i. There are three possible contenders to beat a_i: a_{i-1}, a_{i+1} and a_n. In each case, the contender has the best chance of beating a_i when K contains a_{i-1} and a_{i+1} (by minimising a_i's vote share): in which case, a_i's share is $\frac{1}{2}(a_{i+1} - a_{i-1})$. Alternative a_{i-1}'s vote share is maximised when K contains no alternative a_k for any $j < i-1$; and a_{i-1} then beats a_i if and only if

$$\tfrac{1}{2}(a_i + a_{i-1}) \geqslant \tfrac{1}{2}(a_{i+1} - a_{i-1}) \implies a_{i-1} \geqslant \tfrac{1}{2}(a_{i+1} - a_i)$$

Alternative a_{i+1}'s vote share is maximised when $K = \{a_1, \ldots, a_{i+1}, a_{n-k+i+1}, \ldots, a_n\}$; and a_{i+1} then beats a_i if and only if $a_{n-k+i+1} - a_i > a_{i+1} - a_{i-1}$. Finally, a_n's vote share is maximised when $K = \{a_1, \ldots, a_{k-1}, a_n\}$; and a_n then beats a_i if and only if $2 - a_n - a_{k-1} > a_{i+1} - a_{i-1}$. Hence, a_i cannot be the k-winner if one of these conditions holds.

Conversely, no alternative $a_j < a_{i-1}$ can beat a_i if a_{i-1} cannot beat a_i, since alternative a_i's vote share is greater than the maximum share for alternative a_{i-1}, which is the upper bound for any alternative a_j for any $j < i-1$. Similarly, no alternative a_j for any $j > i$ other than a_{i+1} and a_n can beat a_i if neither a_{i+1} nor a_n can beat a_i. □

In contrast with the previous proposition, our next result deals with the following condition: $m \geqslant \max\{k, n-k\}$. That is, both (i) the size of the subset K, and (ii) the number of alternatives excluded from K are weakly smaller than the ordinal position of the median alternative.

Proposition 8. *Let $m \geqslant \max\{k, n-k\}$. For any $a_m \leqslant a_i \leqslant a_n$ and for any given k, alternative a_i is the k-winner if and only if*

$$\tfrac{1}{2}(a_i - a_{i-1}) > 1 - a_{i+1}, \text{ and } a_{i+1} - a_{i-1} > \max\{a_{i-1} - a_{i-n+k-2}, a_1 + a_{n-k+2}\}.$$

Otherwise, there is no k-winner.

Proof. No alternative $a_i < a_m$ can be the k-winner, which is implied by Lemma 1. Now take any $a_i \geqslant a_m$. The most damaging collection of contending alternatives for a_i must include the following contenders: a_1, a_{i-1}, and a_{i+1}. Now, by including a_{i-1} and a_{i+1} in K, we maximise the contenders vote share as well as minimise a_i's vote share. Then a_i's share would be $\frac{1}{2}(a_{i+1} - a_{i-1})$. Alternative a_{i+1}'s vote share is maximised when K contains no alternative $a_j > a_{i+1}$; and a_{i+1} then beats a_i if and only if

$$1 - \tfrac{1}{2}(a_{i+1} - a_i) \geqslant \tfrac{1}{2}(a_{i+1} - a_{i-1}) \implies 1 - a_{i+1} \geqslant \tfrac{1}{2}(a_i - a_{i-1}).$$

Alternative a_{i-1}'s vote share is maximised when $K = \{a_1, \ldots, a_{i-n+k-2}, a_{i-1}, a_i, \ldots\}$; and alternative a_{i-1} then beats alternative a_i if and only if $a_{i-1} - a_{i-n+k-2} > a_{i+1} - a_{i-1}$. Finally, alternative a_1's vote share is maximised when $K = \{a_1, a_{n-k+2}, \ldots, a_{i-1}, \ldots\}$; and a_1 then beats a_i if and only if $a_1 + a_{n-k+2} > a_{i+1} - a_{i-1}$. Hence, a_i cannot be the k-winner if one of these conditions holds.

Conversely, no alternative $a_j > a_{i+1}$ can beat a_i if a_{i+1} cannot beat a_i, since alternative a_i's vote share is strictly greater than the maximum share for alternative a_{i+1}, which is the upper bound for any alternative $a_j > a_{i+1}$. Similarly, no alternative $a_j < a_i$ other than a_{i-1} and a_1 can beat a_i if neither a_{i-1} nor 1 can beat a_i. □

Finally, the last result of this subsection looks at the last case; which is essentially the converse of conditions from the first case (i.e. Lemmata 1 and 2). More importantly, in this proposition we construct sufficient and necessary conditions under which any alternative a_i is the k-winner.

Proposition 9. *Let $k \geqslant \max\{m, n-m\}$. For any $a_1 < a_i < a_n$ and for any given k, alternative a_i is the k-winner if and only if*

$$a_{i+1} - a_{i-1} > \max\left\{a_1 + a_{n-k+2}, a_i - a_{i-n+k-2}, a_{n-k+i+1} - a_{i+1}, 2 - a_n - a_{k-1}\right\}$$

Proof. The most damaging collection of contending alternatives, for any a_i, must include the following contenders: a_1, a_{i-1}, a_{i+1}, and a_n. Now, by including a_{i-1} and a_{i+1} in a_K, we maximise the contenders vote share as well as minimise a_i's vote share. Then alternative a_i's share would be $\frac{1}{2}(a_{i+1} - a_{i-1})$. Firstly, alternative 1's vote share is maximised when we push all $k-1$ alternatives to the right, i.e. $K = \{a_1, a_{n-k+2}, \ldots,\}$; so a_1 then beats a_i if and only if $a_1 + a_{n-k+2} > a_{i+1} - a_{i-1}$. Then, Alternative a_{i-1}'s vote share is maximised when $K = \{a_1, \ldots, a_{i-n+k-2}, a_{i-1}, \ldots\}$; and a_{i-1} then beats a_i if and only if $a_{i-1} - a_{i-n+k-2} > a_{i+1} - a_{i-1}$. Alternative a_{i+1}'s vote share is maximised when $K = \{a_1, \ldots, a_{i+1}, a_{n-k+i+1}, \ldots, a_n\}$; and $i+1$ then beats i if and only if $a_{n-k+i+1i} - a_{i+1} > a_{i+1} - a_{i-1}$. Finally, alternative a_n's vote share is maximised when $K = \{a_1, \ldots, a_{k-1}, a_n\}$; and a_n then beats a_i if and only if $2 - a_n - a_{k-1} > a_{i+1} - a_{i-1}$. Consequently, a_i cannot be the k-winner if one of these conditions holds.

Conversely, either $a_1 + a_{n-k+2}$ or $a_i - a_{i-n+k-2}$, which are the vote share of the contenders 1 and $i-1$, constitutes the upper bound for alternative j's vote share, for any $j > i$. Now the condition implies that alternative i beats both alternatives 1 and $i-1$. Thus, only alternative a_i can be the k-winner. Similarly, either $a_{n-k+i+1} - a_{i+1}$ or $2 - a_n - a_{k-1}$ represents the upper bound for the vote share of any alternative to the left of a_i. Since the condition implies that alternative a_i beats both alternatives a_{i+1} and a_n then only a_i must to be the k-winner. □

2.4.3 Alternatives exterior to N

In this subsection we present a special case: alternatives exterior to N. The next two propositions provide sufficient and necessary conditions under which the two alternatives exterior to N (i.e., a_1 and a_n) are the k-winners for any given k.

Given their spatial location, alternatives exterior to N must be considered differently. The reader should note that the calculation of the vote share of the alternatives exterior

to N depends on the identity of the neighbouring alternative. For example, the vote share for a_1 would be different if a_2 is present instead of a_3. Therefore, we will use similar strategies to prove the next two propositions. In the next proposition, for example, we fix a subset of k alternatives including alternative a_1. Then, we look at the most damaging collection of alternatives for alternative a_1. In that case, we must consider every possible case: look at every alternative that could lie next to alternative a_1. Then we derive conditions under which, alternative a_1 would beat the remaining alternatives in every case. Finally, we generalise this sequence of conditions. The resulting set of conditions are both sufficient and necessary for alternative a_1 to be the k-winner for every $k > 2$.

Proposition 10. *Alternative a_1 is the k-winner iff* $\quad \forall a_1 < a_j \in N, \forall a_\ell < a_j$:

$$a_1 + a_2 > \max_{1 \leqslant \ell \leqslant k-2} \left\{ a_{n-k+\ell+2} - a_\ell, 2 - a_{k-1} - a_k \right\} \tag{NSC1}$$

Proof. First, fix $K = \{a_1, a_{i_1}, a_{i_2}, \ldots, a_{i_{k-1}}\}$. Given this subset, we will compute the maximum vote share every alternative other than a_i could have. Now, we look at every case where a_j could take the other possible locations in $K \setminus \{a_1\}$.

1. Case $j \equiv i_1$. For alternative 1 to beat alternative $j = i_1$ we need: $\theta(a_1; K) > \theta(a_{i-1}; K)$, which implies: $a_1 + a_{i_1} > a_{i_2} - a_1$. The farthest location for i_2 has to be $n - k + 3$. Then:
$$a_1 + a_2 > a_{n-k+3} - a_1 \quad (\text{and } a_j > a_1) \tag{A-1}$$
2. Case $j \equiv i_2$. Following the previous case, we must have $\theta(a_1; K) > \theta(a_{i_2}; K)$, which implies: $a_1 + a_{i_1} > a_{i_3} - a_{i_1}$. The farthest location for i_3 has to be $n - k + 4$ and for i_1 is 2. Then:
$$a_1 + a_2 > a_{n-k+4} - a_2 \quad (\text{and } a_j > a_2) \tag{A-2}$$
3. Case $i_3 < j \equiv i_\ell < i_{k-2}$ Again, for alternative 1 to beat alternative $j \equiv i_\ell$ it would be necessary and sufficient that $\theta(a_1; K) > \theta(a_{i_\ell}; K)$, which implies: $a_1 + a_{i_{i_1}} > a_{i_{\ell+1}} - a_{i_{\ell-1}}$. The farthest location for i_ℓ has to be $n - k + \ell + 2$ and for i_ℓ is $\ell + 1$. Then:
$$a_1 + a_2 > a_{n-k+\ell+2} - a_\ell \quad (\text{and } a_j > a_\ell) \tag{A-ℓ}$$
4. Case $j \equiv i_{k-2}$. Again, $\theta(a_1; K) > \theta(a_{i_{k-2}}; K) \implies a_1 + a_{i_{i_1}} > a_{i_{k-1}} - a_{i_{k-3}}$. Then, we must set i_{k-3} to be $k - 2$ and i_{k-1} to be n. Then:

$$a_1 + a_2 > a_n - a_{k-2} \quad (\text{and } a_j > a_{k-2}) \tag{A-k-2}$$

5. Case $j \equiv i_{k-1}$. Lastly, we have

$$\theta(a_1; K) > \theta(a_{i_{k-1}}; K) \implies a_1 + a_{i_{i_1}} > 2 - a_{i_{k-2}} - a_{i_{k-1}}.$$

Then, we must set i_{k-2} to be $k-1$ and i_{k-1} to be k. Then:

$$a_1 + a_2 > 2 - a_{k-1} - a_k \tag{A-k-1}$$

□

In the proof of our last proposition, we follow the same case-by-case approach we used in the proof of the previous proposition.

Proposition 11. *Alternative a_n is the k-winner iff* $\quad \forall a_j < a_n,\ \forall \ell > j$:

$$2 - a_n - a_{n-1} > \max_{2 \leqslant \ell \leqslant k-1} \left\{ a_{n-k+2} + a_{n-k+3},\ a_{n-k+\ell+2} - a_{\ell-1},\ a_n - a_{k-1} \right\} \tag{NSCN}$$

Proof. First, fix $K = \{a_{i_1}, \ldots, a_{i_{k-1}}, a_n\}$. Alternative a_n will be the k-winner if and only if $\forall a_j < a_n :\ \theta(a_n; K) > \theta(a_j; K)$. Then we will look at every case where a_j could take the other possible locations.

1. Case $a_j \equiv a_{i_1}$. The first necessary and sufficient condition would be: $\theta(a_n; K) > \theta(a_{i_1}; K) \implies 2 - a_n - a_{n-1} > a_{i_2} + a_{i_1}$. Since i_1 must be $n-k+2$ and i_2 must be $n-k+3$. Then:

$$2 - a_n - a_{n-1} > a_{n-k+2} + a_{n-k+3} \tag{n-1}$$

2. Case $a_j \equiv a_{i_2}$. Then we must have:

$$\theta(a_n; K) > \theta(a_{i_2}; K) \implies 2 - a_n - a_{n-1} > a_{i_3} - a_{i_1}.$$

The farthest location for i_3 has to be $n-k+4$ and for i_2 is 1. Then:

$$2 - a_n - a_{n-1} > a_{n-k+4} - a_1 \tag{n-2}$$

3. Case $a_{i_2} < a_j \equiv a_{i_\ell} < a_{i_{k-1}}$. Again, we must have: $\theta(a_n; K) > \theta(a_{i_\ell}; K) \implies a_{n+1} - a_{n-1} > a_{i_{\ell+1}} - a_{i_{\ell-1}}$. The farthest location for $i_{\ell+1}$ has to be $n-k+\ell+2$ and for $i_{\ell-1}$ is $\ell - 1$. Then:

$$2 - a_n - a_{n-1} > a_{n-k+\ell+2} - a_{\ell-1} \tag{n-ℓ}$$

4. Case $a_j \equiv a_{i_{k-1}}$. Lastly, we must have

$$\theta(a_n; K) > \theta(a_{i_{k-1}}; K) \implies 2 - a_n - a_{n-1} > a_n - a_{i_{k-2}}$$

Then, set i_{k-2} to be $k-1$. Then:

$$2 - a_n - a_{n-1} > a_n - a_{k-1} \qquad \text{(n-}k\text{-1)}$$

□

2.5 Examples

In this subsection, we present four examples which illustrate the following cases:

1. An alternative exterior to N may be the k-winner for every k.

2. The 2-winner and the 3-winner both exist but differ.

3. The 2-winner and the n-winner coincide but no k-winner exists for every $2 < k < n$.

4. The 2-winner and the n-winner differ but no k-winner exists for every $2 < k < n$.

The interval for every graph is the unit interval. The black marks represent the position of the alternatives. Finally, the blue-shaded area represents the vote share of the 2-winner whereas the red-shaded represents the vote share of the n-winner. In the first and third examples, the 2-winner and the n-winner is the same alternative, in that case there is only one shaded area. The main difference between the first two cases and the last two is that in the latter there exists a k for which there is no k-winner. The reader must also notice that given the space limitations, the axes on the four figures below are shown uneven/unequal.

2.5.1 First Case

In this first case, we have a graphical example of Proposition 2. One of the alternatives' vote share is strictly greater than half. Alternative 3's vote share is 0.505 and thus alternative 3 is the k-winner for every k. Moreover, this scenario is also a graphical example of Proposition 3 since alternative 3 is an alternative exterior to N.

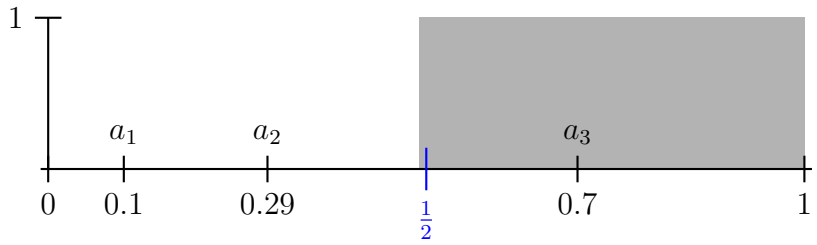


Figure 2.3: An alternative exterior to N is k-winner for every k.

2.5.2 Second Case

Now, we have that the 2-winner and 3-winner exist but they are not the same alternative. Alternative 2 is the 2-winner whereas alternative 3 is the 3-winner. As the reader can inspect, the position of the alternatives in this case are very similar to the previous one. However, alternative 2's position in this case is closer to the median voter's ideal; close enough to become the voter's ideal and thus becoming the median alternative.

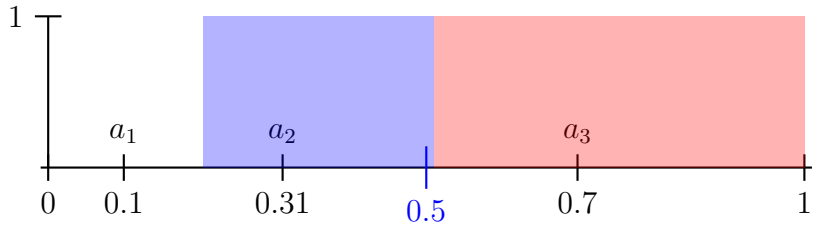


Figure 2.4: The 2-winner and the 3-winner both exist but differ.

