Primaries and the New Hampshire Effect*

Tilman Klumpp
Indiana University

Mattias K. Polborn†
University of Western Ontario

December 29, 2002

Abstract

Candidates for U.S. presidential elections are determined through sequential elections in single states, the primaries. We develop a model in which candidates can influence their winning probability in electoral districts by spending money on campaigning. The equilibrium replicates several stylized facts very well: Campaigning is very intensive in the first district. The outcome of the first election then creates an asymmetry in the candidates' incentives to campaign in the next district, which endogenously increases the equilibrium probability that the first winner wins in further districts.

On the normative side, our model offers a possible explanation for the sequential organization: It leads (in expectation) to a lower level of advertising expenditures than simultaneous elections. Moreover, if one of the candidates is the more effective campaigner, sequential elections also perform better with regard to the selection of the best candidate.

Keywords: Political economy, elections, political campaigns.
JEL code: C7, D7.

*We would like to thank an anonymous referee and the editor for very helpful comments that improved this paper considerably. We are also thankful to Srihari Govindan, Ig Horstmann, Arthur Robson, Al Slivinski, Peter Streufert, John Whalley and seminar participants at the University of Western Ontario, Indiana University and the Spring 2001 Midwest Economic Theory Meetings.

†Corresponding author: Mattias Polborn, Department of Economics, University of Western Ontario, London, Ontario, N6A 5C2, Canada; email: mpolborn@uwo.ca

1 Introduction

Candidates for the U.S. presidential election are determined through a sequence of elections within each political party, the primaries. While the particular regulations vary between states and the two major parties, the basic system is the same in both parties, starting with the Iowa caucus and the New Hampshire primary in February, and continuing with primaries (and very few caucuses) in almost all U.S. states, in which more than 80% of the delegates to the national convention (that elects the party's candidate) are chosen.

The nomination process is one of the most controversial institutions of America's contemporary political landscape. The most common ground for attack on the modern primary system is the perception that its sequential structure is inherently "unfair" in that it shifts too much power to voters in early primary states. A notion that usually comes along with such claims is that the results of early primaries create "momentum" that carries over to later states. 1976 Democratic primary candidate Morris Udall (who eventually lost to Jimmy Carter) notes:

> "We had thirty primaries, presumably all of them equal. After three of those primaries, I'm convinced, it was all over. [...] I take a poll two weeks before the (Wisconsin) primary and he (Carter) is ahead of me, two to one, and has never been in the state except for a few quick visits. That was purely and solely and only the product of that narrow win in New Hampshire and the startling win in Florida." (Witcover, 1977)

Early primary states receive considerable attention by both the political candidates and the media. Malbin (1983) reports that in the 1980 Republican primaries George Bush and Ronald Reagan allocated roughly 3/4 of their respective total campaign budgets to states with primaries before March 31, although these states accounted for considerably less than a fifth of the delegates to the Republican convention in 1980. Among all primaries and caucuses in 1980, Iowa and New Hampshire accounted for 28% of the primary-season coverage in the CBS evening news and the United Press newswire (Robinson and Sheehan, 1983). Similarly, Adams (1987) reports that the 1984 New Hampshire primary attracted almost 20% of the season's coverage in ABC, CBS, NBC, and the New York Times. All these observations are the more surprising as New Hampshire accounts for only 0.4 percent of the U.S. population and only four out of 538 electoral votes in the presidential election, and is far from being demographically representative for the nation's electorate.

The present paper has two interrelated objectives, one positive and one normative: Firstly, we address the question how the observed sequential organization can create sources for strategic momentum that can explain the stylized facts above. Why does the sequential nature of the current primary system induce candidates to campaign so heavily at early stages and the losers of early primaries to withdraw so early from the race? Secondly, in a related vein, we address the question how the temporal organization of elections affects a candidate's welfare, his expected campaign expenditures, and probability of winning under alternative temporal structures. The particular comparison we make is between a sequential system, such as the current presidential primaries, and a counterfactual simultaneous system. A completely simultaneous design (a "national one-day primary") is a natural antipode as well as a prominent counterproposal to the sequential primary arrangement. Therefore, it is an important and interesting question to compare these two temporal organizations.

To this end, we develop an advertising model of political competition in which candidates have to win the majority of a number of electoral districts in order to obtain a certain prize. As in Snyder (1989), candidates can influence their probability of winning a district by their choice of campaign expenditures in that district. In the case of a sequential primary organization, campaign expenditures are very high in early districts, but decrease substantially at later stages, once one candidate has established a clear advantage. Sequential elections leave an expected rent to the candidates, which is bounded from below by a positive constant that is independent of the number of electoral districts. In contrast, simultaneous elections lead to complete rent dissipation if the number of electoral districts is sufficiently high. In other words, the expected campaign expenditures are lower when candidates face a sequential primary system. Interestingly, this cost-advantage is not so much driven by the fact that the need to go through the entire sequence of primaries rarely arises (and hence candidates can save on the spending in the last few primaries, as one might suspect). Rather, it is generated by a strategic "New Hampshire Effect" that is part of the equilibrium play: The outcome of the very first primary election creates an asymmetry between ex-ante symmetric candidates which endogenously facilitates momentum in later districts.

While there is no direct reason why parties should be concerned with candidates' expenditures and ex-ante expected level of rent, each party clearly has an interest that the candidate who wins its nomination keeps resources for the following presidential campaign against the other party's nominee. A long standing "rule" in American politics was that a candidate who lost his party's primary in New Hampshire would not become president.[1] An interpretation of this empirical fact is that a candidate who did not win the first pri-

[1] Recently, this "rule" was broken by Bill Clinton in 1992 and George W. Bush in 2000.

mary, but eventually won his party's nomination, had to go through a long and costly nomination battle in his own party and lacked resources in the actual presidential campaign. If this is the case, each party has an incentive to organize its candidate selection procedure in a way that minimizes wasteful internal battles.[2]

If one candidate is a stronger campaigner than his opponent (in the sense that he is more likely to win if both candidates spend the same amount), a good primary system should also select the stronger candidate with a high probability. In our model, this probability is close to 1 under a sequential regime, provided that the number of primary districts is sufficiently high. A simultaneous system, on the other hand, frequently selects the weaker candidate. These two advantages of a sequential primary system—lower campaign expenditures and higher probability that the efficient candidate is selected—may explain why the sequential organization has been so persistent over time, even though it is often criticized as unfair.

Although our analysis compares mainly two extreme cases—completely sequential elections versus completely simultaneous elections—, the distinct results of the sequential case basically apply to a mixed temporal structure as well, as long as it involves some sequential elements at the early stages. One can argue that such an intermediate system is closer to the modern primary races, in which there are dates (such as "Super Tuesday") when several states vote simultaneously. Nevertheless, even in this case, some primary states vote in sequence at the very beginning of the nomination process. We show that this is enough to generate (and sometimes even amplify) the momentum effect and the spending pattern that arise in a completely sequential system.

Regarding the analysis of the temporal structure of elections, Dekel and Piccione (2000) have analyzed a model of sequential elections in which sophisticated voters try to aggregate their private information through voting. While, in principle, more information is available for voters in later elections, they show that the voting equilibria of sequential elections are essentially the same as those in the case of simultaneous elections. Consequently, the temporal organization of elections does not matter in their model. We provide a complementary model to theirs, which abstracts from the information aggregation aspect of voting and focuses purely on the candidates' actions. We demonstrate that in such a framework, where candidates are modelled as economic agents trying to maximize their payoffs, the two organizational forms are no longer equivalent. From the point of view

[2]The assumption that the organization of the primary system can be chosen by parties appears justified, because the primary system is not part of the constitution, and therefore is subject to fewer political and legal constraints than the following presidential election.

of the candidates and their parties, a sequential organization has considerable advantages over a simultaneous one.

Several alternative approaches have been offered in the literature as explanation of the stylized facts. On the one hand, the political science literature contains theories of "psychological momentum" among primary voters in the sense that voters in later states enjoy voting for candidates that were successful in earlier states (Bartels, 1988). These behavioral theories lack a solid preference-based foundation, resulting in rational decisions, and are hence unsatisfactory from an economic point of view. On the other hand, in their seminal work on informational cascades, Bikhchandani, Hirshleifer and Welch (1991) interpret momentum in primary races as evidence of rational herding on part of the later primary states. However, it is unclear whether primaries are really a valid example for herding: In herding models, people are concerned with *making the right choice*. In a standard model of voting, people do not so much care about whether *they themselves* voted for the correct candidate, but rather whether in the end the right candidate is selected.

Another strand of the literature on primaries is concerned with the question whether the institution of separate primaries for a left-wing and a right-wing party lead to the selection of extremist (vs. moderate) politicians as candidates (see Heckelman (2000), Swank (2001) and Oak (2001)). In these papers, only one primary takes place within each party, so the effects of the sequential process characteristic for the *presidential* primaries (the focus of our model) cannot be explored in these models.

The paper proceeds as follows. In Section 2, we present a model of the primary system. Section 3 analyzes and compares the equilibria under different temporal regimes with respect to their expected outcomes, and shows that several stylized facts concerning primaries can be explained by our model. Section 4 presents two extensions of the basic model, asymmetric campaign strengths, and mixed primary systems involving both sequential and simultaneous stages. Section 5 concludes. All proofs are in the Appendix.

2 The Model

Candidates and Electoral Districts. There are two risk neutral candidates, 1 and 2, who compete in elections in J (odd) districts. The candidate who wins at least $J^* = \frac{J+1}{2}$ elections wins the prize Π, normalized to 1 (and assumed to be equal for both candidates).

The outcome of the election in district j is a random variable from the point of view of the candidates. They can influence the distribution of this random variable by committing campaign funds to each district (see below). Campaign expenditures represent advertising effort, the cost of time, etc. Let $x_j \geq 0$ be the amount spent by candidate 1 in district j, and likewise let $y_j \geq 0$ be the amount spent by candidate 2 in j. The net utility of a player is equal to the prize (if he wins) minus the campaign expenditures: If candidate 1 wins at least J^* districts, he obtains a payoff of $1-\sum_{j=1}^{J} x_j$, otherwise he gets $-\sum_{j=1}^{J} x_j$. The payoff for candidate 2 is defined analogously. The rent dissipation rate, defined as the fraction of the prize which is spent by the two candidates together in their effort to win it, is $\sum x_j + \sum y_j$.

Campaign Technology. Given the spending profile (x_j, y_j), the probability that a candidate wins election j is determined by a campaign technology $f : \mathbb{R}_+^2 \rightarrow [0,1]$, that is, candidate 1 wins with probability $f(x_j, y_j)$ and candidate 2 wins with probability $1 - f(x_j, y_j)$. We assume the following functional form (where $\alpha \in (0,1]$):

$$f(x,y) = \frac{x^\alpha}{x^\alpha + y^\alpha}$$

if $x > 0$ or $y > 0$, and $f(0,0) = 1/2$. Observe that f is continuously differentiable on $\mathbb{R}_{++}^2$, homogeneous of degree 0 in (x,y), increasing and strictly concave in x, and decreasing and strictly convex in y. Candidates are symmetric: $f(x,y) = 1 - f(y,x)$ for all x, y.[3] The parameter α is a measure for the marginal effect of campaign spending. If α is very low, the winning probability is close to 1/2 (as long as both candidates spend a positive amount) and largely independent of the candidates' spending. The higher is α, the higher is the marginal effect of campaign spending on the outcome (and consequently, both candidates have a higher incentive to spend when α is large). The assumption that $\alpha \leq 1$ guarantees that f is globally concave.