2.5.3 Third Case

In this particular case n=4, such that the 2-winner and n-winner(i.e., 4-winner) are the same alternative but the k-winner for $2 < k < n$ does not exist.

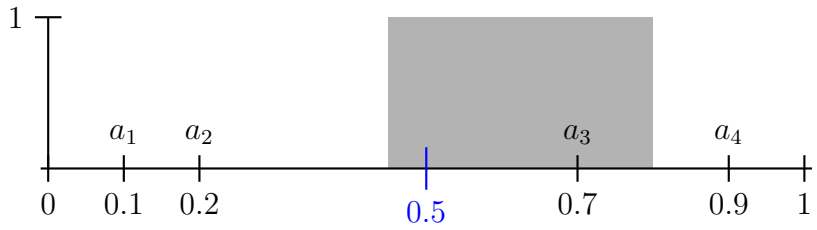


Figure 2.5: The 2-winner and the n-winner coincide but no k-winner exists for every $2 < k < n$.

2.5.4 Fourth Case

Also in this case we have that n=4. In this configuration the 2-winner and 4-winner exist but they are not the same alternative. However the 3-winner does not exist.

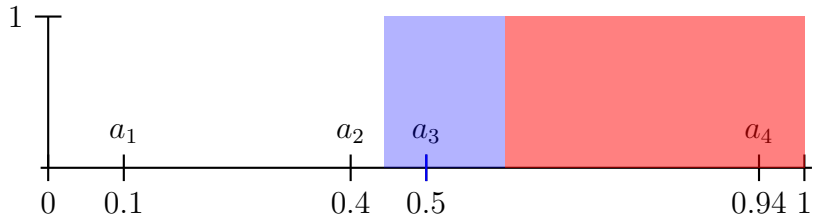


Figure 2.6: The 2-winner and the n-winner differ but no k-winner exists for every $2 < k < n$.

2.6 Concluding remarks

In 1785, the Marquis of Condorcet conceptualised a rather simple mechanism to identify the electoral winner among many alternatives: look at the winner of every pair. This criterion would remain dormant for centuries, until in the 1950's the Scottish economist Duncan Black would not only resurrect the Marquis concept but also derive another simple but potent concept: single-peakedness. Since Black, the idea of a Condorcet winner has become central to many models in Political Economy (e.g., Median Voter Theorem). In this chapter, we have used an axiomatic approach to present a natural extension of the Condorcet winner: a k-winner. Instead of looking at pairs of alternatives, in our model we look at subsets composed of k alternatives.

The main contribution is twofold: (i) we present sufficient and necessary conditions for any alternative to be the k-winner for some $k \geqslant 2$, and (ii) we show that if a k-winner exists for some $k > 2$ then this alternative is also the k'-winner for all $k < k' \leqslant n$. Moreover, this present chapter also contributes to the literature by studying the identification, taxonomy, and nomenclature of electoral winners beyond the well-known Condorcet-winner, using the spatial model of voting. This contribution has a palpable application to the theoretical and empirical analysis of modern American politics. For example, Kurrild-

Klitgaard (2018) has claimed that in the 2016 Republican Party presidential primaries, Donald Trump was simultaneously the Condorcet-loser (i.e., the 2-loser in our notation) and the plurality winner (i.e., the n-winner in our notation). This taxonomic analysis helps to elucidate the strengths and weaknesses of every candidate in the primary process.

This constitutes a natural link between the idea of k-winners and the concepts and results that will be presented in Chapter 3. American primaries are both unique and complex. At every stage there are (i) different voting rules, (ii) different assignment of delegates, and more importantly (iii) the subset of candidates remaining can be dramatically different (and more precisely the subset of candidate shrinks as the primary calendar goes on).

There are several avenues on which the results of this chapter can be extended further. Mainly, Proposition 4 can be translated into a testable prediction. However, the econometric challenge seems almost insurmountable since the polling and electoral data would not contain all the ranked preferences of the respondents. That is, most of the polls only ask for top-rank preferences.

2.7 References

Arrow, K. J. (1951). *Social Choice and Individual Values*. New York: John Wiley & Sons.

Barberá, S. (1983). Strategy-Proofness and Pivotal Voters: A Direct Proof of the Gibbard-Satterthwaite Theorem. *International Economic Review*, *24*(2), 413–417.

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

Blackorby, C., Donaldson, D., & Weymark, J. A. (1990). A Welfarist Proof of Arrow's Theorem. *Recherches Économiques de Louvain/Louvain Economic Review*, *56*(3-4), 259–286.

Blackwell, C. W. (2003). Athenian Democracy: A Brief Overview. *CW Blackwell Classical Athenian Democracy–STOA*.

Blau, J. H. (1971). Arrow's Theorem with Weak Independence. *Economica*, *38*(152), 413–420.

Blau, J. H. (1972). A Direct Proof of Arrow's Theorem. *Econometrica*, *40*(1), 61–67. Retrieved from http://www.jstor.org/stable/1909721

Condorcet, M. d. (1785). *Essai Sur L'Application de L'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix*. de l'Imprimerie Royale.

Dodgson, C. (1876). A Method of Taking Votes on More Than Two Issues.

Feldman, A. M. & Serrano, R. (2008). Arrow's Impossibility Theorem: Two Simple Single-Profile Versions.

Fishburn, P. C. (1970). Arrow's Impossibility Theorem: Concise Proof and Infinite Voters. *Journal of Economic Theory*, *2*(1), 103–106.

Geanakoplos, J. (2005). Three Brief Proofs of Arrow's Impossibility Theorem. *Economic Theory*, *26*(1), 211–215.

Gibbard, A. (1973). Manipulation of Voting Schemes: A General Result. *Econometrica*, *41*(4), 587–601. Retrieved from http://www.jstor.org/stable/1914083

Kurrild-Klitgaard, P. (2018). Trump, condorcet and borda: Voting paradoxes in the 2016 republican presidential primaries. *European Journal of Political Economy*, *55*, 29–35.

Lull, R. (1299). *De arte eleccionis*. Cod. Cus. 83.

Mas-Colell, A. & Sonnenschein, H. (1972). General possibility theorems for group decisions. *Review of Economic Studies*, *39*(2), 185–192.

McLean, I. (1990). The Borda and Condorcet Principles: Three Medieval Applications. *Social Choice and Welfare*, *7*(2), 99–108.

Meyers, A., Orrison, M. E., Townsend, J., Wolff, S., & Wu, A. (2014). Generalized Condorcet winners. *Social Choice and Welfare*, *43*(1), 11–27. doi:https://doi.org/10.1007/s00355-013-0765-8

Newing, R. A. & Black, D. (1951). *Committee Decisions with Complementary Valuation*. London: William Hodge.

Reny, P. J. (2001). Arrow's Theorem and the Gibbard-Satterthwaite Theorem: A Unified Approach. *Economics Letters*, *70*(1), 99–105.

Saari, D. G. (1976). Then-body problem of celestial mechanics. *Celestial mechanics*, *14*(1), 11–17.

Saari, D. G. (1989). A Dictionary for Voting Paradoxes. *Journal of Economic Theory*, *48*(2), 443–475. doi:https://doi.org/10.1016/0022-0531(89)90037-9

Saari, D. G. (2016). From Arrow's theorem to 'dark matter'. *British Journal of Political Science*, *46*(1), 1–9.

Satterthwaite, M. A. (1975). Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. *Journal of Economic Theory*, *10*(2), 187–217.

Sen, A. (2017). *Collective choice and social welfare*. Penguin Random House.

Vishnia, R. F. (2012). *Roman Elections in the Age of Cicero: Society, Government, and Voting*. Routledge.

Wilson, R. (1972). Social Choice Without the Pareto Principle. *Journal of Economic Theory*, *5*, 14–20.

Appendix A: Non-Uniformly Distributed Ideals

Model with a non-uniform distribution

Let the voter's ideal alternatives follow a distribution f, with support on $[0,1]$, which we assume to be continuous and to have a unique median m and cumulative distribution function F. Now, fix a subset $K \subseteq N$ with cardinality equal to k, such that the alternatives $\{a_{i^-}, a_i, a_{i^+}\} \subseteq K$ and let a_{i^-}, and a_{i^+} be the respective left and right "neighbour" alternatives for a_i such that

$$a_{i^-} \equiv \max\{a_j \in K : a_j < a_i\} \text{ and } a_{i^+} \equiv \min\{a_j \in K : a_j > a_i\}.$$

Given a $K \in N$, we now re-define the proportion of voters whose ideal alternative is $a_i \in K$ as,

$$\theta(a_i; K) = \begin{cases} F\left(\dfrac{a_i + a_{i^+}}{2}\right) & \text{if } a_i \equiv \min\{a_j \in K\} \\ 1 - F\left(\dfrac{a_i + a_{i^-}}{2}\right) & \text{if } a_i \equiv \max\{a_j \in K\} \\ F\left(\dfrac{a_i + a_{i^+}}{2}\right) - F\left(\dfrac{a_i + a_{i^-}}{2}\right) & \text{otherwise} \end{cases}$$

Proposition 10. *Let $k \leqslant \min\{m, n-m\}$. For a given $k > 2$, the median alternative a_m (which is interior to N) is the k-winner if and only if*

$$\forall K \in N \text{ such that } \{a_{m-1}, a_m, a_{m+1}\} \in K : \begin{cases} F\left(\dfrac{a_m + a_{m-1}}{2}\right) \leqslant \theta(a_m; K) \\ 1 - F\left(\dfrac{a_m + a_{m+1}}{2}\right) \leqslant \theta(a_m; K) \end{cases}$$

Otherwise, there is no k-winner.

Proof. If $k \leqslant \min\{m, n-m\}$ then no alternative a_i different than a_m can be a k-winner as established in Lemmata 1 and 2. Now, we need to construct the most damaging collection for alternative a_m, which would minimise a_m's vote. This collection would

necessarily include her neighbours a_{m-1} and a_{m+1}, yielding alternative a_m a share of $\frac{1}{2}(a_{m+1} - a_{m-1})$. Now, suppose the following two conditions hold,

$$F\left(\frac{a_m + a_{m-1}}{2}\right) \geqslant \theta(a_m; K) \text{ and}$$
$$1 - F\left(\frac{a_m + a_{m+1}}{2}\right) \geqslant \theta(a_m; K)$$

Take the first condition, which implies that the proportion of voters that top-rank alternative a_{m-1} is strictly greater than the proportion of voters that top-rank alternative a_m, if the subset of alternatives were $K = \{a_{m-1}, a_m, \ldots\}$. Then clearly alternative a_m cannot be a k-winner. Nonetheless since $k \leqslant m \leqslant n - k$ also holds, then there cannot be a k-winner, since the only potential alternative to be a k-winner was alternative a_m. Taking the second condition, we have the exact same argument as before but for alternative a_{m+1} and it would also imply that there cannot be a k-winner (given the fact that alternative a_m fails to be a k-winner).

If neither of these conditions hold then a_m is the k-winner. To show this, take any K which includes a_{m-1}, a_m and a_{m+1} and any $a_j \in K$ such that $j \leqslant m+1$. Removing alternatives a_j for any $j < m-1$ and replacing them with some alternatives a_j for any $j > m+1$, increases the vote share of some remaining alternatives a_j for any $j < m-1$ without increasing the vote share for alternative a_j for any $j > m$. However, the increased vote shares of a_j for any $j < m$ are bounded above by $F\left(\frac{a_m + a_{m-1}}{2}\right)$; and alternative a_m then wins against a_{m-1}. The same argument applies to removing any alternative a_j for any $j < m - 1$. Finally, removing alternative a_{m-1} increases a_m's vote share and cannot increase the vote share of alternatives a_j for any $j < m-1$ above $1-F\left(\frac{a_m + a_{m+1}}{2}\right)$ which, by supposition is less than $\theta(a_m; K)$, which itself is less than a_m's new increased vote share (given alternative a_{m-1}'s removal) . An equivalent argument precludes eliminating alternative a_{m+1}. □

Chapter 3

A Model of Sequential Elections: The Case of American Primaries

Abstract

We present a theoretical model to analyse two phenomena surrounding sequential elections: momentum and Duverger's law. A candidate is said to experience "momentum effects" when she wins the first primary elections in such a way that it would guarantee a greater probability of becoming the overall winner. On the other hand, our modified version of Duverger's law asserts that primary candidates tend to quit at earlier stages of the election under plurality rule compared to proportional representation. Our theoretical results will challenge both assertions. More specifically, we find conditions under which both phenomena are precluded. Furthermore, we will exploit the dynamic nature of primary elections by using the concept of a Condorcet-winner and k-winners. The k-winner is the alternative that is top-ranked by the majority in every electoral competition composed exactly of a subset of k candidates.

JEL codes: C72, D72

Keywords: elections, primaries, Condorcet winner, subgame perfect equilibrium

3.1 Introduction

The 2016 presidential election in the United States marked the fourth occasion which the winner of the popular vote did not get enough Electoral College votes and therefore did not become President (the three previous occasions were 1876, 1888, and 2000). This fact brought back controversy and discussion about the redesign and reform of the electoral structure and rules in America. However, there was already controversy around the nomination process of each political party. Naturally, this process has also been the centre of attention for many academics from several academic disciplines.

The presidential elections process in the United States is unique in the world. Not only does the Electoral College, composed of delegates directly elected by voters, ultimately determine which candidate will become the President of the United States, but another set of delegates congregates at the national convention of each party and chooses the official nominee. The most striking aspect about this electoral scheme is the sequential aspect of the primary elections: voters in Iowa go to the ballots early in February while Californian voters wait until June. However, the process as we know it only emerged in the late 1960's.

The Constitution of the United States is often claimed as the "oldest written (national) constitution still in use today," and thus it has almost reached the status of a sacred text amongst American politicians. It was supposed to avoid the abuse of power by a single tyrant but also to reach a compromise between the different regional interests and concerns at the time, explicitly trying to balance the power within the federal government and between the states. Ultimately, we can find in its pages the primordial source of all political rules and designs for the government at the national level.

Nonetheless, there is a significant problem: the concept of political parties does not appear at all in the constitutional text. The Founding Fathers could not have envisioned (while writing the constitution) all the complex political ramifications that ensued. For example, the practicalities of governing changed Madison's approach to parties. The original intention was to leave these particular matters to the states, to promote a balance

of political power between federation and states. Therefore, political parties emerged as national private clubs/ organisations trying to advance a political agenda and thus each party could establish its own arbitrary set of rules for every aspect of the party organisation.

In the last year of his administration, George Washington (1796) foresaw the complex political climate in his Farewell Address, where he warned that "... any ground should have been furnished for characterizing parties by geographical discriminations." Washington then cautioned about "... designing men may endeavor to excite a belief that there is a real difference of local interests and views." Not only did he warn about the "... danger of parties in the State, with particular reference to the founding of them on geographical discriminations..." but he also emphasised "the baneful effects of the spirit of party generally."

Despite Washington's best advice, political parties took over and quickly polarised voters. For example, the controversies around the design of economic policies of the first government created the most fundamental division among the writers and designers of the constitution: Jefferson (and Madison) founded the Democratic-Republican Party in 1791 while Hamilton founded the Federalist party one year later.

The Federalist approach to party hierarchy translated into the design of the first political parties. The founders of any given political party (and then their successors) had absolute control over the party and its regional/local branches. They chose the rules of the procedures, the candidates, the policies, etc.

From the early 1800s to the early 1900s, candidates were chosen officially at the party's National Convention, where hand-picked delegates from the regional branches would get together and nominally select their candidate; behind closed doors, all the political bosses and party ranking members would have previously agreed on the decision. The procedure was efficient since the negotiations only took place between a handful of politicians, and it was relevant because the convention served as a perfect simulated mechanism of democracy for the party rank and file.

In the early 1910's, several states (e.g., Florida, Wisconsin, Oregon) enacted laws to

introduce primary elections, which were considered as a natural step towards a more democratic system. However, there was a major setback during the inter-war period. Most of the states that had decided to implement a primary saw voters rather uninterested by the additional elections and candidates circumventing the process (see Davis (1997)). These factors along with rising operational costs gave local politicians enough reasons to repeal the primary elections format. It was only after WWII that primary elections became more prominent across the country.