Temporal Structure. There are two principal ways to organize J elections temporally: They can be held sequentially (as in the present presidential primary system), or simultaneously. In the sequential elections game G_J^{seq}, the candidates first choose campaign expenditure levels x_1 and y_1 in district 1. Then, they observe the outcome in district 1 and move on to district 2, where they choose x_2 and y_2. Within each district j, x_j and y_j are chosen simultaneously. The procedure continues until a candidate has accumulated the required majority of J^* districts.

In the simultaneous election game G_J^{sim}, candidates choose all x_j and y_j $(j = 1, \ldots, J)$ simultaneously. Then, the outcomes in all districts are observed and the candidate who has gained at least the required majority J^* wins the prize.

[3]In section 4.2, we consider the case of asymmetric candidates.

Combinations of these two basic structures are also possible; for example, one could start with n sequential elections and hold the remaining $J - n$ elections simultaneously. We view the current primary structure as mainly sequential, although there are certain simultaneous elements added to it (e.g., on "Super Tuesday", where several states hold their primaries simultaneously). The focus of our analysis is on the comparison between the completely simultaneous and the completely sequential case; the issue of mixed temporal arrangements will be addressed briefly in section 4.1.

Related literature. Our modelling approach is a multi-object variant of the rent-seeking game formulated by Tullock (1980), an all-pay auction model of lobbying. There, two bidders compete for a prize by submitting monetary bids (bribes) to a bureaucrat who has the power to allocate a political favor. The bureaucrat allocates winning probabilities (according to the exogenously given functional form) and draws the winner. The bribes are not recoverable: All bidders have to pay their submitted bids, regardless of whether they were allocated the object or not. Here, we use all-pay auctions as a model of political competition in elections, where campaign expenditures are naturally not recoverable.

As a description of the effect of campaign spending on election outcomes, this model has been analyzed by Snyder (1989). He analyzes the campaign expenditure allocation game between two parties, which compete in a number of districts (e.g., Republicans and Democrats in an election for the House of Representatives). His focus is on the effect different objective functions of the parties have on the allocation of campaign resources (e.g., what happens if parties wish to maximize the expected number of seats, or the probability that they win the majority in the house?).

Our model is also related to Szentes and Rosenthal (2001a), who study all pay majority auction games.[4] As in our model, the objective of players in their paper is to win a simple majority of districts. However, they assume that the candidate who spends the most in a particular district wins that district with certainty (this corresponds to $\alpha \to \infty$ in our model). Szentes and Rosenthal characterize the mixed strategy equilibrium in a simultaneous election.

3 Equilibrium

We start with an analysis of the (counterfactual) simultaneous election game in section 3.1. This highlights the problem of excessive campaigning that would arise under this

[4]See also Szentes and Rosenthal (2001b).

organizational form. We then show in sections 3.2-3.5 that a sequential organization avoids this problem and generates an equilibrium behavior of candidates that resembles the stylized facts.

3.1 Simultaneous Elections

The pure strategy space for a player in G_J^{sim} is $\mathbb{R}^J$, which, depending on J, can be rather high dimensional. Fortunately, we can show below that there are equilibria in which candidates' strategies take the following very simple form:

Definition 1. *Candidate 1 (2) plays a* ***uniform campaign strategy*** *if he chooses $x \geq 0$ ($y \geq 0$) according to some cumulative distribution Λ_1 (Λ_2), and then sets $x_j = x$ ($y_j = y$) for all j.*

An equilibrium of G_J^{sim} in which both players choose uniform campaign strategies is called a ***uniform campaign equilibrium*** *(UCE).*

Note that the word "uniform" in these definitions means that the total investment level is equally distributed across districts (and not that Λ_i is uniform on its support). We will often refer to a *symmetric UCE* (SUCE), using the word "symmetric" in the usual sense to indicate that both players use the same strategy ($\Lambda_1 = \Lambda_2$).

The advantage of uniform campaign strategies is that the dimensionality of the players' strategy spaces is reduced from J to 1. If players use uniform strategies, the analysis of the game G_J^{sim} is simplified considerably. We prove in Proposition 1 below that a UCE exists for all games, and that, if the UCE is in pure strategies, then it is the unique Nash equilibrium of the game.

Before we proceed to Proposition 1, it is instructive to characterize the pure strategy UCE, if it exists, and to show why existence of a pure strategy equilibrium fails, if J is large. Given two pure uniform campaign strategies $x, y \in \mathbb{R}_+$, candidate 1's payoff is

$$u_1(x, y) = F(x, y) - Jx, \tag{1}$$

where

$$F(x, y) = \sum_{k=J^*}^{J} \binom{J}{k} \frac{x^{\alpha k} y^{\alpha(J-k)}}{(x^\alpha + y^\alpha)^J} \tag{2}$$

is the overall winning probability for candidate 1: He wins the prize if he wins in $k \geq J^*$ districts. If his opponent mixes with distribution Λ_2, we write 1's expected payoff from playing pure strategy x as

$$u_1(x, \Lambda_2) = \int_y F(x, y) d\Lambda_2(y) - Jx.$$

Of course, an analogous payoff structure arises for candidate 2. To find a SUCE in pure strategies, differentiating (1) with respect to x, invoking symmetry ($y = x$) and simplifying yields the following first order condition:

$$\binom{J-1}{\frac{J-1}{2}}\left(\frac{1}{2}\right)^{J-1}\frac{\alpha}{4x} = 1. \tag{3}$$

The first two terms, $\binom{J-1}{\frac{J-1}{2}}\left(\frac{1}{2}\right)^{J-1}$ represent the probability that, given equal spending by both opponents in each district, exactly one half of the $J-1$ districts are won by candidate 1, and the other half by candidate 2: In this case (and only in this case), the outcome in the last district is pivotal. The third term, $\frac{\alpha}{4x}$, is the marginal effect of additional spending in the last district on the winning probability there; to see this, differentiate the winning probability in that district, $\frac{x^\alpha}{x^\alpha+y^\alpha}$, with respect to x and use symmetry ($y = x$). Put differently, the marginal benefit of spending in any district j is the marginal increase in probability of winning district j times the probability of j being pivotal. At the optimum, this marginal benefit of spending has to be equal to its marginal cost, which is 1. Consequently, a pure strategy UCE (if it exists) is given by

$$x = y = \frac{\alpha}{2^{J+1}}\binom{J-1}{\frac{J-1}{2}}. \tag{4}$$

Since the equilibrium rent for each candidate in a symmetric UCE is $\frac{1}{2} - Jx$, the value for x, defined by (4), must satisfy $x \leq \frac{1}{2J}$ in order for a pure strategy equilibrium to exist. Hence, (4) constitutes an equilibrium if and only if

$$\alpha \leq 2^J \Big/ J\binom{J-1}{\frac{J-1}{2}}. \tag{5}$$

In the appendix, we show that the right hand side of this inequality is decreasing in J and goes to zero for $J \to \infty$. Hence, (5) implicitly defines, for each α, a maximum number of districts $K(\alpha)$ for which pure strategy symmetric UCE exist. While there is no closed form solution for $K(\alpha)$, we show in the appendix that $K(\alpha) \approx \frac{2\pi}{\alpha^2}$, where $\pi = 3.1415\ldots$, is a very good approximation. For a number of districts larger than $K(\alpha)$ only mixed strategy equilibria exist.

The intuition for this result is as follows: Independent of J, a candidate needs to win just one more district than his opponent. Suppose that J is large and both candidates spend the same amount, which cannot be larger than 1/2. Now, if player 1 increases his total campaign spending by a small amount and distributes the additional spending equally over all districts, he wins every district with a slightly higher probability than his opponent. The law of large numbers then implies that player 1 wins the majority of districts with probability close to one (even though, of course, he wins only slightly more than half of the districts in expectation). Hence, for J sufficiently large, a symmetric pure strategy profile cannot be an equilibrium.

We also show in Proposition 1 that for $J > K(\alpha)$, the candidates' rent is completely dissipated in the mixed strategy SUCE (in expectation). Intuitively, players have a strong incentive to "outcampaign" their opponent, as J grows, because the marginal effect of spending on the probability to win the whole race increases. This is very similar to the effect of an increase of α for a fixed number of districts, and also leads to complete rent dissipation.

In fact, for $J \to \infty$, our game approaches the standard all-pay auction in which candidates choose their overall expenditures and the candidate with the higher expenditure wins. This observation alone implies that equilibrium payoffs in the limit (for $J \to \infty$) must be zero. However, our result is stronger, as there exists a finite number $K(\alpha)$ such that $v_J^{\mathrm{sim}} = 0$ for $J > K(\alpha)$. The intuition why K decreases in α is as follows: A high value of α means that the marginal effect of campaigning *per district* is high. Consequently, for high α, the number of districts from which on the pure strategy UCE vanishes is smaller than for small α, so that K is decreasing.

Our results for this section are formally stated in the following Proposition 1. If a pure strategy equilibrium exists, the SUCE is in fact the unique pure strategy equilibrium. A SUCE also exists for $J > K(\alpha)$ (this time in mixed strategies), and we will focus on this equilibrium for the mixed strategy case. This seems reasonable, first in analogy to the pure strategy case, and second, because it is certainly the simplest mixed strategy equilibrium (since it involves only the draw of a unidimensional random variable). Moreover, in the mixed strategy SUCE, rents are fully dissipated in expectation.[5]

Proposition 1. *There exists a decreasing function $K : (0, 1] \to \mathbf{R}$, implicitly defined by equality in (5), such that*

(a) if $J \leq K(\alpha)$, a UCE in pure strategies exists in G_J^{sim}. Rents are (generically) not fully dissipated in this equilibrium. Furthermore, the UCE is the only pure strategy equilibrium.

(b) if $J > K(\alpha)$, a pure strategy UCE does not exist. A symmetric, mixed strategy UCE exists.

(c) If $J > K(\alpha)$, the mixed strategy SUCE involves full rent dissipation: $E(\sum x_j) = E(\sum y_j) = 1/2$, and candidates' expected rent is $v_J^{sim} = 0$ for all $J > K(\alpha)$.

[5]We conjecture, but have not been able to prove, that the mixed SUCE is the unique equilibrium.

3.2 Sequential Elections: Equilibrium existence and uniqueness

The sequential election game with J districts, G_J^{seq}, can be analyzed using backward induction. After $j-1$ elections have been held, call a *state* for a candidate a tuple (j,k) where k is the number of elections won by the candidate so far. Consequently, the opponent is in state $(j, j-k-1)$. Let $x_{j,k}$ be the candidate's spending in state (j,k), and $v_{j,k}$ his continuation value.[6] The value $v_J^{\text{seq}} = v_{1,0}$ is then the value of the game G_J^{seq}.

The continuation value does not take account of any prior investments, because these have to be considered as sunk costs by the candidates. A useful consequence of the sunk cost property is that, if we extend the game from G_J^{seq} to G_{J+2}^{seq}, all we have to do is to introduce a number of new states and relabel. That is, $v_{j,k}$ becomes $v_{j+2,k+1}$, and likewise $x_{j,k}$ becomes $x_{j+2,k+1}$. For instance, the problem when each candidate has won exactly one election in G_{J+2}^{seq} is the same as the problem at the very beginning of G_J^{seq}.