However, there were crucial times when the will of the voters did not match the intentions of the party leaders. A clear example would be the Democratic National Convention of 1940, where delegates unanimously voted for the sitting president Roosevelt. However, the front-runner candidate for the Vice-Presidency, Henry Wallace, posed a threat to the party leaders. Wallace had not only had created a name for himself during his time as the head of the Department of Agriculture, but also had the support of Roosevelt, and was viewed positively by the majority of party members. Nonetheless, party bosses and ranking members of the Democratic party considered him too progressive and an insufficiently strong candidate to compete with the Republicans. Ultimately, after winning almost all the necessary votes at the Convention, the chairman of the Democratic National Committee (DNC), Robert Hannegan, flexed his muscle and completely turned the procedure against Wallace. The winner and eventual Vice Presidential candidate would be a rather modest and mostly unknown politician named Harry S. Truman.

The number of primaries in both main parties fluctuated across the first four decades of the last century. However, the system faced its most challenging times during the Democratic Convention of 1968, where Hubert H. Humphrey became the official Democrat nominee without actually winning a single primary/caucus. Party members realised that the ultimate decision indeed resided in the hands of the party's leaders. That major incident resulted in the creation of two commissions formed within the Democratic party: the Hughes Commission in 1968 and the McGovern-Fraser Commission from 1969 to 1972. The reports of both commissions paved the way for the current complex set of guidelines to choose a presidential candidate. Not only Democrats but also Republicans

decided to have a more democratic procedure to select their candidates.

This chapter is organised as follows. In Section 3.2, we review the relevant literature, including the literature on sequential and primary elections, momentum, and Duverger's Law. In Section 3.3, we introduce the preliminaries of our model; in Section 3.4, we present a set of benchmark results for the case of one region only and in Section 3.5, some initial results for the case of two regions are presented. In Sections 3.6 and 3.7, we present the results on momentum and Duverger's law. In Section 3.8, we compare the results from this chapter with the framework presented in the previous chapter: the spatial setting. Finally, in Section 3.9, we discuss our results and conclusions.

3.2 Literature Review

Primary elections have been studied in depth both by political scientists and game theorists in the last decades. The most attractive feature of primaries is its inherently dynamic features: it usually is the first stage to the general election. In the literature, there have been different approaches to analyse primaries.

3.2.1 Sequential Elections

There is a stream of research concerned with the contrast between sequential and simultaneous elections. The main theme in this stream of the literature is information asymmetries. The seminal paper by Dekel and Piccione (2000) provided a remarkable theoretical result that challenged theoretical intuitions about informational cascades and even some empirical findings at that time. They used a basic theoretical set-up with n identical players voting across two alternatives, using a majoritarian voting rule to determine the outcome. Voters in this model choose their alternatives supposing they are pivotal. However, voters do not gain any information in a sequential voting game because "the equilibrium is symmetric" and "voters are identical ex ante" (Dekel and Piccione, 2000). Therefore, the punchline is that the symmetric equilibria in simultaneous voting games are also equilibria in its sequential counterpart. More specifically all symmetric equilibria of the simultaneous voting games form a subset (not necessarily proper) of the

equilibria of the sequential voting games. In contrast to their paper, we do not consider information asymmetries in our model since we assume the players have complete and perfect information.

Nonetheless, Battaglini (2005) showed that the introduction of an "arbitrarily small cost of voting" was sufficient to make these two subsets of equilibria disjoint. This small cost would change how the voter calculates her probability of being pivotal; and this in turn affects her decision to either abstain or not. Ultimately, abstention can distort how information is aggregated. In our model, voters vote sincerely: they are almost not pivotal because there is a continuum of voters.

In the context of American primary elections, candidates do not decide to leave the race simultaneously. This dynamic feature is central to our model. In this sense, a paper close to the present work is Ishihara and Miura (2017), where the citizen-candidate framework was equipped with sequential entry announcements. The main result in the paper was how the minor candidate (defined as the candidate that could only win when running unopposed) becomes a central player in the election because her decision could radically alter the result of the competition. This paper had a novel perspective on what would have been considered an unimportant candidate in this voting context: a minor candidate (a 2-loser in our notation). The minor candidate enters the race, despite the cost of entry and the fact that she would surely lose, to ensure that her favourite rival (the candidate whose win would provide higher payoff) win ultimately. The authors also identified a scenario where the Condorcet-winner happens to be also the Grand-winner, where she would surely win under any circumstance (i.e. this candidate is "sufficiently strong and always wins regardless of other candidates' preferences.") In our notation, we have defined this type of candidates as the "k-winner for any k". The focus of the paper is the balance of electoral power between the candidates in the context of sequential elections, whereas in this chapter we focus on how voting rules directly impact on the candidates' decision to stay or quit the race (which relates to the number of candidates remaining). Another significant difference from our paper is that the order in which candidates decide to stay/leave the contest is exogenously determined.

Lengle and Shafer (1976) was an early and quite significant contribution to the literature that studies the direct relation between voting rules and how delegates are awarded to candidates in American primaries and the actual outcome of the election. Their analysis focused on the Democratic Presidential primaries of 1972, more specifically on the first fifteen elections. The two authors considered the three main voting rules used in the primary: Winner-Take-All, Proportional, or Districted allocations. They also distinguished between what they considered the three main factions within the party: (i) centrists from "old, eastern, industrial states", (ii) liberals from "Mid-western Western", and (iii) social conservatives from the "Southern Democratic" region. The three main candidates in the election were: (i) the eventual nominee Senator George McGovern, (ii) former Vice President (and previous nominee in 1968) Hubert H. Humphrey, and Alabama Governor George Wallace. Their main focus of analysis was to compare the number of delegates the three candidates would have been able to get if all the first fifteen primaries had the same voting rule. Then, the main result links voting rules and the type of candidate nominated: plurality rule would tend to favour centrist candidates (i.e. Humphrey), and proportional representation (and Districted allocations) would substantially favour liberal candidates (i.e. McGovern). The explanation behind this result relies mostly on the geopolitical distribution of Democratic voters: plurality rule advantages larger states (e.g., Pennsylvania with 137 delegates, Ohio with 140 delegates, and California with 271 delegates).

Furthermore, another interesting conjecture was made by Hammond (1980): the three candidates were favoured under plurality rule in their own states, but more strikingly they were hindered under proportional rules in their own states. Additionally, both McGovern and Humphrey were also favoured under proportional rules outside their home states.

A comprehensive empirical analysis can be found in Ansolabehere and King (1990) , where the main focus was two properties of voting rules: responsiveness to changes in voters' preferences and candidate bias. The authors defined a responsive voting system as "... one where small changes in voter preferences lead to relatively large changes in candidate support among a state's delegation to the convention." and candidate bias

as the "... deviations from partisan symmetry" (as specified in King (1990)). In other words, it measures how (un)fairly delegates are distributed among candidates. The paper relied on data from the primary elections of both Democratic and Republican parties from 1976 to 1988. The results indicate that Republican primaries were more responsive than their Democratic counterparts and non-PR primaries were "more responsive" than PR primaries. Finally, they provided evidence of a strong bias against one candidate, Jesse Jackson, in the non-PR Democratic primaries of 1988.

Our paper is particularly interested in understanding the relation between delegation allocation rules (i.e., voting rules) and the possible existence of two political phenomena: momentum and Duverger's Law.

3.2.2 Momentum

Momentum has received a lot of attention in the literature, mainly because many academic and non-academic analyses describe different aspects of the same political phenomenon. In the media it is known as "momentum" or the so-called "New Hampshire effect"; in the Political Science literature it is described as front-loading, or winnowing. In the Economics literature, it has been studied as an instance of social learning: that is, as a particular case of information asymmetries.

The main idea behind Momentum is based on the fact that Californian voters would have more (and better) information about the candidates than voters from Iowa, not only from hypothetical polling but actual electoral results. On the other hand, Iowa voters get to choose from a larger selection of candidates (i.e., typically the ballot in Iowa has the largest number of candidates) whereas Californian voters have to choose only from the candidates that have not decided to quit the race.

Another crucial aspect of this set of phenomena is the non-static nature of the list of candidates. Candidates will eventually decide to "suspend" their campaign (i.e., quit the race) because of the financial cost of maintaining a campaign.

Our results are directly motivated by all the possible scenarios when any given candidate decides to leave the race at any stage. The effect of monetary contributions has been

singled out as one of the main reasons for candidates to drop out (see Norrander (2006), Haynes et al. (2004)). Moreover, several papers claim that financial contributions also seem to affect the actual dynamics of elections. For example, Schwabe (2015) studied the economic and political incentives behind the electoral schedule. The author noted that changes to the schedule would alter the welfare of voters and the party, given different levels of campaign costs (which are financed by donors). Therefore, there exists a trade-off between the welfare of donors and the welfare of society. That is because more extensive (and thus expensive) campaigns can positively affect voters by providing them with more information about candidates allowing them to identify the best candidate whereas donors are interested in supporting a winning candidate while keeping expenditures to a minimum. Candidates in our model do not receive private financial contributions but instead fund each primary election out of their own pocket.

Momentum can be defined when a candidate, not necessarily the front-runner, starts the election season with a streak of successful events (polling or election results) resulting in a larger number of accumulated delegates which could translate into a higher probability of winning the race. Political pundits tend to highlight early success (in the primary calendar) as a definite measure of eventual success. However, the empirical evidence has not identified the causal relation definitively.

An interesting stylised fact is that out of the last ten competitive [i] Republican primary elections, only twice has the winner of the Iowa caucus become the eventual presidential nominee. Given the Iowa caucus is among the first contests, all the candidates had participated in this specific contest. Following Ishihara and Miura's (2017) definition, the winner of the Iowa election can be classified as the Grand-Winner (i.e., the candidate that wins when all candidates participate). Another supporting interpretation of this definition can be found in Steger et al. (2002), where the front-runner "... is the candidate who is doing better (in the polls and in fundraising efforts during the invisible primary and in the delegate count once the primaries begin) than any of the other candidates." For these authors, the decision whether to quit the race can be quite simple for front-

[i] Excluding candidates running unopposed.

runners because "they do not have to improve to gain the nomination; indeed, they may be able to capture the nomination even if their future performance is not equal to past performance." Ultimately, they conclude that front-runners would "never quit the race." Our theoretical results differ: we have identified specific parametric configuration where the Grand-Winner would have the lowest probability of winning the nomination, which is consistent with the stylised fact about Iowa winners.

The media and political pundits have discussed possible alternative primary formats. The most popular alternative scheme is the concept of regional primaries, where states would be grouped into regional units, and each unit would be assigned in sequence. The national primary format is the most common format used throughout the world, because of its simple rules and straightforward application. However, critics of a national primary always point out the enormous imbalance between the larger states (such as Texas and California, which coincidently have larger economies) and the smaller (and more rural-based) states. Lastly, other plans would order the electoral schedule according to the population size of the states. In Norrader (2008) the reader could find a thorough analysis of the regional primaries proposals.

However, in the literature, readers can find papers that provide reasoning that highlights the counterintuitive positive effects of the current sequential arrangement. Klumpp and Polborn (2006) concluded that fewer resources would be spent in sequential elections. In their theoretical model, candidates are more likely to focus resources on the first elections, but once a candidate emerges as the front-runner spending would be reduced. By contrast in a simultaneous election scheme, money would be exhausted entirely across all districts, supposing the number of districts is large enough. Given this, candidates would face lower expected expenditure levels in sequential elections. Additionally, it would allow candidates without national recognition and a smaller budget to participate in the race.

In a similar spirit, Deltas, Herrera, and Polborn (2015) showed that in sequential elections, there is a trade-off between voter coordination and the revealed quality of candidates. Then, sequential elections could minimise the adverse effects of vote splitting, by winnow-

ing out a candidate with less valence who shares the same political position as another candidate. By performing simulations, the authors showed that sequential elections outperform any other scheme (simultaneous or regional elections) by selecting the candidate with the highest valence and having the highest probability of a Condorcet-winner getting elected. In contrast, our model does not consider any heterogeneity in candidate quality or valence.

Candidacy length during presidential primaries has been analysed empirically. Given the extensive length of American primaries, candidates usually require (i) vast amounts of financial contributions throughout the campaign and (ii) media attention. Intuitively, a longer candidacy would require a larger amounts of campaign money and continuous notoriety in the media. However, there has not been a consensus in the literature to formalise this relation. Norrander (2006) provided evidence that voting rules and the primary schedule did not have an effect on the length of candidacy in the primaries from 1980 to 2004. But cash holding, poll standings, and quite importantly the results in Iowa and New Hampshire was statistically significant. However, Damore, Hansford, and Barghothi (2010) provided empirical evidence that contradicts Norrander's findings. Specifically, the only variables that explain candidacy length are money and media coverage.

Another sub-stream of papers focuses on the structure of the primary calendar. Selman et al. (2010) (and quite similarly Hummel and Holden (2014)) provided theoretical results which indicate that the optimal primary schedule should allow (a) small states, (b) states with "fewer partisans", (c) states with more informed voters and (d) states with a more diverse population to vote first. This arrangement would result in better information cascading through bigger electorates such that candidates with higher quality would be elected. In the same spirit, Schwabe (2015) found that an optimal voting schedule could include a block of states voting simultaneously earlier (i.e., Super Tuesday and Super-Duper Tuesday). Schwabe's results rely on the existence of high campaign costs, which is an observable fact from American elections.

There are several empirical results in the literature concerning voting rules in primary elections and the effect on the number of candidates and their decision to stay or quit the

competition. Firstly, Cooper (2002) concluded that PR rules tend to increase the length of candidacies. Using computer simulations, and allowing for different voter preferences to vary, the results seem to indicate that proportional representation rules are associated with a longer time for any candidate to reach enough candidates to become the nominee officially. However, Steger et al. (2002) used data from the presidential primaries contest of both parties from 1912 to 2000 and found that the effective number of candidates has increased since the reforms of the 1970s, for both parties. This set of results seems to indicate that the increased number of primaries used by political parties has fostered political competition.

Finally, Haynes et al. (2004) used survival analysis (using a Weibull distribution[ii]), to test which variables explained candidacy duration in the 2000 presidential primaries. These authors concluded that: (a) media coverage was only the most critical factor for dark horse candidates at earlier stages of the competition and (b) money is a significant component of a candidate's decision (to stay/quit the race), but is more important for candidates closer to the front-runner than for dark horse candidates. According to the authors, the effectiveness of financial support is not constant: it would wear out as the electoral calendar moves on, such that any candidate with initial lower poll standings would positively benefit from financial support at the beginning of the primary but its effect would vanish as the campaigns moves on. In the 2016 Republican presidential primary, for example, Jeb Bush despite having the most substantial contribution fund ($130 million) had to quit the race after losing every single primary.

An obvious question raised in literature is whether every voter has the same influence on the eventual result of the nomination. Knight and Schiff (2010) found that because of "momentum effects," early voters have disproportionate influence compared to late voters. This influence can be understood as the relative importance of the signals from the early elections (which candidate performed better) to the voters in the late primaries. Finally, Atkeson and Maestas (2008) put forward evidence that indicates front-loading and its associated factors (i.e. "delegate accumulation, the effective number of candidates,

[ii] A through reference on Survival Analysis can be found in Miller Jr (2011)

election compression, campaign spending, separation of presidential primary from state and local primaries, and earlier effective endings of the campaign") can curtail some voters (from the last elections in the calendar).

Because of the complexity of the issue, sometimes relevant data is not available or is not sufficient to model voter's preferences adequately. Hence, Potthoff (2011) proposed a polling procedure that would be efficient enough to implement and would retrieve richer datasets: Condorcet Polling. In this type of polling, voters would have to pick their preferred choice over each pair of candidates.

3.2.3 Duverger's Law

Duverger (1951) presented a rigorous taxonomical analysis of political entities. His work was a novel and refreshing approach in political science, not only because it was the first analytical description of the different party systems around the world, but also because Duverger made a significant breakthrough by proposing what he called a "true sociological law." His law states that "the simple-majority single-ballot system favours the two-party system." Additionally, he proposed the following hypothesis (without the status of law): "the simple-majority system with second ballot and proportional representation favours multi-party systems."

More than three decades later, Riker (1982) brought Duverger's (sociological) law to the mainstream academic literature in Political Science and Political Economy. Since then there has been a stream of research trying to either provide a (theoretical or empirical) rationalization for the law or test the findings in different countries such as Japan (Reed, 1990), Canada (Gaines, 1999), India (Diwakar, 2007), United States (Scarrow, 1986), and United Kingdom (Bowler et al., 2009). Duverger famously explained his law and hypothesis via a mechanical and a psychological effect. The former is just the mathematical effect of plurality rule; as a party needs the largest share of votes it would directly undermine every third party or small party. Then the real competition would be between the main/major parties which have a higher probability of obtaining the largest share of the votes. On the other hand, according to Duverger, there is a psychological drive from

voters to not vote for smaller parties because they would not want to "waste" their vote on a probable loser.

Voters assume that the important competition would be between the two major parties; so voting for a third party would effectively undermine one of the major parties and support the other one. In a sense, Duverger is arguing that third parties could end up splitting the vote and voters would prefer to use their vote more strategically.

The critical matter in Duverger's law is the number of political parties in a party system. This chapter aims to turn this law into the primary elections context. In that case, the analogous feature will be the number of candidates at a primary contest.

Naturally, the main difference is the fact that, irrespective of the electoral system, parties will only nominate one candidate (i.e., coalitions of candidates do not exist). However, we hypothesise that the rate at which candidates decide to leave the race will be different across electoral rules and different voting schedules/arrangements. For example, it would be natural to assume that a sequence of primaries using plurality rule would undermine the electoral prospects of the lagging candidates because their cumulative sum of delegates would be overrun entirely by that of the front-runner (or the set of front-runners). We can contrast this with a sequence of proportional representation primaries. In that case, lagging candidates would accumulate modest number of delegates in every election. Therefore, the rate at which the lagging candidates would start quitting in the former case would be faster than in the latter case.

Thus the modified Duverger's law version for sequential primary elections would assert that primary candidates tend to quit at earlier stages of the election under plurality rule compared to proportional representation.

According to Kamarck (2018) the Democratic and Republican parties have different incentives to arrange the voting rules for their respective nomination processes. After a sequence of commissions and committees investigating the effects of the different rules and regulations across almost five decades, the Democratic party decided in 1992 to restrict all states to use Proportional Representation rules. On the other hand, Republican officials had a stronger stance on states' rights and have left the choice of voting rule

to each state. However, according to Kamarck (2018) there is a natural "antipathy to Proportional Representation rules in the Republican Party". This antipathy has translated into the current rules of the Republican Convention which force "... early states to use proportional representation systems if they want[ed] to hold their contest before April" and this strategy "... actually succeeded in moving states to later in the season" (Kamarck, 2018, p.87).

Moreover, when Elaine Kamarck asked former chairman of the Republican National Committee Frank Fahrenkopf about the use of proportional representation rules in the Republican party, he answered: "We Republicans would never do that. Since the Electoral College is winner-take-all, why would you adopt a primary system that's different? When we go into a state, all the candidates build statewide organizations, and then they are prepared to run in the Electoral College" (Kamarck, 2018, p.87).

More substantially, we can find support for this modified Duverger's law in Kamarck's (2018) claim that Republican primaries (composed of a big proportion of states using PL rules) select a winner much faster and earlier than the Democratic contest (where all states must use PR):

> "Republicans tend to use winner-take-all systems that reward candidates who win by even the slimmest margins. This difference means that Democratic contests that make it past the early states can go on much longer than Republican contests. It also helps explain why four out of the five contested nomination contests listed above were in the Democratic Party. The same pattern held in 2008, when John McCain was able to wrap up his nomination several weeks before Barack Obama, in part because his primary wins were yielding greater results in the delegate count." (Kamarck, 2018, p.84).

Our theoretical results will challenge this assertion. More specifically in this chapter, we find conditions under which, in equilibrium, fewer candidates may remain after the first stages of a sequential primary election, under proportional representation than under plurality rule.

3.2.4 Additional Literature on Primaries

Primaries can be divisive and showcase the internal divisions of the party. Agranov (2016) highlighted what she calls the "divisive-primary effect." If a candidate does not emerge as the obvious nominee in the earlier stages from the primary process with a clear mandate, it could be driving up polarisation amongst party members. Instead of focusing on the general election strategies (grooming the candidate, preparing the attacks for the opposition, etc.), time would be wasted on internal battles. Ultimately, party leaders would prefer to propose a hand-picked candidate and gather support and donations within the party to focus on the general election. However in the post-reform period that would be beyond their control.

It is natural to ask if having a primary has any merits. Consequently, many papers have provided empirical and theoretical evidence that would explain the increasing use of primaries by political parties throughout the world. For example, Adams and Merrill (2008) found that by holding primaries political parties (including small ones) ultimately improve their chances of winning the general election. This result holds despite the many "dangers" of holding a primary such as: (i) "... candidates cannot modify their policies between the primary and the general election" and (ii) "primary voters may hold extremist policy views compared to the general electorate". The rationale behind this result relies on the primary election working as a "quality control" and thus helping to identify the candidate with the highest valence for the general election.

Moreover, Serra (2011) argued that party leaders are more likely to have a primary when the current candidates are "low-quality" (in this paper's context, quality is directly related to campaigning skills). Furthermore, Hortala-Vallve and Esteve-Volart (2011) developed a theoretical model to explain how primary elections would arise to bind a political party together after an inside faction threatens to diverge from party lines. Otherwise, a split would undermine their political efforts and would facilitate the electoral road for their adversaries.

There is another subset of the literature interested in the effects of the format of the

primaries. The most relevant types of format are open and closed. In the former, every registered voter can participate; in the latter, only registered members of the party can vote.

The debate up to Rogowski and Langella (2015) had not been conclusive about finding a link between electoral formats and the characteristics of the ultimate winner. The authors found that open primaries tend to favour more moderate candidates and conversely closed primaries would favour more extreme candidates. The authors defined candidates by their platforms; that is extreme candidates would implement polarised (or polarising) platforms. We do not incorporate party affiliation; so our primary elections are essentially open elections.

Most researchers agree that closed primaries would elect more extreme candidates because of (a) the lower turnout in primary elections (see DeSilver (2016)) and (b) median party members tend to prefer more extreme candidates compared to the median voter of the general election (see Serra (2011) and Brady et al. (2007)). Gerber and Morton (1998) found that congressional candidates from states with closed primaries are not closer to the district's estimated median voter's ideal position. They also concluded that open primary elections result in ideologically moderate candidates since they have to appeal to a broader audience. However, Cherry and Kroll (2003) questioned this conclusion; their claim was that open primaries would not necessarily result in moderate nominees. More specifically they found that "... semi-closed primary format generates more moderate candidate winners relative to the other primary formats."

3.3 Model

There are $r \in R = \{1, 2\}$ regions[iii], and each region is composed of a continuum of citizens who vote sincerely according to their ranking of preferences. Let $N = \{1, \ldots, n\}$ be the set of candidates, with representative element i. The reader must note that some results hold for $n = 4$ (see Proposition 4.5) or any number of candidates (see Proposition

[iii]The reader should interpret $r = 1$ as the case of a simultaneous election. Additionally, most of our main results hold for $r = 2$ exclusively but some minor results hold for any number of regions.

4.1, 4.2, 4.3, 4.4, and 4.6); however our main results hold for $n = 3$ exclusively. Each candidate i faces running costs, in the r-th region, $c_i^r \in \mathbb{R}_+$ and her benefit of holding office is $b_i \in \mathbb{R}_+$.

We represent electoral competition as an extensive-form game with perfect information. In our model each candidate makes an entry decision sequentially given a random order[iv] of moves for each region. Given an order of moves $o \in O$, we define history H_i to be the set of candidates who have decided to enter when candidate i makes a decision. For example, if we fix $r = 1$, $N = \{1, 2, 3\}$, and the order of moves is 123 then the history H_i satisfies $H_1 = \varnothing$, $H_2 = \{\varnothing, \{1\}\}$, $H_3 = \{\varnothing, \{1\}, \{2\}, \{1, 2\}\}$[v]. Let $A_i(h_i) \equiv \{\text{Stay, Quit}\}$ be the actions that candidate i can choose at every $h_i \in H_i$. Now, we define a pure strategy for candidate i to be a mapping $\sigma_i : H_i \to \{\text{Stay, Quit}\}$ that assigns an action $\sigma_i(h_i) \in A_i(h_i)$ after observing history $h_i \in H_i$. A strategy profile is a tuple of pure strategies: $\sigma \equiv (\sigma_1, \sigma_2, \sigma_3)$. The solution concept we focus on is subgame perfect Nash equilibrium.

A total number of $\rho = \rho_1 + \rho_2$ delegates are awarded sequentially to the candidates. Firstly, ρ_1 delegates are assigned in region 1 , then region 2 awards ρ_2 delegates. The candidate with the most delegates wins the primaries and becomes the nominee.

Another major component of the model is the choice of primary rules and how delegates are assigned. There are two electoral rules we will focus on: Plurality Rule (PL) and Proportional Representation (PR). Under the former, the candidate with the largest proportion of votes gets all the delegates. While under PR, each candidate gets a proportion of delegates, according to their share of votes. We are particularly interested in looking at the different combinations of voting rules: plurality rule or proportional representation.

We are particularly interested in looking at the different combinations of voting rules across a sequence of regions. Therefore, for $r = 2$ regions we define a set of games $\mathcal{G}^2 = \{G_1, G_2, G_3, G_4\}$; which contains every game with the different combinations of either

[iv] Nature moves first in the game to randomly choose an order of moves for the set of candidates. Moreover, to simplify the presentation of results we will assume that every order of moves has the same probability of being realised.

[v] NB. This notation hold exclusively for the case of only one region. Moreover, the notation for the case of two regions follows.

voting rules: plurality rule or proportional representation.

Timeline

In the first stage, Nature chooses the order of sequential moves for the first region and reveals the order to all candidates. Then, each candidate chooses whether to leave/quit after observing her predecessors' choices. Then, all citizens from the first region vote sincerely across the set of candidates that had chosen to stay. According to the voting rule (PR or PL), delegates are assigned to the remaining candidates. In the next stage, the same procedure is repeated in the second region. Any candidate who quits remains out in the succeeding regions. At the final stage, all the accumulated delegates from all regions elect a candidate using plurality rule; the candidate with the largest number of delegates will be the nominee (i.e., overall winner of the primary).

Furthermore, in order to simplify exposition, it is assumed that delegates from candidates who quit the race must cast their votes at the convention (i.e. in our model delegates are strictly pledged).

Voter Preferences

We continue with the notation from last chapter by writing $\theta(i; \mathcal{S})$ for the proportion of voters who top-rank candidate i if the subset of candidates $\mathcal{S} \subseteq N$ are standing. We abuse notation by writing $\theta(i; N)$ as θ_i. To simplify exposition and results, we have assumed that the proportions of votes (i.e., θ_i for candidate i) remains the same for all the regions.

In order to rule out payoff ties, we assume that for any two remaining candidates, their proportion of votes cannot be the same.

Assumption 3.1. *For any two candidates, i, i' in any subset S:*

$$\forall S \subseteq N \text{ and } i, i' \in S : \theta(i; S) \neq \theta(i'; S)$$

In our next assumption we have that the benefit of holding office, for every candidate, is

large enough to motivate staying in every region.

Assumption 3.2. *For every candidate $i \in N$:*

$$b_i > \sum_{r \in R} c_i^r$$

Following our definitions from the previous chapter, a k-winner is the candidate who in any k-wise vote has the largest share of the votes. Conversely, a k-loser is the candidate who in any k-wise vote has a smallest share of the votes. Finally, the n-winner will be the candidate with the largest share of the votes, if all candidates were to stay. The reader should recall that there is at most one k-winner for every $k > 1$ (The proof can be found in Proposition 1 from the Chapter 2). Moreover, for n=3, the 2-winner exists if and only if the 2-loser exists. We will then use the following labels to characterise our results,

Definition 2. *For every $k \in \{2, \ldots, n\}$:*

- *Candidate i is the $k-$winner if* $i = \arg\max_{j \in \mathcal{S}} \theta(j; \mathcal{S}) : \forall \mathcal{S} \subseteq N : |\mathcal{S}| = k$ *and* $i \in S$
- *Candidate i is the $k-$loser if* $i = \arg\min_{j \in \mathcal{S}} \theta(j; \mathcal{S}) : \forall \mathcal{S} \subseteq N : |\mathcal{S}| = k$ *and* $i \in S$

The reader must note that Definition 2 and Assumption 3.1 imply that k-winners and k-losers are unique. To see this, fix any k and suppose there are two k-losers. If they face each other in a two-way election then by Assumption 3.1 one of the candidates must have a smaller proportion of the votes and this candidate must indeed be the k-loser. A similar argument can be made to show the uniqueness of k-winners. Moreover, a candidate might now win for any k but nevertheless not be a k-loser. But this distinction is not relevant for our results.

The results will be categorised in next two sections. In Section 3.4, we will consider the case of one region (i.e., simultaneous election) as a benchmark. Then in Section 3.5, we will exploit the dynamic richness there is with sequential voting in $r = 2$ regions.

3.3.1 List of Conditions

In this subsection, we review the conditions which will be used in the next Section.

Condition (CC)

Essentially condition (CC) specify which candidate is the k-winner for $k \in \{2,3\}$. The first inequality establishes that candidate 3 is the 3-winner. Next, the the last three inequalities establish which candidate has the biggest vote share in every pair-wise competition:candidate 1 beats candidate 2; candidate 2 beats candidate 3; and candidate 3 beats candidate 1. The reader must note that (CC) imply a Condorcet cycle.

$$\begin{aligned} \theta_3 &> \max\{\theta_1, \theta_2\} \\ \theta(1;\{1,2\}) &> \theta(2;\{1,2\}) \\ \theta(2;\{2,3\}) &> \theta(3;\{2,3\}) \\ \theta(3;\{1,3\}) &> \theta(1;\{1,3\}) \end{aligned} \tag{CC}$$

Conditions (C.1)

$$\begin{aligned} \theta_1 &> \max\{\theta_2, \theta_3, \theta_4\} \\ \theta(3;\{2,3,4\}) &> \max\{\theta(2;\{2,3,4\}), \theta(4;\{2,3,4\})\} \\ \theta(3;\{1,3,4\}) &> \max\{\theta(1;\{1,3,4\}), \theta(4;\{1,3,4\})\} \\ \theta(3;\{1,2,3\}) &> \max\{\theta(1;\{1,2,3\}), \theta(2;\{1,2,3\})\} \\ \theta(2;\{1,2,4\}) &> \max\{\theta(1;\{1,2,4\}), \theta(4;\{1,2,4\})\} \\ \theta(1;\{1,2\}) &> \theta(2;\{1,2\}) \\ \theta(1;\{1,3\}) &> \theta(3;\{1,3\}) \\ \theta(1;\{1,4\}) &> \theta(4;\{1,4\}) \\ \theta(4;\{2,4\}) &> \theta(2;\{2,4\}) \\ \theta(4;\{3,4\}) &> \theta(3;\{3,4\}) \\ \theta(2;\{2,3\}) &> \theta(3;\{2,3\}) \end{aligned} \tag{C.1}$$

This long list of inequalities sets out which candidate is the k-winner for $k \in \{2,3,4\}$. The first inequality establishes that candidate 1 is the 4-winner. Moreover, the next four inequalities show which candidate would win in every three-way contest: candidate 3

would beat both candidates 2 and 4; candidate 3 also would beat both candidates 1 and 4, and both candidates 1 and 2. Finally, candidate 2 would beat both candidates 1 and 4. It follows that candidate 3 is the 3-winner. Additionally, since candidate 1 would lose in every three-way competition, she would be the 3-loser. Finally, the last six inequalities show which candidates would win in every pair-wise contest. Candidate 2 would beat candidate 3; candidate 4 would beat (i) candidate 2, and (ii) candidate 3. Next, since candidate 1 would beat any other candidate in every pairwise contest, she is the 2-winner. Additionally, because candidate 3 would lose in every pair-wise competition she is the 2-loser.

Conditions (C.2)

Given that the next set of conditions are used in the games of $r = 2$ region, the next pair of inequalities incorporate the number of delegates awarded in the first region (i.e., ρ_1) and the number of delegates awarded in the second region (i.e., ρ_2.) These four inequalities show which candidates would have more accumulated delegates in the second region in every pair-wise competition, given that the 3-winner (i.e., candidate 3) would get ρ_1 delegates in the first region (by the first inequality).

The second inequality establishes that candidate 2 would have more accumulated delegates if she were to face candidate 3, even though the latter already has ρ_1 delegates from region 1. The third inequality shows that candidate 3 would have more accumulated delegates than candidate 1 in a two-way race. The last inequality, implies candidate 1 would beat candidate 2 in a two-candidate race.

$$\begin{aligned} \theta_3 &> \max\{\theta_1, \theta_2\} \\ \rho_2 \cdot \theta(2; \{2,3\}) &> \rho_2 \cdot \theta(3; \{2,3\}) + \rho_1 \\ \rho_2 \cdot \theta(3; \{1,3\}) + \rho_1 &> \rho_2 \cdot \theta(1; \{1,3\}) \\ \rho_2 \cdot \theta(1; \{1,2\}) &> \rho_2 \cdot \theta(2; \{1,2\}) \end{aligned} \tag{C.2}$$

Conditions (C.3)

Just like the previous set of conditions, (C.3) is used in the games of $r = 2$ region, so it incorporates ρ_1 and ρ_2. In this case, we focus on the case of both regions having proportion voting rules (PR). Then, each candidate i would have gotten $\rho_1\theta_i$ delegates in the first region.