Given (j,k), we can now set up a pair of Bellman equations, one for each player:

$$v_{j,k} = \max_{x_{j,k}} \left\{ \frac{x_{j,k}^{\alpha}}{x_{j,k}^{\alpha} + x_{j,j-k-1}^{\alpha}} v_{j+1,k+1} + \frac{x_{j,j-k-1}^{\alpha}}{x_{j,k}^{\alpha} + x_{j,j-k-1}^{\alpha}} v_{j+1,k} - x_{j,k} \right\} \tag{6}$$

and

$$v_{j,j-k-1} = \max_{x_{j,j-k-1}} \left\{ \frac{x_{j,j-k-1}^{\alpha}}{x_{j,k}^{\alpha} + x_{j,j-k-1}^{\alpha}} v_{j+1,j-k} \right. \\ \left. + \frac{x_{j,k}^{\alpha}}{x_{j,k}^{\alpha} + x_{j,j-k-1}^{\alpha}} v_{j+1,j-k-1} - x_{j,j-k-1} \right\}. \tag{7}$$

When we set $v_{j,J^*} = 1$ for $j = J^*, \ldots, J+1$ and $v_{j,j-J^*-1} = 0$ for $j = J^*+1, \ldots, J+1$, the Bellman equations become a finite horizon dynamic programming problem.

Since it is always feasible for a candidate to spend zero, we have $v_{j,k} \geq 0$ for all j,k, and, as shown in Lemma 5, $v_{j,k} \geq v_{j,k-1}$ and $v_{j,k} \geq v_{j+1,k}$. Together with the fact that $0 < \alpha \leq 1$, this implies that the right-hand side of each Bellman equation is strictly concave in its respective decision variable.

Define $\Delta_{j,k} = v_{j,k} - v_{j,k-1}$ to be the difference in the continuation payoff from winning the stage election in state (j,k) and losing it, and let

$$\theta_{j,k} = \frac{\Delta_{j+1,j-k}}{\Delta_{j+1,k+1}}.$$

[6] In this section, it is notationally more convenient to denote both candidates' spending by x. Since the subscripts indicate the respective state candidates are in, no confusion should arise.

Taking first order conditions of (6) and (7), the ratio of the candidates' expenditures is

$$\frac{x_{j,j-k-1}}{x_{j,k}} = \theta_{j,k}.$$

Using this relation in the first order conditions, the unique solutions of the first order conditions are

$$x_{j,k} = \alpha \frac{\theta_{j,k}^{\alpha}}{(1+\theta_{j,k}^{\alpha})^2}\Delta_{j+1,k+1} \text{ and } x_{j,j-k-1} = \alpha \frac{\theta_{j,k}^{\alpha}}{(1+\theta_{j,k}^{\alpha})^2}\Delta_{j+1,j-k}. \tag{8}$$

Due to the strict concavity of the Bellman equations, the first order necessary conditions are also sufficient. Since on each stage of the game, there is a unique continuation equilibrium, the usual backwards induction argument shows that the subgame perfect equilibrium of game G_J^{seq} is unique.

Proposition 2. *For all J, the sequential elections game G_J^{seq} has a unique subgame perfect equilibrium in pure strategies.*

It is easy to show that sequential elections leave candidates with a rent, even if there are very many districts; thus, sequential elections are at least asymptotically better than simultaneous elections. Call (j,k) *symmetric* if j is odd and $k = \frac{j-1}{2}$. (In particular, the initial state $(1,0)$ of every G_J^{seq} is symmetric.) In symmetric (j,k), $\theta_{j,k} = 1$, so that

$$x_{j,k} = x_{j,j-k-1} = \frac{1}{4}\alpha\Delta_{j,k}.$$

Since each candidate is equally likely to win the stage election, we have

$$\begin{aligned} v_{j,k} &= \left(\frac{1}{2} - \frac{1}{4}\alpha\right) v_{j+1,k+1} + \left(\frac{1}{2} + \frac{1}{4}\alpha\right) v_{j+1,k} \\ &\geq \left(\frac{1}{2} - \frac{1}{4}\alpha\right) v_{j+1,k+1} \geq \frac{1}{4} v_{j+1,k+1} \geq \frac{1}{4} v_{j+2,k+1}. \end{aligned}$$

Using $v_{J+1,J^*+1} = 1$, we have $v_J^{\mathrm{seq}} \geq 4^{-J^*} > 0$. Hence

Proposition 3. *For any number of districts J, $v_J^{seq} > 0$*

In the following, we want to characterize further the equilibrium in the sequential election model. Apart from the question of rent dissipation, we are also interested in whether the equilibrium behavior of candidates looks similar to the stylized facts that we observe in reality, in particular whether campaigning is very intensive in the initial primaries and then decreases considerably.

Ideally, we would like to solve these questions by calculating how the equilibrium strategies and the value of the game depend on the parameters J and α. Unfortunately, a general closed form solution is too difficult to obtain for all but a very small number of districts. We will therefore have to resort to two further methods, numerical simulations and analysis of a limit game, that will help us to better understand the basic features of the candidates' behavior in very long nomination races.

3.3 Sequential Elections: Numerical Results

In this section, we present quantitative results obtained from the numerical computation of the SGPE in the sequential game and the UCE of the simultaneous game.[7] Table 2 (in the Appendix) compares the simultaneous and sequential election games for $\alpha \in \{0.25, 0.5, 0.75, 1\}$ and various values for J. In the table, v_J^{sim} is the equilibrium payoff to a candidate of the simultaneous game G_J^{seq}, and $\Sigma^{\text{sim}}x$ is the expected sum of expenditures per candidate. Similarly, v_J^{seq} and $\Sigma^{\text{seq}}x$ are the value and the expected expenditures for G_J^{sim}. For the sequential game, Table 2 also gives the strategy components $x_{1,0}$ (expenditures in the first primary district), as well as $x_{2,1}$ and $x_{2,0}$ (expenditures in the second primary district after having won ($x_{2,1}$) or lost ($x_{2,0}$) the first election). Finally, $P_{2,1}$ denotes the probability that a candidate wins the entire race, conditional on a victory in the first district.