These three inequalities show which candidate would win (i.e., to have more accumulated delegates) in every pair-wise contest. By the first inequality, candidate 3 would win against candidate 2. Next, by the second inequality, candidate 3 would win against candidate 1. Finally, candidate 1 would win against candidate 2.

$$\begin{aligned} \rho_1\theta_3 + [\rho_2 \cdot \theta(3;\{2,3\})] &> \rho_1\theta_2 + [\rho_2 \cdot \theta(2;\{2,3\})] \\ \rho_1\theta_3 + [\rho_2 \cdot \theta(3;\{1,3\})] &> \rho_1\theta_1 + [\rho_2 \cdot \theta(1;\{1,3\})] \\ \rho_1\theta_1 + [\rho_2 \cdot \theta(1;\{1,2\})] &> \rho_1\theta_2 + [\rho_2 \cdot \theta(2;\{1,2\})] \end{aligned} \tag{C.3}$$

In Appendix 1, we confirm that the pairs of conditions
(i) $(C.1)$, (ii) $(CC), (C.2)$, and (iii) $(CC), (C.3)$ are each mutually consistent.

3.4 Results ($r = 1$ case)

In this section, we focus on the case of simultaneous elections. There are examples, such as the Conservative Party's leadership election in 2019 that can be construed as a simultaneous election (once the Conservative MP narrow the list of candidates) in the sense that Conservative party member across the United Kingdom cast their votes simultaneously.

Moreover, all the results in this subsection hold for the $r = 1$ case, unless otherwise specified. In the first two propositions, we prove the existence and uniqueness of the subgame perfect Nash equilibrium. Next, we present a result about the irrelevance of choice of voting rules in the case of simultaneous elections. Then, we discuss the effect of the identity of the different k-winners (and k-losers) on the eventual outcome of the election. Finally, we discuss the implications of Condorcet cycles.

Proposition 4.1. *There exists a unique SPNE in pure strategies for either voting rule (PR or PL) with n candidates.*

Proof. Firstly, since this game is finite (more specifically an n-candidate sequential game) we can use Zermelo's theorem (see Proposition 9.B.1 in Mas-Colell et al. (1995)) to guarantee the existence of a pure-strategy equilibrium found via backward-induction. By Assumption 3.1, uniqueness is also guaranteed by Zermelo's theorem. □

Proposition 4.2. *In equilibrium, only one candidate stays in the game.*

Proof. As a contradiction, suppose that a subset $|M| > 1$ of candidates stay such that candidate $i \in M$ is the winner. Any candidate $j \in M \setminus \{i\}$ could profitably deviate by quitting the game: if she were to stay her payoff would be $-c_j$ but if she were to quit then her payoff would be zero. □

3.4.1 Dynamics

Let us define the two possible games in the one-region case as: G_{PR} and G_{PL}. The former is the game with Proportional Representation voting rules whereas the latter is the game with Plurality voting rules. The reader must note that these two games, not just the equilibria, are equivalent because the strategies are the same (i.e., run/not run for election depending on the order of moves and on the choices of the candidates that have already moved) and the mapping from strategy combinations to payoffs is the same. Nevertheless, the following proposition works for future reference.

This result might seem surprising and controversial at first sight. However, the reader must note that our model reflects a primordial aspect of the mechanics of the primary process: only delegates (and not direct votes) are determinant in the nomination process. That is, if a candidate gets a plurality of delegates then she becomes the nominee.

Proposition 4.3 simply states that in the case of simultaneous elections (i.e., the case with $r = 1$ regions) the candidate who gets the plurality of the votes will get the plurality of the delegates and thus will become the nominee, irrespective of the voting rule.

Proposition 4.3. *The identity of the nominee (i.e. the only candidate that stays) is the same in both games G_{PR} and G_{PL}.*

Proof. We noted previously that G_{PR} and G_{PL} are equivalent and thus have the same equilibria. □

This result will be used as a benchmark in the next section, as the difference in outcomes arise only in cases with more than one region (see Proposition 5.5 in the next section).

Our next result provides sufficient conditions for one candidate to win irrespective of the order of moves.

Proposition 4.4. *If candidate i is the k-winner for every $k \in \{1, \ldots, n\}$ then she wins for every order of moves.*

Proof. By definition, this candidate would have more votes than any other remaining candidates. Then clearly she wins whenever she stays, irrespective of her opponents' decisions. Consequently she never quits. □

The next proposition shows that being the n-winner and the 2-winner does not necessarily guarantee winning for every order of moves.

Proposition 4.5. *If n=4 candidates, and (C.1) then in equilibrium the 2-winner (who is also the n-winner) does not win for some order of moves.*

Proof. It would be sufficient to show there exists an order of moves in which if the 4-winner is also the 2-winner she cannot win the game. Now consider the conditions (C.1) and fix the order of moves to be 4312. The first inequality in (C.1) implies candidate 1 is the 4-winner. Additionally, the last six inequalities describe pair-wise contests between the four candidates and these inequalities imply that candidate 1 is also the 2-winner. In the same fashion, the second, third, fourth, and fifth inequalities imply that candidate 3 is the 3-winner.

To show that candidate 1 quits in equilibrium, we begin by supposing that candidate 4 quits as a first mover. Next, by backward induction, we look at candidate 2's choices.

She stays whenever candidate 1 quits, and she quits whenever candidate 1 stays. Next, we look at candidate 1's choices: she stays regardless of candidate 3's choice because she would win. So, if candidate 4 quits: candidate 3 quits, candidate 1 stays and candidate 2 quits.

Conversely, suppose that candidate 4 stays. Now, by backward induction, we look at candidate 2's choices. She stays only when candidate 3 quits solely; that is because candidate 2 beats candidate 1 and 4 whenever whenever the three candidates run together. Next, we look at candidate 1's choices: she quits regardless of candidate 3's choice because she would lose otherwise. That is, she will lose when (i) candidates 1, 3, and 4 run together (since candidate 3 is the 3-winner), and (ii) candidates 1, 2, and 4 run together (by the fourth inequality). Next, candidate 3 loses if she stayed therefore she quits. Finally, if candidate 4 stays she wins and if she quits she loses. Therefore, in equilibrium candidate 4 stays, candidate 3 quits, candidate 1 quits and candidate 2 quits. □

We now consider games with a 2-loser. We show that such a candidate cannot win.

Proposition 4.6. *For any number of candidates and for every order of moves, if a 2-loser exists then she does not win.*

Proof. (By contradiction)

The 2-loser would win if all other candidates decided to quit. Now by Proposition 4.2, there can only be one candidate remaining in equilibrium; so we assume the 2-loser is the only remaining candidate. Now, we fix a particular order and look at the last candidate's decision. If this candidate is not the 2-loser, she will only face the 2-loser and thus could profitably deviate to stay. On the other hand, if the last candidate is the 2-loser we look at the penultimate candidate's decision. Given that the other candidates have already quit, she would win irrespective of the 2-loser's decision, so she can profitably deviate. In both situations there is a contradiction. □

In light of Proposition 4.6, one might conjecture that the presence of a 2-loser is irrelevant. We now show this conjecture is wrong.

Proposition 4.7. *Suppose n=3. The following holds in every SPNE: a 2-winner loses for some orders of moves if and only if the 2-loser is the 3-winner.*

Proof. For the sufficiency part of the proposition, we suppose that candidate 1 is the 3-winner and 2-loser. We additionally suppose that candidate 2 is the 2-winner. Now, we fix the order of moves to be 321. Then in equilibrium the 2-winner quits. By backward induction, candidate 1 will stay only if the other two candidates were both to stay (because she is the 3-winner) or both to quit. Now, candidate 2 quits if candidate 3 stays because the 3-winner (candidate 1) would have stayed if all stayed. Then, candidate 2 stays when candidate 1 quits because the 3-winner would have quit as well. Thus, candidate 3 wins in equilibrium: if she stays then she wins and if she quits then she loses.
Conversely, suppose that the 2-loser is not the 3-winner. Then there are two cases: either the 3-winner is the 2-winner or the 3-winner is neither the 2-winner nor the 2-loser. By Proposition 4.4, in the former case the 2-winner always wins. In the latter case, the 2-loser would always quit in equilibrium (by Proposition 4.6). Then, the 3-winner would quit because the 2-winner stays and win. □

Proposition 4.7 implies that if n=3 then a 2-winner wins the nomination for every order of moves if the 2-loser is not the 3-winner.

3.4.2 Condorcet cycles

Thus far, we have focused on games in which there is a 2-loser. We now turn to games with a Condorcet cycle, where there is no 2-loser. This section will present results that will be crucial in the results with $r = 2$ regions.

Suppose that (i) $N = \{1, 2, 3\}$, (ii) there is a Condorcet cycle, and (iii) candidate 3 is the

3-winner. Then, this is the resulting parametric configuration:

$$\begin{aligned} \theta_3 &> \max\{\theta_1, \theta_2\} \\ \theta(1;\{1,2\}) &> \theta(2;\{1,2\}) \\ \theta(2;\{2,3\}) &> \theta(3;\{2,3\}) \\ \theta(3;\{1,3\}) &> \theta(1;\{1,3\}) \end{aligned} \tag{CC}$$

This set of parameters, (CC), describes a scenario with a cyclical nature. By the second inequality, if candidate 3 were to quit then candidate 1 would receive a larger vote share than candidate 2. By the third inequality, if candidate 1 quits then candidate 2 would receive more votes than candidate 3. Finally by the last inequality, if candidate 2 were to quit then candidate 3 would receive more votes than candidate 1. Moreover, candidate 3 would receive more votes if all candidates stayed by the first inequality.

Proposition 4.8. *Suppose that (i) $N = \{1, 2, 3\}$, (ii) there is a Condorcet cycle, and (iii) candidate 3 is the 3-winner. The following holds in every SPNE: every candidate wins the game for at least one order of moves.*

Proof. In equilibrium, candidate 3 loses as first mover (orders 312 and 321) because staying would induce 1 to quit and 2 to stay and beat 3; so 3 quits and 1 wins. Now, whenever candidate 1's turn is before candidate 3 (orders 123, 132 and 213) the winner would be candidate 2, because the presence of candidate 3 induces 1 to quit. Finally, candidate 3 wins only in order 231 because if candidate 2 stays then candidate 3 quits, and 1 stays; so candidate 2 quits. □

In other words, if the electorate preferences are reflected by (CC) then every candidate has a strictly positive probability of winning the nomination.

Order of Moves	Candidate 1	Candidate 2	Candidate 3
123	Quits	**Stays**	Quits
132	Quits	**Stays**	Quits
213	Quits	**Stays**	Quits
231	Quits	Quits	**Stays**
312	**Stays**	Quits	Quits
321	**Stays**	Quits	Quits

Now we have a clear link between the (non-)existence of a 2-winner/2-loser and the orders of moves in which each candidates wins; and thus the probability of winning any given region (whenever n=3). Essentially, there are only three different scenarios. If there is no 2-loser then each candidate wins for some order. If there is a 2-loser who is the 3-winner then the 2-winner loses for some orders. Finally, if the 2-loser is not the 3-winner then the 2-winner must win the nomination.

3.5 Results (r=2 case)

In the previous section we presented results for the benchmark case of only one region. However, these results do not capture the dynamic properties of sequential primaries. The reader must note that all the results in this subsection hold for the $r = 2$ case, unless otherwise specified.

The definition of winning must be clarified. In the case of one region, when a candidate wins the election that is synonymous with getting the largest number of delegates and thus becoming the nominee. However, in the case of two regions, a candidate could win (i.e., get the largest number of delegates from the first region) in the first region but not get the largest amount of delegates overall and not become the nominee. In order to become the nominee, a candidate needs to get the largest total number of delegates from the two regions combined. This particular distinction will be essential in proving the results of Section 3.6

As we will see in our results, the differences between voting rules arise only in the cases

of more than one region. Given that our main results are presented assuming only two regions, we focus in this particular case.

Definition 3. *For r=2 regions, there exist four games* $G_1, G_2, G_3, G_4 \in \mathcal{G}^2$

Game	*Voting rule in Region 1*	*Voting rule in Region 2*
G_1	PR	PR
G_2	PR	PL
G_3	PL	PR
G_4	PL	PL

Firstly, we present existence and uniqueness results which generalise Proposition 4.1 and 4.2 from the previous section. The number of candidates remaining at any region is an important feature of this model; so, in Proposition 5.3, we rule out the case of having exactly two candidates at any region. Next, Proposition 5.4 generalises the result found in Proposition 4.4, above, to cases with any number of regions. Finally, Proposition 5.5 provides two important takeaways from this section (and indeed from this chapter): (i) the differences between voting rules (PR and PL) arise only in the case of games with more than one region, and (ii) we find conditions under which the momentum effects are not effective.

Proposition 5.1. *There exists a unique SPNE, for either voting rule.*

Proof. As in Proposition 4.1, this result follows from Zermelo's theorem, as it would still be a finite game with perfect information. □

The same argument made in the proof of Proposition 5.1 can be generalised to any number of regions. An interesting characteristic of all the equilibria analysed in this chapter is that at the last region only one candidate can remain. This following result mirrors Proposition 4.2.

Proposition 5.2. *In equilibrium, only one candidate stays in the second region.*

Proof. Let a subset $|M| > 1$ of candidates stay until the second region such that candidate $i \in M$ wins. Now take any candidate $j \in M \setminus \{i\}$: she could profitably deviate by quitting

the game beforehand (if she were to stay her payoff would be $-c_j^2$ but if she were to quit her payoff would be zero). □

It is worth noting that at any given region, no two candidates can remain on the equilibrium path. All the cases where three or more candidates remain will not be excluded. In other words, either just one candidate can remain on the race at any given region, or any number of candidates strictly greater than two. This is because one of the two candidates left in the race would win, irrespective of the order.

Proposition 5.3. *For any region r there cannot be exactly two remaining candidates.*

Proof. As a contradiction, fix any r, any voting rule and suppose that exactly two candidates decide to stay. Then, for either of the two orders of moves, the same candidate would win and thus would always stay. The other candidate could clearly deviate and quit, which is a contradiction. □

The next result is a generalisation of Proposition 4.4.

Proposition 5.4. *Any candidate who is a k-winner for every k, stays in the second region in every SPNE, irrespective of the voting rule in each region.*

Proof. The proof works with the same line of argument as in the proof of Proposition 4.4. This specific candidate will win whenever she stays in the race, irrespective of the identity or the number of candidates that quit before or after her. That is because she will win against any pair-wise competition, and races where she might face any number of rivals. □

3.6 Momentum ($r = 2, n = 3$ case)

The main idea behind momentum is the intuitive notion that performing well at the initial stages (Iowa, New Hampshire, etc) is crucial for any candidate, even for strong, well-funded front-runners. We would like to challenge that notion by simply asking: when is it disadvantageous to win in the firsts regions? In this section we look at conditions under

which a front-runner, namely a 3-winner, would amass the largest proportion of delegates from the first region but eventually "fails" to become the nominee. More specifically, we provide conditions under which an 3-winner would have the lowest probabilities of becoming the nominee, even though she has won all the delegates from the first region.

In our model, we say a candidate experiences "momentum-killing effects" whenever she gets all the delegates from the first region but there is at least one order of moves in the second region in which she quits.

In the following striking result, found in Proposition 5.5, there is an order of moves in the second region in which the 3-winner (i.e., the front-runner) does not become the nominee even after getting all the delegates from the first region. Again, the reader must note that all the results in this subsection hold for the $r = 2$, and $n = 3$ case exclusively.

Definition 4 (Momentum-killing effects)**.** *A candidate experiences "momentum-killing effects" whenever she gets all delegates from the first region (ρ_1) but there is an order of moves in the second region in which she quits (i.e. does not become the nominee).*

Consider the following conditions:

$$\begin{aligned} \theta_3 &> \max\{\theta_1, \theta_2\} \\ \rho_2 \cdot \theta(2; \{2,3\}) &> \rho_2 \cdot \theta(3; \{2,3\}) + \rho_1 \\ \rho_2 \cdot \theta(3; \{1,3\}) + \rho_1 &> \rho_2 \cdot \theta(1; \{1,3\}) \\ \rho_2 \cdot \theta(1; \{1,2\}) &> \rho_2 \cdot \theta(2; \{1,2\}) \end{aligned} \tag{C.2}$$

Proposition 5.5. *In every SPNE of G_3 and G_4, the 3-winner experiences "momentum-killing effects", provided the following conditions hold: (CC), (C.2), and $\rho_2 > \rho_1$.*

Proof. In G_4, the winner of the second region would get the majority of the delegates needed to win the game. Given this and the fact that by (CC) the winner is determined by the order of moves, all candidates stay in the first region. Then the 3-winner would get all ρ_1 delegates. However, the 3-winner would only stay in the second region and become the nominee in one out of the possible six orders of moves (see Proposition 4.8).