Consider first the table for $\alpha = 1$. The ex ante expected rent of a candidate in the sequential organization, v_J^{seq}, is always substantially higher than the rent in the simultaneous organization, and converges to approximately 0.1838 for very many districts. Hence, for $\alpha = 1$, about 37% of the maximum possible rent is not dissipated by the candidates, even for a number of districts where rent dissipation is complete in a simultaneous organization of primaries. Consequently, the expected total campaigning expenditures in the sequential organization converge to 0.6324. In the simultaneous organization, these expected expenditures are 1 for $J \geq 7$. In the very first district of the sequential primary race, both candidates spend 0.1775, respectively, which is about 56% of their total expected expenditures. Note that the proportion of campaign expenditures in the first district to total expenditures is essentially the same, whether there are 19, 49 or 99 districts. Hence, campaigning in the first district is fierce even if this district is rather unimportant in terms

[7]We compute SGPE in G_J^{seq} by backward induction. Pure strategy UCE of G_J^{sim} and their values are readily obtained from (4). As far as v_J^{sim} is concerned, it is not necessary to compute mixed strategy UCE, since we already know from Proposition 1 that the entire rent is dissipated there. The computation of UCE in discretized versions of G_J^{sim} is nevertheless possible given the payoff structure (Dasgupta and Maskin, 1986), e.g. by using the Linear Complementarity algorithm. This exercise confirms Proposition 1.

of its number of delegates.[8]

In the second district, the situation is necessarily asymmetric and second district expenditures for both candidates are considerably lower. The front runner spends 0.1218, while the runner up reduces his expenditures to 0.0279. The probability that the candidate who won the first district wins the second district as well is therefore about 81.4%. Should this candidate win the second district as well, a similar asymmetry effect can be observed again, endogenously furthering the advantage of the winner: Relative to his opponent, the front runner spends much more, virtually guaranteeing him yet another victory. For example, if one candidate has won the first 3 elections, the candidate who has fallen behind spends about 1/5000 of the front runner's expenditures in that district, virtually conceding the race to his opponent. Yet, this increase in relative spending by the race leader occurs at much lower absolute levels, keeping overall campaign expenditures low. Thus, once a candidate has managed to establish an early lead, he will, in all likelihood, win the entire race, and can do so with comparatively little use of further campaign resources.

Such endogenous momentum (the fact that early victories make further successes more likely) implies that the cost-advantage of sequential elections is not so much driven by the fact that the need to go through the entire sequence of primaries rarely arises (and hence, candidates do not need to maintain a costly campaign in late primary districts). Rather, it is generated by a strategic "New Hampshire effect" that is part of the equilibrium play: The outcome of the first election creates an asymmetry between ex-ante symmetric candidates, which triggers different spending patterns in subsequent primaries. Because it is now less likely for the candidate who has fallen behind to win, the absolute level of campaign expenditures decreases sharply.[9]

While the winner of the first election always spends more than his opponent in the second election and hence will likely expand his lead, it could of course happen by chance that the first-district runner up wins the second district. In this case, we would see vigorous campaigning in the third district, and, more generally, in all districts in which the candidates have won the same number of districts. In light of our model, the campaign

[8]Note that convergence is effectively completed for $J = 19$. For large J, the SGPE strategies in G_J^{seq} become stationary, so that the SGPE in these games converges to a Markov perfect equilibrium (MPE), in which strategies are measurable with respect to the lead one candidate has over the other. Intuitively, large games become very similar to each other, and candidates use very similar continuation strategies in G_{99}^{seq} as they do in, say, G_{97}^{seq}.

[9]This result is somewhat reminiscent of Che and Gale (2000a), who—in a context of lobbying games—show that asymmetry between lobbyists increases the expected rents in equilibrium. In our model, the sequential structure of primaries is used to create this asymmetry endogenously.

expenditures in the New Hampshire presidential primary are high not so much because it is the first primary state, but rather because no candidate has yet established a clear lead in the race. That such a head-to-head situation with high expenditures occurs in later stages is possible, but improbable.

What is the effect of a change in the campaign technology? In both simultaneous and sequential primaries, a decrease in α leads to (weakly) less rent dissipation. Intuitively, as advertising becomes (marginally) less effective, candidates spend less aggressively and a larger part of the possible rent is not dissipated.

For 49 districts and $\alpha = 1$, spending in the first district as a proportion of total (expected) spending is approximately 56%. For the same number of districts, this proportion is about 37% for $\alpha = 0.75$, 19% for $\alpha = 0.5$, and 4.7% for $\alpha = 0.25$. As argued above, a lower value of α will lead to lower total campaign expenditures; but since this applies to expenditures in all districts, it would not by itself explain an effect on the ratio of first-district spending to total spending.

However, an indirect, strategic effect amplifies the direct effect: Consider a 3 district race. If α is high, a large proportion of the continuation rent is dissipated in the third district, if that district is contested after both candidates have won one district, respectively. Therefore, in the second district, the loser of the first district will not campaign very intensely, because he does not have to win that much. Consequently, once an advantage is established, the front runner will spend more than the runner up, and is more likely to be the eventual winner. These strategic effects later on make it vital to win the first district.

On the other hand, when α is low, a fixed advantage over the opponent becomes less valuable: In this case, the winning probability is close to 1/2 for both candidates in every district, and so winning the first district becomes less important, because it is harder (in equilibrium) to maintain this advantage after it has been established.