In G_3, if all candidates were to stay in the first region, candidate 3 gets all the delegates from the first region (i.e., ρ_1) but it is not enough to get the nomination. Now, given (C.2) we know which candidates would win in each pair-wise contest. Now, we need to examine each order of moves to identify the winner. Candidate 1 wins whenever candidate 3 is the first mover (orders 312 and 321) because if 3 were to stay 1 would have to quit, so 3 quits and 1 wins. Next, candidate 2 will win whenever candidate 1's turn is before candidate 3 (orders 123, 132 and 213), because the presence of candidate 3 induces 1 to quit. Finally, candidate 3 wins only in order 231 because if candidate 2 stays induces candidate 3 to quit, so candidate 2 quits. Since the order of moves in the second region determines the winner, all candidates decide to stay for the first region (even though candidate 3 gets all ρ_1 delegates). □

3.7 Modified Duverger's Law ($r = 2, n = 3$ case)

The critical matter in Duverger's law is the number of political parties in a party system. However, Duverger never meant to analyse dynamic settings such as sequential primary elections; in that case, the analogous concept is the number of candidates in a primary contest. Thus a modified Duverger's Law version for primary elections would assert that primary candidates tend to quit at earlier stages of the election at a faster rate under plurality rule compared to proportional representation. Again, the reader must note that all the results in this subsection hold for the $r = 2$, and $n = 3$ case exclusively.

Definition 5 (Duvergerian Property)**.** *The Duvergerian Property holds for the ordered pair of games (G_1, G_4) whenever the number of candidates remaining in the first region prescribed by every SPNE of G_1 (i.e. the game where the voting rule is PR for both regions) is strictly greater than the number of candidates remaining in the first region prescribed by every SPNE of G_4 (i.e. the game where the voting rule is PL for both regions).*

Our next result challenges this dynamic version of Duverger's Law. Consider the following

conditions:

$$\begin{aligned}
\rho_1\theta_3 + [\rho_2 \cdot \theta(3;\{2,3\})] &> \rho_1\theta_2 + [\rho_2 \cdot \theta(2;\{2,3\})] \\
\rho_1\theta_3 + [\rho_2 \cdot \theta(3;\{1,3\})] &> \rho_1\theta_1 + [\rho_2 \cdot \theta(1;\{1,3\})] \\
\rho_1\theta_1 + [\rho_2 \cdot \theta(1;\{1,2\})] &> \rho_1\theta_2 + [\rho_2 \cdot \theta(2;\{1,2\})]
\end{aligned} \tag{C.3}$$

Proposition 5.6. *The Duvergerian Property does not hold for (G_1, G_4) provided the following conditions hold: (CC), (C.3), and $\rho_2 > \rho_1$.*

Proof. Firstly, we will show that no SPNE in G_1 prescribes all three candidates to stay in the first region. To prove this, as a contradiction we will assume that all candidates stay. Each candidate i would then get $\rho_1\theta_i$ delegates in the first region. Because of the difference in delegate size, the winner of the first region would not get enough delegates to be nominated. In order to determine the winner in every subgame, we need to determine who would win in every k-wise contest. For $k = 3$, the 3-winner (i.e., candidate 3) would win the largest share of delegates. Then, we look at every pairwise comparison to determine who would win whenever there would be two remaining candidates in the second region.

Firstly, if candidate 1 were to quit then candidate 3 ends up with more delegates than candidate 2 given the first inequality in (C.3). In this case, delegates awarded to candidate 3 are composed of the following: (i)$\rho_1\theta_3$ delegates inherited from the first region, (ii) $\rho_2 \cdot \theta(3;\{2,3\})$ delegates awarded from voters that top-rank her, and voters that top-ranked candidate 1 but second-ranked her.

Secondly, if candidate 2 were to quit then candidate 3 would get more delegates than candidate 1 given the second inequality in (C.3). Finally, if candidate 3 were to quit then candidate 1 would end up with more delegates than candidate 2 given the third inequality in (C.3).

If these three inequalities hold, then candidate 3 has more delegates in every pairwise contest and whenever the three candidates stay. Now, via backward induction let us look at what happens in equilibrium at every order of moves in the second region.

- If the order of moves is 123: 1 quits, 2 quits, and 3 stays
- If the order of moves is 132: 1 quits, 3 stays, and 2 quits
- If the order of moves is 213: 2 quits, 1 quits, and 3 stays
- If the order of moves is 231: 2 quits, 3 stays, and 1 quits
- If the order of moves is 312: 3 stays, 1 quits, and 2 quits
- If the order of moves is 321: 3 stays, 2 quits, and 1 quits.

Given this, if all candidates were to stay in the first region only candidate 3 would stay in the second region for every order of moves. Because all candidates can anticipate this, candidates 1 and 2 would quit in the first region. This is a contradiction. Therefore only one candidate can stay in the first region.

On the other hand, in G_4 the winner of the first region is irrelevant (because $\rho_2 > \rho_1$); and only the result from the second region will be decisive. Thus all candidates will stay in the first region, and the 3-winner gets ρ_1 delegates. Finally, we will find the winner of the nomination for each order of moves in the second region. (CC) implies that candidate 1 wins whenever candidate 3 is the first mover (orders 312 and 321) because if 3 were to stay then 1 would have to quit; so 3 quits and 1 wins. Next, candidate 2 wins whenever candidate 1 moves before candidate 3 (orders 123, 132 and 213), because the presence of candidate 3 induces 1 to quit. Finally, candidate 3 only wins in order 231 because candidate 2 staying induces candidate 3 to quit; so candidate 2 quits.

Thus, every candidate stays in the first region, and the 3-winner gets all ρ_1 delegates from the first region. Then, by (CC), all candidates win for at least one order of moves in the second region. □

The main lesson from this subsection is that candidates may drop out earlier in sequential elections with PR rules. If we identify PL and PR respectively with the Republicans and Democrats then this lesson can be translated directly into a testable implication and the primary results from the previous decades can be used to explore it.

3.8 Spatial Setting

In this section we will (re)introduce the spatial setting presented in Chapter 2 to check which results generalise to the additional spatial restrictions. For that, we will firstly (re)state the spatial setting. We then exhaustively look at every result and set of conditions from Sections 3.4 and 3.5, and show which would be (in)consistent with the spatial setting (see Corollaries 2 and 3).

The results in this section could be interpreted as unsurprising or even trivial. In fact, we face the inescapable consequences of the Median Voter Theorem: the fact that the Condorcet-winner (which is the candidate preferred by the median voter) is the undisputed winner. However, this constitutes an additional contribution of this present chapter. We have shown in Sections 3.6 and 3.7 that our two main (negative) results (i.e., Propositions 5.5 and 5.6) would hold in a non-spatial setting. Then, in Corollary 3, we conclude these results cannot be extended to the spatial setting. Therefore, the phenomena described in Propositions 5.5 and 5.6 can only occur in a non-spatial setting.

Preliminaries

The policy space is the unit interval, and a policy is defined as a point in $[0, 1]$. There are n candidates represented by their ideal policy in the unit interval; and these candidates can be ordered as $0 \leqslant a_1 < a_2 < \ldots < a_N \leqslant 1$. There is a continuum of voters, indexed by j, whose preferences over alternatives is represented by the loss function $-(a - a^j)^2$; a^j is voter j's *ideal alternative in* $[0, 1]$. That is, each voter would prefer the remaining candidate who is closest to the voter's ideal policy. In order to simply the exposition of our main results, we assume that a^j is uniformly distributed on $[0, 1]$. The *median* voter is then the voter whose ideal alternative on $[0, 1]$ is $1/2$.

For any collection $K \subseteq N$, we say that the lowest and highest alternatives in K are *exterior to* K, writing the former as $a_{\underline{K}}$ and the latter as $a_{\overline{K}}$; and we say that any other alternatives are *interior to* K.

Each voter's loss function determines an ordering over the alternatives in any subset

$K \subseteq N$. Write $a^j(K)$ for voter j's preferred alternative on K. We will refer to $a^{1/2}(N)$, the median voter's preferred alternative in N, as the *median alternative*. It will prove convenient to write $a^{1/2}(N)$ as a_m. Take any nonempty subset of N, say K. For any $a \in K$, write $\theta(a; K)$ as the proportion of voters j such that $a = a^j(K)$. We say that a is a *winner on* K if

$$a \in \arg\max_{\alpha \in K} \theta(\alpha; K),$$

and that a is a *loser on* K if

$$a \in \arg\min_{\alpha \in K} \theta(\alpha; K).$$

We now expand this notion to every collection of alternatives with k elements.

Definition 6. *We say that alternative a is a k-winner if alternative a is a winner on K for every nonempty $K \subseteq N$ with k elements; we define a k-loser analogously.*

In this subsection, we must consider two relevant features that arise by the introduction of a spatial structure: (i) candidates are now classified as either interior to N or exterior to N, and (ii) both the 2-winner and the 2-loser always exist. Therefore any type of electoral cycles are precluded. Additionally, there is another interesting and important feature in the case of three candidates: the identity of the 2-loser (see Lemma 3 below.)

Quite importantly, the reader must note that Propositions 4.1, 4.2, 4.3, 4.4, 4.6, 5.1, 5.2, 5.3, and 5.4 (all of which cover existence, uniqueness, and winner-identity) do not rely on preference restrictions. Therefore also hold in the spatial setting. In the next two subsections we will show that the following results are not consistent with the spatial setting: Propositions 4.5, 4.7, and 4.8; as well as Propositions 5.5, and 5.6.

3.8.1 Results with $r = 1, n = 3$

In this subsection, in Proposition 6.1, we characterise the equilibrium with one region and three candidates. Lemma 3 will help to prove Proposition 6.1. Furthermore, Proposition 6.2 generalises this result to the case of n candidates.

Lemma 3. *The identity of the 2-loser is not the same as the 3-winner.*

Proof. As a contradiction, assume the 3-winner is the 2-loser. First, we can trivially rule out the 3-winner to be the 2-winner (because she is also the 2-loser). Next, We examine two cases:

- Assume that the 2-winner is a candidate exterior to N. Then, by Proposition 3 in Chapter 2, she must also be the 3-winner. A contradiction.
- Assume that the 2-winner is a candidate interior to N. Then, the 3-winner must be a candidate exterior to N. Moreover, whenever the 3-winner is running with the other candidate exterior to N, the 3-winner must have the majority of the votes. Thus, she cannot be the 2-loser. □

In light of Lemma 3, we can rule out Proposition 4.7 to be consistent with the Spatial Setting.

Proposition 6.1. *In every SPNE, the 2-winner is the only candidate that stays, irrespective of the voting rule and the order of moves.*

Proof. By Proposition 4.3, in the case of only one region, the identity of the winner is the same under PR than under PL. Now, we need to show that the winner is indeed the 2-winner. We look at two cases:

1. Suppose that the 2-winner is also the 3-winner. Then she wins by Proposition 4.6.
2. Suppose that the 2-winner is not the 3-winner. Then by Proposition 4.7, as long as the 3-winner is the 2-loser the 2-winner loses for some order of moves. This possibility is precluded by Lemma 3. Thus, the 2-winner wins. □

In Proposition 6.1, we have proved that for every order of moves the 2-winner is the only candidate that stays. In that light, Proposition 4.8 is not consistent with the spatial setting. Additionally, this proposition is inconsistent under the spatial setting because the results rely on the existence of electoral cycles. In other words, conditions such as (CC) are invalid in a spatial setting.

In the next result, we present a generalisation of Proposition 6.1.

Proposition 6.2. *Irrespective of (i) the voting rule, (ii) the number of candidates, and (iii) the order of moves: the 2-winner is the only candidate that stays in every SPNE of the spatial setting.*

Proof. By Proposition 4.6, the 2-loser quits in every SPNE. Now, we focus on the subset of $n-1$ other candidates where the identity of the 2-winner clearly remains the same. Then, there must be a 2-loser (label this candidate as ℓ_{n-1}) with respect of this subset of $n-1$ candidates.

Clearly, candidate ℓ_{n-1} would only win if all other candidates were to quit. By Proposition 4.2, there can only be one candidate remaining in equilibrium. As a contradiction, assume the 2-loser is the only candidate that stays. Next, fix a particular order and look at the last candidate's decision. If this candidate is not the 2-loser, she will only face the 2-loser and thus could profitably deviate to stay. On the other hand, if the last candidate is the 2-loser we look at the penultimate candidate's decision. Given that the other candidates have already quit, she would win irrespective of the 2-loser's decision, so she can profitably deviate. In both situations there is a contradiction. Therefore, candidate l_{n-1} does not stay in every SPNE.

After removing candidate l_{n-1}, we focus on the subset of $n-2$ other candidates where the identity of the 2-winner clearly remains the same. Then, there must be a 2-loser (label her as candidate ℓ_{n-2}) with respect of this subset of $n-2$ candidates. Following the same argument from the previous paragraph, candidate ℓ_{n-2} does not stay in every SPNE.

We repeat the same argument until we reach the subset of two candidates, where candidate ℓ_2 would lose to the 2-winner and thus she does not stay in every SPNE. □

An crucial consequence of Proposition 6.2 is that Proposition 4.5 is render inconsistent with the spatial setting. That is, the 2-winner must win for every order of moves in the spatial setting for any number of candidates.

Moreover, Proposition 4 in Chapter 2 implies that a candidate exterior to N who is a 2-winner (e.g., candidate 1 in Proposition 4.5) must also be a k-winner for any $k > 2$. Additionally, Proposition 4.4 would then imply that the 2-winner wins for every order.

In conclusion, if a 2-winner is also the n-winner then she is the k-winner for any $k > 1$ and thus she wins for every order of moves for any number of candidates.

The punchline of this subsection is formalised in the corollary below.

Corollary 2. *Propositions 4.5, 4.7, and 4.8 are inconsistent in the spatial setting.*

3.8.2 Results with $r = 2, n = 3$

In this subsection, we characterise the equilibrium with two region and three candidates. In Proposition 6.3, we show that in equilibrium, only one candidate stays in the first region. This result rules out any game to have the Duvergerian Property in the spatial setting and thus precludes Proposition 5.6. Moreover, it also rules out any candidate to experience "momentum-killing effects" in the spatial setting and thus also precludes Proposition 5.5. Furthermore, in Proposition 6.4 we characterise the equilibrium for every game with two regions and three candidates.

Proposition 6.3. *In every SPNE of games G_1, G_2, G_3, and G_4 only one candidate stays in the first region.*

Proof. Firstly, we show that no SPNE prescribes all three candidates to stay in the first region. To prove this, as a contradiction we will assume that all candidates stay in the first region.

- In G_1 and G_2, each candidate i would then get $\rho_1\theta_i$ delegates in the first region. In the case of $\rho_1 > \rho_2$, the 3-winner would get most delegates and therefore the other candidates can profitably deviate and quit in the first region. A contradiction.

  In the case of $\rho_2 > \rho_1$, the winner of the first region would not get enough delegates to be nominated. In equilibrium, regardless of the identity of the candidate with the largest number of delegates from the second region, the 2-loser would profitably deviate and quit in the first region. Given this, the 3-winner cannot beat the 2-winner. Then, she also deviates and quits in the first region. A contradiction.

- In G_3 and $G4$, the 3-winner would get ρ_1 delegates in the first region. In the case of $\rho_1 > \rho_2$, the 3-winner would get most delegates (i.e., ρ_1) and therefore the other

candidates can profitably deviate and quit in the first region. A contradiction.

In the case of $\rho_2 > \rho_1$, the winner of the first region would not get enough delegates to be nominated. In equilibrium, regardless of the identity of the candidate with the largest number of delegates from the second region, the 2-loser would profitably deviate and quit in the first region. Given this, the 3-winner cannot beat the 2-winner. Then, she also deviates and quits in the first region. A contradiction.

□

Proposition 6.4. *In every SPNE of games G_1, G_2, G_3, G_4, the 2-winner becomes the nominee, for every order of moves.*

Proof. The 2-loser could only get all delegates in the first region (and then become nominee) if and only if the other two candidates were to quit in the first region. We suppose by contradiction that only the 2-loser stays in the first region. Then, by using the same argument found in the proof of Propositions 4.6 and 6.2, any other candidate can profitably deviate and stay; thus, in equilibrium, the 2-loser quits.

The next possible winner in the first region is the 3-winner, who by Lemma 1 is not the same candidate as the 2-loser. The 3-winner would only win in the first region if and only if (i) all candidates were to stay or (ii) all other candidates were to quit.

In the first case, if all three candidates were to stay then the 3-winner would win. In equilibrium, however, the 2-loser would quit, and so the 3-winner would also quit. The 2-winner is the only candidate that stays in equilibrium.