3.4 Sequential elections: Limit behavior

In this section, we want to explore the candidates' behavior in nomination races with very many districts theoretically. To do so, we consider a slight variant of our original game:

Definition 2. *A* b-advantage game *is defined as follows: The technology of influencing an election in a district is the same as in our original model, and the candidate who is the first to establish a lead of* b *districts wins the game (i.e. gets a prize of 1, and the game ends).*

The b-advantage game is motivated by the numerical results. As reported in the last section, most campaign expenditures take place when both candidates are close together in terms of their victories in previous districts; once a candidate has secured an advantage of several victories, his opponent spends almost no money (and so securing further victories is very cheap for the front-runner). This is true even if there are still very many periods to go and the front runner's advantage is relatively small in comparison to J^*, the number of districts that is sufficient to secure a victory.

The b-advantage game takes this story to the limit: Candidates just fight until one of them has secured b more victories than the other candidate, then the fighting stops and payoffs are distributed. A property of the b-advantage game that we exploit extensively in the following is that it has a Markov perfect equilibrium which is relatively easy to characterize. Intuitively, the equilibrium of the b-advantage game (for b sufficiently large) will look very similar to the equilibrium of the original game with a large number of districts: At those nodes where both games are defined, spending levels will be similar, and at nodes where only the original game is defined (i.e., where one candidate has more than b districts advantage), spending levels will be very close to zero in the original game and the front runner wins the next district with a probability that is close to one.

Let w_i^b be the continuation utility of a candidate who has won i elections more than his competitor, in the b-advantage game; we will call i the "state of the game". The initial value of the game is hence denoted w_0^b. Furthermore, x_i^b (y_i^b) is the campaign expenditure by the front runner (the runner-up) in the present district, given that the front-runner has won i districts more than the runner-up.

Proposition 4 shows that the transition probabilities between different states are equal in different b-advantage games (of course, as long as they are defined). For example, the probability that the winner of the first district wins the second election is exactly the same, whether $b = 2$ or $b = 35$. The reason for this result is that the continuation utilities of two different b-advantage games are linear transformations of each other (i.e., $w_i^{b'} = \alpha + \beta w_i^b$), and so players' expenditures in state i in the b'-advantage game are just β times their expenditures in the b-advantage game; hence, the expenditure ratio is unchanged. Consequently, if we want to know the probability that the winner of the first district wins also in the second district (in the original game with many districts), then all we have to do in order to answer this question is to solve the 2-advantage game.

The second result relates to the "momentum" observed in primary races: We show that the front-runner is more likely to achieve another victory than the runner-up. It is interesting to note that both results could be obtained without using the special functional form $f(x, y) = \frac{x^\alpha}{x^\alpha + y^\alpha}$, and so the momentum effect is quite robust in sequential contests.

Proposition 4. *Let $i < b < b'$.*

1. *The winning probabilities in state i are equal in the b- and the b'-advantage game.*
2. *The front-runner has a higher probability to win the next district than the runner-up: If $i > 0$, then $x_i > y_i$.*

Our next result deals with rent dissipation in primaries with many districts. For $\alpha = 1$, we show that the candidates' rent is not completely dissipated, even for $b \to \infty$, and we can explicitly calculate a lower bound for candidates' ex ante rent, which is also quite close to our numerical result. Our proof is considerably simplified by focusing on the case of $\alpha = 1$. It is intuitive that, the smaller is α, the less both players spend on campaigning, and consequently, the payoff bound should also be valid for all $\alpha \leq 1$. Our numerical results reported in the previous section confirm this intuition, however, we have not been able to show this formally. The second part of Proposition 5 provides a formal link between the b-advantage game and the original game: It shows that in the early districts of a very long primary race, strategies are the same (in the limit) as in a b advantage game.

Proposition 5. *Consider the b-advantage game for $\alpha = 1$.*

1. *The limit of w_0^b for b to infinity is greater than $\frac{2\sqrt{33}-10}{27\sqrt{33}-147} \approx 0.18377$.*
2. *For all j and k, $\lim_{J\to\infty} v_{j,k}^J = \lim_{b\to\infty} w_{2k+1-j}^b$.*

Let us interpret the second part: Consider the continuation utility of a player who has won k of the first $j-1$ elections. Since his opponent has won $j-1-k$ elections, the first player has won $k-(j-1-k) = 2k+1-j$ elections more than his opponent. The proposition says that this continuation utility is the same (provided that there are many districts, $J \to \infty$) as the continuation utility of a player in a $b-$advantage game who has the same number of districts as advantage. Note that, since the continuation utilities in a large game converge to those of the $b-$advantage game, also the strategies used in the first districts of a large primary race converge to those of the $b-$advantage game.

3.5 Relation to the Stylized Facts

Let us now compare these numerical results with two stylized facts about primaries, namely the development of winning probability over the course of the primary campaign, and the campaign expenditure profile. We will argue that the equilibrium behavior of players in our model matches these stylized facts qualitatively, and even a quantitative interpretation of the results looks reasonable. The empirical observations reported below are taken from McGillivray and Scammon (1994), Cook and McGillivray (1997) and Cook (2000).

Winning probabilities over the course of the primaries. One result of our model is that the candidate who is (through pure luck) successful in the first primaries is very likely to be successful in later primaries, too. This feature is interesting, as it is derived in a model in which both candidates have, in all districts, the same technology of converting money into electoral success. Early successes are (in our model as well as in reality), a very good predictor of who will eventually win the nomination. Consider the 7 races in both parties between 1976 and 1996 in which no competitor was a sitting U.S. president. In 5 of these races, the winner of the New Hampshire primary was the same as the eventual nominee. Our model matches this frequency approximately for $\alpha = 0.5$ (see $P_{2,1} = 0.7543$ in Table 2).[10]

However, an alternative explanation of these facts is, of course, that the eventual nominee is simply a better candidate and therefore (without any of the strategic effects that we focus on in our model) is more likely to win both in New Hampshire and in later primaries. It is therefore desirable to look at a prediction from the model that allows us to distinguish between these two theories. This prediction from our model is that the probability that the nominee wins a late primary is higher than his probability of winning an early primary, because the advantage (in terms of districts already won) later in the race is bigger than in the beginning.