For the second case, suppose as a contradiction, that only the 3-winner stays. Then, the 2-winner could profitably deviate and stay and so the 3-winner must quit in equilibrium. □

Additionally, since Propositions 5.5 and 5.6 rely on Condition (CC), the result in both must be inconsistent with a spatial setting with two regions and three candidates. Ultimately, Proposition 6.1 precludes Proposition 5.5; and Corollary 3 precludes Proposition 5.6.

Corollary 3. *Propositions 5.5 and 5.6 are inconsistent in the spatial setting.*

3.9 Concluding remarks

The American Presidential primary system is both unique and quite complex. Its dynamic richness makes it a fertile focus of research. There are three distinctive features of such elections: (i) heterogeneous voting rules, (ii) sequential order and (iii) delegation size. In this chapter, we use a non-cooperative model to analyse the implications of these features.

The main contribution of this chapter is threefold. Firstly, we show how some of the benchmark results from the one-region case (simultaneous election) disappear in cases with more than one region (sequential election). For example, Proposition 5.5 indicates that a 2-loser can win the nomination with some positive probability in the sequential case, whereas in a simultaneous case, a 2-loser cannot win according to Proposition 4.6. Secondly, we look at cases where the so-called "momentum effects" fail to materialise, and more specifically we look at cases where a candidate experiences "momentum-killing" effects. That is, in the case of two regions, we derive conditions under which a front-runner (i.e., an n-winner) gets all the delegates from the first region but there is an order of moves in the second region in which she quits and fails to become the nominee.

Thirdly, we study a sequential election version of Duverger's law. This version asserts that the attrition rate amongst candidates is strictly higher/faster in games with plurality rule compared to games with proportional rules. We derive conditions to challenge this Duvergerian property. More specifically, we find conditions under which more candidates drop out in the first region under proportional representation that under plurality rule.

Furthermore, as a minor contribution, we have shown our main results (i.e., Propositions 5.5 and 5.6) cannot be extended to the spatial setting. Both of these negative results can only occur in a non-spatial setting.

A natural extension to our work is to translate these three takeaways into testable implications. The econometric challenge, however, is not the collection of data but the inconsistency of the data. There is electoral data from the last two decades for each state's primary; but both the order and voting rule (and even in some case delegation size) vary through time. Moreover, the number of candidates is always different for each

party and each state.

3.10 References

Adams, J. & Merrill, S. (2008). Candidate and Party Strategies in Two-Stage Elections Beginning with a Primary. *American Journal of Political Science*, *52*(2), 344–359.

Agranov, M. (2016). Flip-Flopping, Primary Visibility, and the Selection of Candidates. *American Economic Journal: Microeconomics*, *8*(2), 61–85.

Ansolabehere, S. & King, G. (1990). Measuring the Consequences of Delegate Selection Rules in Presidential Nominations. *The Journal of Politics*, *52*(2), 609–621.

Atkeson, L. R. & Maestas, C. (2008). Racing to the Front: The Effect of Front-Loading on presidential primary turnout. In *University of Iowa Typescript*.

Battaglini, M. (2005). Sequential Voting with Abstention. *Games and Economic Behavior*, *51*(2), 445–463. Special Issue in Honor of Richard D. McKelvey. doi:https://doi.org/10.1016/j.geb.2004.06.007

Bowler, S., Grofman, B., & Blais, A. (2009). *Duverger's Law of Plurality Voting: The Logic of Party Competition in Canada, India, the United Kingdom and the United States*. Springer.

Brady, D. W., Han, H., & Pope, J. C. (2007). Primary Elections and Candidate Ideology: Out of Step with the Primary Electorate? *Legislative Studies Quarterly*, *32*(1), 79–105.

Cherry, T. L. & Kroll, S. (2003). Crashing the Party: An Experimental Investigation of Strategic Voting in Primary Elections. *Public Choice*, *114*(3), 387–420.

Cooper, A. L. (2002). The Effective Length of the Presidential Primary Season: The Impact of Delegate Allocation Rules and Voter Preferences. *Journal of Theoretical Politics*, *14*(1), 71–92.

Damore, D. F., Hansford, T. G., & Barghothi, A. (2010). Explaining the Decision to Withdraw from a U.S. Presidential Nomination Campaign. *Political Behavior*, *32*(2), 157–180.

Davis, J. W. (1997). *U.S. Presidential Primaries and the Caucus-Convention System: A Sourcebook*. Greenwood Publishing Group.

Dekel, E. & Piccione, M. (2000). Sequential Voting Procedures in Symmetric Binary Elections. *Journal of Political Economy*, *108*(1), 34–55.

Deltas, G., Herrera, H., & Polborn, M. K. (2015). Learning and Coordination in the Presidential Primary System. *The Review of Economic Studies*, *83*(4), 1544–1578. doi:10.1093/restud/rdv055

DeSilver, D. (2016). Turnout Was High in the 2016 Primary Season, But Just Short of 2008 Record. Retrieved June 10, 2019, from https://www.pewresearch.org/fact-tank/2016/06/10/turnout-was-high-in-the-2016-primary-season-but-just-short-of-2008-record/

Diwakar, R. (2007). Duverger's Law and the Size of the Indian Party System. *Party Politics*, *13*(5), 539–561.

Duverger, M. (1951). Les Partis Politiques. A. Colin, Paris.

Gaines, B. J. (1999). Duverger's Law and the Meaning of Canadian Exceptionalism. *Comparative Political Studies*, *32*(7), 835–861.

Gerber, E. R. & Morton, R. B. (1998). Primary Election Systems and Representation. *Journal of Law, Economics, & Organization*, 304–324.

Hammond, T. H. (1980). Another Look at the Role of "The Rules" in the 1972 Democratic Presidential Primaries. *Western Political Quarterly*, *33*(1), 50–72.

Haynes, A. A., Gurian, P.-H., Crespin, M. H., & Zorn, C. (2004). The Calculus of Concession: Media Coverage and the Dynamics of Winnowing in Presidential Nominations. *American Politics Research*, *32*(3), 310–337.

Hortala-Vallve, R. & Esteve-Volart, B. (2011). Voter Turnout and Electoral Competition in a Multidimensional Policy Space. *European Journal of Political Economy*, *27*(2), 376–384.

Hummel, P. & Holden, R. (2014). Optimal primaries. *Journal of Public Economics*, *109*, 64–75.

Ishihara, A. & Miura, S. (2017). Minor Candidates as Kingmakers. *Public Choice*, *170*(3), 253–263. doi:10.1007/s11127-016-0393-9

Kamarck, E. (2018). *Primary politics: Everything You Need to Know About How America nominates its Presidential Candidates*. Brookings Institution Press.

King, G. (1990). Electoral Responsiveness and Partisan Bias in Multiparty Democracies. *Legislative Studies Quarterly*, 159–181.

Klumpp, T. & Polborn, M. K. (2006). Primaries and the New Hampshire Effect. *Journal of Public Economics*, *90*(6), 1073–1114.

Knight, B. & Schiff, N. (2010). Momentum and Social Learning in Presidential Primaries. *Journal of Political Economy*, *118*(6), 1110–1150.

Lengle, J. I. & Shafer, B. (1976). Primary Rules, Political Power, and Social Change. *American Political Science Review*, *70*(1), 25–40. doi:10.1017/S000305540026396X

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory*. Oxford University Press.

Miller Jr, R. G. (2011). *Survival Analysis*. John Wiley & Sons.

Norrander, B. (2006). The Attrition Game: Initial Resources, Initial Contests and the Exit of Candidates during the U.S. Presidential Primary Season. *British Journal of Political Science*, *36*(3), 487–507.

Potthoff, R. F. (2011). Condorcet Polling. *Public Choice*, *148*(1), 67–86. doi:10.1007/s11127-010-9646-1

Reed, S. R. (1990). Structure and Behaviour: Extending Duverger's Law to the Japanese Case. *British Journal of Political Science*, *20*(3), 335–356.

Riker, W. H. (1982). The Two-Party System and Duverger's Law: An Essay on the History of Political Science. *American Political Science Review*, *76*(4), 753–766.

Rogowski, J. C. & Langella, S. (2015). Primary Systems and Candidate Ideology: Evidence from Federal and State Legislative Elections. *American Politics Research*, *43*(5), 846–871.

Scarrow, H. A. (1986). Duverger's Law, Fusion, and the Decline of American Third Parties. *Western Political Quarterly*, *39*(4), 634–647.

Schwabe, R. (2015). Super Tuesday: Campaign Finance and the Dynamics of Sequential Elections. *Social Choice and Welfare*, *44*(4), 927–951.

Selman, D. et al. (2010). Optimal Sequencing of Presidential Primaries. *Bogaziçi University Typescript*.

Serra, G. (2011). Why Primaries? The Party's Tradeoff between Policy and Valence. *Journal of Theoretical Politics*, *23*(1), 21–51.

Steger, W. P., Hickman, J., & Yohn, K. (2002). Candidate Competition and Attrition in Presidential Primaries, 1912-2000. *American Politics Research*, *30*(5), 528–554.

Washington, G. (1796). Farewell Address. *September 19*.

Appendix 1

Numerical Verification

Proposition 4.5. *Consider the following set of inequalities. We verify the set of inequalities is consistent by providing a numerical example:*

$\theta(1;\{1,2,3\}) = 0.34$	$\theta(2;\{1,2,3\}) = 0.314$	$\theta(3;\{1,2,3\}) = 0.346$	$\theta_1 = 0.289$
$\theta(1;\{1,2,4\}) = 0.35$	$\theta(2;\{1,2,4\}) = 0.362$	$\theta(4;\{1,2,4\}) = 0.288$	$\theta_2 = 0.192$
$\theta(1;\{1,3,4\}) = 0.372$	$\theta(3;\{1,3,4\}) = 0.373$	$\theta(4;\{1,3,4\}) = 0.255$	$\theta_3 = 0.288$
$\theta(2;\{2,3,4\}) = 0.27$	$\theta(3;\{2,3,4\}) = 0.371$	$\theta(4;\{2,3,4\}) = 0.359$	$\theta_4 = 0.231$
$\theta(1;\{1,2\}) = 0.515$	$\theta(2;\{1,2\}) = 0.485$	$\theta(1;\{1,4\}) = 0.515$	$\theta(4;\{1,4\}) = 0.485$
$\theta(1;\{1,3\}) = 0.513$	$\theta(3;\{1,3\}) = 0.487$	$\theta(2;\{2,3\}) = 0.516$	$\theta(3;\{2,3\}) = 0.484$
$\theta(2;\{2,4\}) = 0.436$	$\theta(4;\{2,4\}) = 0.564$	$\theta(3;\{3,4\}) = 0.497$	$\theta(4;\{3,4\}) = 0.503$

$$
\begin{aligned}
\theta_1 &> \max\{\theta_2, \theta_3, \theta_4\} \\
0.289 &> \max\{0.192, 0.288, 0.231\} \\
\theta(3;\{2,3,4\}) &> \max\{\theta(2;\{2,3,4\}), \theta(4;\{2,3,4\})\} \\
0.371 &> \max\{0.27, 0.359\} \\
\theta(3;\{1,3,4\}) &> \max\{\theta(1;\{1,3,4\}), \theta(4;\{1,3,4\})\} \\
0.373 &> \max\{0.372, 0.255\} \\
\theta(3;\{1,2,3\}) &> \max\{\theta(1;\{1,2,3\}), \theta(2;\{1,2,3\})\} \\
0.346 &> \max\{0.34, 0.314\} \\
\theta(2;\{1,2,4\}) &> \max\{\theta(1;\{1,2,4\}), \theta(4;\{1,2,4\})\} \\
0.362 &> \max\{0.35, 0.288\}
\end{aligned}
\tag{C.1}
$$

$$
\begin{aligned}
\theta(1;\{1,2\}) &> \theta(2;\{1,2\}) \quad 0.515 > 0.485 \\
\theta(1;\{1,3\}) &> \theta(3;\{1,3\}) \quad 0.513 > 0.487 \\
\theta(1;\{1,4\}) &> \theta(4;\{1,4\}) \quad 0.515 > 0.485 \\
\theta(4;\{2,4\}) &> \theta(2;\{2,4\}) \quad 0.564 > 0.436 \\
\theta(4;\{3,4\}) &> \theta(3;\{3,4\}) \quad 0.503 > 0.497 \\
\theta(2;\{2,3\}) &> \theta(3;\{2,3\}) \quad 0.516 > 0.484
\end{aligned}
$$

Proposition 5.5. *Consider the following set of inequalities. We verify the set of inequalities is consistent by providing a numerical example:*

$$\begin{array}{lll} \theta(1;\{1,2\}) = 0.52 & \theta(2;\{2,3\}) = 0.63 & \theta(1;\{1,3\}) = 0.4 \\ \theta(2;\{1,2\}) = 0.48 & \theta(3;\{2,3\}) = 0.37 & \theta(3;\{1,3\}) = 0.6 \\ \theta_1 = 0.32 & \theta_2 = 0.33 & \theta_3 = 0.35 \\ \rho_1 = 50 & \rho_2 = 200 & \end{array}$$

$$\begin{gathered} \theta_3 > \max\{\theta_1, \theta_2\} \\ 0.35 > \max\{0.32, 0.33\} \\ \theta_1 + \theta_{312} > \theta_2 + \theta_{321} \\ \theta(1;\{1,2\}) > \theta(2;\{1,2\}) \\ 0.52 > 0.48 \\ \theta_2 + \theta_{123} > \theta_3 + \theta_{132} \\ \theta(2;\{2,3\}) > \theta(3;\{2,3\}) \\ 0.63 > 0.37 \\ \theta_3 + \theta_{231} > \theta_1 + \theta_{213} \\ \theta(3;\{1,3\}) > \theta(1;\{1,3\}) \\ 0.6 > 0.4 \end{gathered} \tag{CC}$$

$$\begin{gathered} \rho_2 * \theta(2;\{2,3\}) > [\rho_2 * \theta(3;\{2,3\})] + \rho_1 \\ 200 * (0.63) > (200 * 0.37) + 50 \\ 126 > 124 \\ [\rho_2 * \theta(3;\{1,3\})] + \rho_1 > \rho_2 * \theta(1;\{1,3\}) \\ (200 * 0.6) + 50 > 200 * (0.4) \\ 170 > 80 \end{gathered} \tag{C.2}$$

Proposition 5.8. *Consider the following set of inequalities. We can verify the set of inequalities is consistent by providing a numerical example:*

$$\begin{array}{lll} \theta(1;\{1,2\}) = 0.57 & \theta(2;\{2,3\}) = 0.52 & \theta(1;\{1,3\}) = 0.45 \\ \theta(2;\{1,2\}) = 0.43 & \theta(3;\{2,3\}) = 0.48 & \theta(3;\{1,3\}) = 0.55 \\ \theta_1 = 0.33 & \theta_2 = 0.30 & \theta_3 = 0.37 \\ \rho_1 = 100 & \rho_2 = 140 & \end{array}$$

$$\begin{gathered} \theta_3 > \max\{\theta_1, \theta_2\} \\ 0.37 > \max\{0.33, 0.30\} \\ \theta(1;\{1,2\}) > \theta(2;\{1,2\}) \\ 0.57 > 0.43 \\ \theta(2;\{2,3\}) > \theta(3;\{2,3\}) \\ 0.52 > 0.48 \\ \theta(3;\{1,3\}) > \theta(1;\{1,3\}) \\ 0.55 > 0.45 \end{gathered} \tag{CC}$$

$$\begin{gathered} \rho_1\theta_3 + \rho_2 * [\theta(3;\{2,3\})] > \rho_1\theta_2 + \rho_2 * [\theta(2;\{2,3\})] \\ (100 * 0.37) + 140 * (0.48) > (100 * 0.30) + 140 * (0.52) \\ 104.2 > 102.8 \\ \rho_1\theta_3 + \rho_2 * [\theta(3;\{1,3\})] > \rho_1\theta_1 + \rho_2 * [\theta(1;\{1,3\})] \\ (100 * 0.37) + 140 * (0.55) > (100 * 0.33) + 140 * (0.45) \\ 114 > 96 \\ \rho_1\theta_1 + \rho_2 * [\theta(1;\{1,2\})] > \rho_1\theta_2 + \rho_2 * [\theta(2;\{1,2\})] \\ (100 * 0.33) + 140 * (0.57) > (100 * 0.30) + 140 * (0.43) \\ 112.8 > 90.2 \end{gathered} \tag{C.3}$$

Appendix 2

Primary Election Rules

The last stage of the primary process and thus the most relevant is the National Convention, where all the delegates get together and choose which candidate will become the official party nominee. It is the case that for the two major political parties, delegation allocation rules tend to start with a set of general guidelines (and certain restrictions) sourced from the Party's National Committee (Democratic National Committee (DNC) and Republican National Committee (RNC)). Then each party's state committee would decide on the final rule on (a) the type of election (i.e. caucus, primary, etc), (b) the number of delegates per state/territory, (c) how to allocate delegates, (d) who would be eligible to vote in the election, (e) the date of the election, etc.

On the surface, there are no significant differences between the electoral contests in the two main parties. However there are certain peculiarities in the Democratic process: (a) By the 90's, all the Democratic primaries had to use Proportional Representation (PR) rules; in the Republican primaries each state decides on the type of voting rule and (b) only the Democratic Party has "superdelegates", who make up around 15% of the delegates. In a closer inspection, Kamarck (2018) claimed that Democratic party officials have a strong preference for primaries with PR rules whereas Republican party officials, albeit respecting states' preferences, have a strong preference for primaries with plurality rule.

The figure of superdelegates has received its fair share of criticism as a serious conflict of interest. Unlike regular elected delegates, they are appointed (by party leaders) and unbound (i.e., they are free to choose any candidate without any legal compromise) representatives that come directly from the party ranks. In the following two sub-sections, we will discuss the set of rules used by the two main parties for the Presidential Election of 2016.

Democratic Party

- **Number of Delegates**

  The rules of the National Convention and the allocation of delegates are established in the previous Convention; and the most recent document is the "Delegate Selection Rules for the 2016 Democratic National Convention". State are allocated Democratic Delegates according to a very specific formula based on a figure called the "base vote". This number reflects, in two ways, the electoral support in that state for Democratic candidates at several levels (i.e. Presidential, Senate, House, etc) in the last three electoral cycles. More precisely: "one half of a state's base vote is determined by the number of Presidential Electors assigned to that state and half are computed by the number of people who voted for the Democratic Presidential candidate in the last three elections" (Berg-Andersson & Roza, 2016). Moreover, the DNC would incentivise states to move their election to a later date and additionally to join a cluster of states by giving bonus delegates, based on how late the state moves the election date.

  Moreover, three different type of delegates are determined according to the state's Base Votes. Firstly, 25% of the base delegates are "at-large" delegates, who are basically delegates assigned for the whole state; whereas 75% of the base delegates are assigned to specific congressional districts (i.e., District delegates). Furthermore, there are unpledged Party Leader and Elected Official delegates, more commonly known as "Super Delegates".

- **Timing Sanctions** The main penalty would be to lose half of the pledged delegation if a state moved its election date before some window (for 2016 it was March 1st).

- **Timing** The DNC had established that the 2016 primary calendar for Democrats would start on "the first Tuesday in March and end on the second Tuesday in June" (Rule 11(A) of the 2016 Delegate Selection Rules for the Democratic National). However, only four states are allowed to have their election before the

"window": Iowa, New Hampshire, Nevada, and South Carolina. For the full Schedule of Democratic primaries in 2016, see Table 6.4 in Appendix 3.

- **Delegate Allocation Rules** All states have to allocate delegates using Prop. Representation rules, according to the DNC. For the full allocation of Democratic Delegates in the primary of 2016, see Table 6.2 in Appendix 3.

Republican Party

The rules of the National Convention and the Allocation of delegates are decided at the previous National Convention; so the most recent document is "The Rules of the Republican Party". Unlike Democrats, the GOP convention does not allow "super delegates". Moreover, GOP delegates are awarded more evenly: each state would get 10 at-large delegates. Moreover, states can get additional delegates based on previous electoral performance of local GOP candidates in congressional races (both local and federal). For example, states get bonus delegates for every Republican member of the House of Representatives of the U.S. Congress from that state.

- **Number of Delegates** According to Rule 14 paragraph (a), contained in The Rules of the Republican Party (2012), delegates should be allocated in the following way:

    1. Ten delegates at large from each of the fifty states.
    2. Three district delegates for each Representative in the United States House of Representatives from each state.
    3. Six delegates at large from American Samoa, sixteen delegates at large from the District of Columbia, six delegates at large from Guam, six delegates at large from the Northern Mariana Islands, twenty delegates at large from Puerto Rico, and six delegates at large from the Virgin Islands.
    4. Plus the so-called "Bonus delegates", which are awarded to each state depending on the electoral results on the previous elections, at all levels (Presidential, Senate, House, Governor and State Legislature)

For the full allocation of Republican Delegates in the presidential primary of 2016, see Table 6.1 in Appendix 3.

- **Timing** There are two main rules about the primary schedule:

  1. **Rule No. 16 paragraph (c)(1):** "No primary, caucus, convention, or other process to elect, select, allocate, or bind delegates to the national convention shall occur prior to March 1 or after the second Saturday in June in the year in which a national convention is held".

  2. **Rule No. 16 paragraph (c)(2):** States other than "Iowa, New Hampshire, South Carolina, and Nevada may not conduct their processes earlier than one month before the next earliest state in the year in which a national convention is held and shall not be subject to the provisions of paragraph (c)(2) of this rule."

  For the full Schedule of Republican Delegates in 2016, see Table 6.3 in Appendix 3.

- **Timing Sanctions** The main sanctions are for violations of the previous paragraph. If any of the first four states move their election earlier, they would lose half of their delegation automatically. If any other state moves its election earlier then that state would be penalised: its delegation would be reduced to either nine or six delegates (depending on the size of the original allocation) (See Rule 17(a))

- **Delegate Allocation Rules** Contrary to Democratic Allocation rules, the RNC offered more flexibility to states. More specifically Rule 16 paragraph (c)(2) states that any electoral contest before March 15 must use Proportional Representation (PR) rules. But then, it was further specified in Rule No. 16 paragraph (c)(3)(i) and paragraph (c)(3)(ii) that states could establish a minimum and maximum threshold of "no higher than twenty percent and no lower than fifty percent", respectively in order to receive their respective proportional assignment of delegates. Moreover, if a candidate obtains more than 50% of the votes then that candidate will get all the delegation; whereas if a candidate obtains less than 20% of the votes then that candidate would get zero delegates.

Appendix 3

Tables

Table 3.1: Distribution of Republican Delegates for the Presidential Primary of 2016

State	At-Large Del.	District Del.	Party Leaders	Bonus	Total Del.
Alabama	10	21	3	16	50
Alaska	10	3	3	12	28
Arizona	10	27	3	18	58
Arkansas	10	12	3	15	40
California	10	159	3	0	172
Colorado	10	21	3	3	37
Connecticut	10	15	3	0	28
Delaware	10	3	3	0	16
D.C.	16	0	3	0	19
Florida	10	81	3	5	99
Georgia	10	42	3	21	76
Guam	6	0	3	0	9
Hawaii	10	6	3	0	19
Idaho	10	6	3	13	32
Illinois	10	54	3	2	69
Indiana	10	27	3	17	57
Iowa	10	12	3	5	30
Kansas	10	12	3	15	40
Kentucky	10	18	3	15	46
Louisiana	10	18	3	15	46
Maine	10	6	3	4	23
Maryland	10	24	3	1	38
Massachusetts	10	27	3	2	42
Michigan	10	42	3	4	59
Minnesota	10	24	3	1	38
Mississippi	10	12	3	15	40
Missouri	10	24	3	15	52
Montana	10	3	3	11	27
Nebraska	10	9	3	14	36
Nevada	10	12	3	5	30
New Hampshire	10	6	3	4	23
New Jersey	10	36	3	2	51
New Mexico	10	9	3	2	24
New York	10	81	3	1	95
North Carolina	10	39	3	20	72
North Dakota	10	3	3	12	28
Ohio	10	48	3	5	66

State	At-Large Del.	District Del.	Party Leaders	Bonus	Total Del.
Oklahoma	10	15	3	15	43
Oregon	10	15	3	0	28
Pennsylvania	10	54	3	4	71
Puerto Rico	20	0	3	0	23
Rhode Island	10	6	3	0	19
South Carolina	10	21	3	16	50
South Dakota	10	3	3	13	29
Tennessee	10	27	3	18	58
Texas	10	108	3	34	155
Utah	10	12	3	15	40
Vermont	10	3	3	0	16
Virgin Islands	6	0	3	0	9
Virginia	10	33	3	3	49
Washington	10	30	3	1	44
West Virginia	10	9	3	12	34
Wisconsin	10	24	3	5	42
Wyoming	10	3	3	13	29
Totals	560	1,305	168	439	2,472

Note. Adapted from https://www.thegreenpapers.com/P16/R-Del.phtml.
Copyright (1999-2019) by Berg-Andersson and Roza (2016)

Table 3.2: Distribution of Democratic Delegates for the Presidential Primary of 2016

State	Pledged Delegates	Super Delegates	Total Del.
Alabama	53	7	60
Alaska	16	4	20
A. Samoa	6	5	11
Arizona	75	10	85
Arkansas	32	5	37
California	475	76	551
Colorado	66	12	78
Connecticut	55	16	71
Delaware	21	11	32
Dem. Abroad	13	4	17
D.C.	20	24	44
Florida	214	32	246
Georgia	102	15	117
Guam	7	5	12
Hawaii	25	9	34
Idaho	23	4	27
Illinois	156	27	183
Indiana	83	9	92
Iowa	44	7	51
Kansas	33	4	37
Kentucky	55	5	60
Louisiana	51	8	59
Maine	25	5	30
Maryland	95	25	120
Massachusetts	91	24	115
Michigan	130	17	147
Minnesota	77	16	93
Mississippi	36	5	41
Missouri	71	13	84
Montana	21	6	27
Nebraska	25	5	30
Nevada	35	8	43
N. Hampshire	24	8	32
N. Jersey	126	16	142
N. Mexico	34	9	43
N. York	247	44	291
N. Carolina	107	13	120
N. Dakota	18	5	23
N. Marianas	6	5	11
Ohio	143	17	160
Oklahoma	38	4	42
Oregon	61	13	74
Pennsylvania	189	19	208
Puerto Rico	60	7	67
Rhode Island	24	9	33
S. Carolina	53	6	59
S. Dakota	20	5	25
Tennessee	67	8	75
Texas	222	29	251
Utah	33	4	37
Vermont	16	10	26
Virgin Islands	7	5	12
Virginia	95	13	108
Washington	101	17	118
West Virginia	29	8	37
Wisconsin	86	10	96
Wyoming	14	4	18
Totals	4,051	712	4,763

Note. Adapted from https://www.thegreenpapers.com/P16/D-Del.phtml.
Copyright (1999-2019) by Berg-Andersson and Roza (2016)

Table 3.3: Schedule of Republican Presidential Primaries in 2016

State	Date
Iowa	Monday 1 February 2016
New Hampshire	Tuesday 9 February 2016
South Carolina	Saturday 20 February 2016
Nevada	Tuesday 23 February 2016
Alabama	Tuesday 1 March 2016
Alaska	Tuesday 1 March 2016
Arkansas	Tuesday 1 March 2016
Georgia	Tuesday 1 March 2016
Massachusetts	Tuesday 1 March 2016
Minnesota	Tuesday 1 March 2016
Oklahoma	Tuesday 1 March 2016
Tennessee	Tuesday 1 March 2016
Texas	Tuesday 1 March 2016
Vermont	Tuesday 1 March 2016
Virginia	Tuesday 1 March 2016
Kansas	Saturday 5 March 2016
Kentucky	Saturday 5 March 2016
Louisiana	Saturday 5 March 2016
Maine	Saturday 5 March 2016
Puerto Rico	Sunday 6 March 2016
Hawaii	Tuesday 8 March 2016
Idaho	Tuesday 8 March 2016
Michigan	Tuesday 8 March 2016
Mississippi	Tuesday 8 March 2016
Virgin Islands	Thursday 10 March 2016
District of Columbia	Saturday 12 March 2016
Guam	Saturday 12 March 2016
Wyoming	Saturday 12 March 2016
Florida	Tuesday 15 March 2016
Missouri	Tuesday 15 March 2016
North Carolina	Tuesday 15 March 2016
Northern Marianas	Tuesday 15 March 2016
Ohio	Tuesday 15 March 2016
American Samoa	Tuesday 22 March 2016
Arizona	Tuesday 22 March 2016
Utah	Tuesday 22 March 2016
North Dakota	Friday 1 April 2016
Wisconsin	Tuesday 5 April 2016
Colorado	Thursday 7 April 2016
New York	Tuesday 19 April 2016
Connecticut	Tuesday 26 April 2016
Delaware	Tuesday 26 April 2016
Maryland	Tuesday 26 April 2016
Rhode Island	Tuesday 26 April 2016
Indiana	Tuesday 3 May 2016
Nebraska	Tuesday 10 May 2016
West Virginia	Tuesday 10 May 2016
Oregon	Tuesday 17 May 2016
Illinois	Friday 20 May 2016
Pennsylvania	Saturday 21 May 2016
Washington	Tuesday 24 May 2016
California	Tuesday 7 June 2016
Montana	Tuesday 7 June 2016
New Jersey	Tuesday 7 June 2016
New Mexico	Tuesday 7 June 2016
South Dakota	Tuesday 7 June 2016

Note. Adapted from https://www.thegreenpapers.com/P16/R-Timing.phtml.
Copyright (1999-2019) by Berg-Andersson and Roza (2016)

Table 3.4: Schedule of Democratic Presidential Primaries in 2016

State	Start Date	Stage	Date Bonus	Cluster Bonus
Iowa	Monday 1 February 2016			
New Hampshire	Tuesday 9 February 2016			
Nevada	Saturday 20 February 2016			
South Carolina	Saturday 27 February 2016			
Alabama	Tuesday 1 March 2016	Stage I		
Arkansas	Tuesday 1 March 2016	Stage I		
Colorado	Tuesday 1 March 2016	Stage I		
Georgia	Tuesday 1 March 2016	Stage I		
Massachusetts	Tuesday 1 March 2016	Stage I		
Minnesota	Tuesday 1 March 2016	Stage I		
Oklahoma	Tuesday 1 March 2016	Stage I		
Tennessee	Tuesday 1 March 2016	Stage I		
Texas	Tuesday 1 March 2016	Stage I		
Vermont	Tuesday 1 March 2016	Stage I		
Virginia	Tuesday 1 March 2016	Stage I		
Kansas	Saturday 5 March 2016	Stage I		
Louisiana	Saturday 5 March 2016	Stage I		
Nebraska	Saturday 5 March 2016	Stage I		
Michigan	Tuesday 8 March 2016	Stage I		
Mississippi	Tuesday 8 March 2016	Stage I		
Florida	Tuesday 15 March 2016	Stage I		
Illinois	Tuesday 15 March 2016	Stage I		
Missouri	Tuesday 15 March 2016	Stage I		
North Carolina	Tuesday 15 March 2016	Stage I		
Ohio	Tuesday 15 March 2016	Stage I		
Arizona	Tuesday 22 March 2016	Stage I cluster		15%
Idaho	Tuesday 22 March 2016	Stage I cluster		15%
Utah	Tuesday 22 March 2016	Stage I cluster		15%
Alaska	Saturday 26 March 2016	Stage I cluster		15%
Hawaii	Saturday 26 March 2016	Stage I cluster		15%
Washington	Saturday 26 March 2016	Stage I cluster		15%
Wisconsin	Tuesday 5 April 2016	Stage II cluster	10%	
Wyoming	Saturday 9 April 2016	Stage II cluster	10%	
New York	Tuesday 19 April 2016	Stage II cluster	10%	
Connecticut	Tuesday 26 April 2016	Stage II cluster	10%	
Rhode Island	Tuesday 26 April 2016	Stage II cluster	10%	
Delaware	Tuesday 26 April 2016	Stage II cluster	10%	15%
Maryland	Tuesday 26 April 2016	Stage II cluster	10%	15%
Pennsylvania	Tuesday 26 April 2016	Stage II cluster	10%	15%
Indiana	Tuesday 3 May 2016	Stage III cluster	20%	
West Virginia	Tuesday 10 May 2016	Stage III cluster	20%	
Kentucky	Tuesday 17 May 2016	Stage III cluster	20%	
Oregon	Tuesday 17 May 2016	Stage III cluster	20%	
California	Tuesday 7 June 2016	Stage III cluster	20%	
New Jersey	Tuesday 7 June 2016	Stage III cluster	20%	
New Mexico	Tuesday 7 June 2016	Stage III cluster	20%	
Montana	Tuesday 7 June 2016	Stage III cluster	20%	15%
North Dakota	Tuesday 7 June 2016	Stage III cluster	20%	15%
South Dakota	Tuesday 7 June 2016	Stage III cluster	20%	15%

Note. Adapted from https://www.thegreenpapers.com/P16/D-Timing.phtml.
Copyright (1999-2019) by Berg-Andersson and Roza (2016)