The following table compares the percentage of primaries won by the eventual nominee for "early" primaries (held in February and March) to "late" primaries (April and May), in the 7 races without incumbent between 1976 and 1996. In 6 of these 7 races, the percentage of primaries won by the eventual nominee was higher after April 1st than before: These observations are largely consistent with our model in which campaign momentum increases the likelihood of victories later in the race, when there is already some advantage of the front runner.

	D-1976	R-1980	D-1984	D-1988	R-1988	D-1992	R-1996
early primaries	0.67	0.73	0.4	0.36	0.91	0.58	0.90
late primaries	0.6	0.86	0.42	1.00	1.00	1.00	1.00

Table 1: Percentage of early and late primaries won by the eventual nominee

[10]Lower values of α correspond to lower values of $P_{2,1}$. Intuitively, for $\alpha = 0$, campaign expenditures do not matter, and each candidate wins with probability 0.5 in each district; if J is large, the probability that the first district winner wins at least $J^* - 1$ of the remaining $J - 1$ districts (i.e., the necessary majority), is close to 1/2.

Campaign spending in early primaries. Malbin (1983) reports that in the 1980 Republican primaries George Bush and Ronald Reagan allocated roughly 3/4 of their respective total campaign budgets to states that held their primaries before March 31, although these states accounted for considerably less than a fifth of the delegates to the Republican convention in 1980. This indicates that candidates place a disproportionate emphasis on early primary states.[11] A similar emphasis on early primaries is also reflected in the media coverage. For example, Adams (1987) reports that the New Hampshire primary (which accounts for less than one percent of the convention delegates) attracted almost 20% of the total media coverage of all primaries in ABC, CBS, NBC and the New York Times during the 1984 presidential nomination campaign.

Our model predicts that candidates will spend a large amount of money whenever the race is tied or very close; in particular, expenditures will be large in early primary states. For example, for $\alpha = 0.5$ (the number which could approximately match the data for $P_{2,1}$) and 49 districts, each candidate's total expected campaign expenditure is 0.2739, and the expenditure in the first district alone is 0.051, or about 19% of the total expected campaign expenditure.

For comparison with the stylized fact reported above, one can also compute the number of districts after which the candidates have, on expectation, spent a fraction of 75% of their total expenditures. For 49 districts and $\alpha = 1$, more than 75% of the total expenditures are allocated in the first two districts alone (4% of all districts). For $\alpha = 0.75$, the first three districts (6%), and for $\alpha = 0.5$, the first seven districts (14%) account for this share.[12] While we do not wish to overextend these quantitative results as "matching the data" (given that our model does not capture many aspects that may be empirically important), the disproportionate share of campaign resources allocated to early primary states is consistent with our model.

[11] Unfortunately, we did not find more disaggregate data on candidates' campaign expenditures in single states. Even if financial data were available, they would not capture all resources spent, for instance, the time candidates spend visiting a district would not be included in these figures. Presumably, time is an important resource in primary campaigns, especially in small states such as New Hampshire, where a relatively large fraction of the voting population can be reached through public appearances by the candidates.

[12] One needs to exercise some caution in interpreting these results, however, because the stationarity of strategies implies that the expenditure share accruing to any *absolute* number of districts converges to some stationary level. Consequently, the expenditure share of the first n% of all districts is not invariant to changes in the total number of districts.

4 Extensions

In this section, we discuss two extensions. First, we analyze a system, in which some elections are held sequentially, followed by a number of districts that vote simultaneously. Then, in section 4.2, we consider the case where candidates are asymmetric.

4.1 Super Tuesday

While the basic structure of the U.S. primaries is sequential, it is not completely so: Some states hold their primaries at the same day as other states. The most important such day is "Super Tuesday". For example, in 1992, eight states (among them big states like Texas and Florida) held their primaries on that day. In 1996, there were two big primary Tuesdays, one week apart from each other, on which eight and seven states voted, respectively. Typically, going into the Super Tuesday election, one candidate has a clearly established lead. The race is usually conceded by one candidate shortly afterwards.

What can we say in our model framework concerning such mixed temporal arrangements? Basically, adding a simultaneous stage introduces an implicit threat of a fierce battle should the race still be close when this stage is reached. This threat can shorten the overall length of the race. Let us consider an organization in which the first two primaries are held sequentially and then, at a third period, the remaining "many" districts vote simultaneously. By "many", we mean more than $K(\alpha)$, such that, if both candidates go head-to-head into the third election period, only a mixed strategy equilibrium involving full rent dissipation exists.

In a slight variation of our model, assume here that, after every election period, the candidate who has fallen behind can withdraw from the race, and will do so, if (and only if) he has a zero continuation utility. If he concedes, then the remaining elections automatically go to his opponent.[13]

[13]Some comments on this additional assumption are in order. First, if we added a withdrawal option to the completely sequential model, nothing would change there, as even a candidate who lags behind still has a positive continuation utility (as long as it is still possible that he catches up and wins the race). Second, it would be possible to amend the model slightly such that candidates with a zero continuation utility have a *strict* incentive to take the withdrawal option. One possible way to do this is to assume slightly altruistic candidates, who maximize a weighted average of their own payoff and their opponent's payoff, because they share some policy objectives with their competitor (and would rather see their fellow party member see as president than the nominee of the other party). A candidate who concedes the race increases the opponent's payoff while leaving his own payoff unchanged; via the altruistic channel, withdrawal is a strictly preferred option, even if the altruistic effect is very weak.

An alternative to the altruistic motive is to assume that party loyalists reward, in future races, an un-

Consider the candidate's incentives after the first election has taken place. The best possible outcome in the second district for the first-round loser is to win and equalize the score. Then, both candidates are head-to-head in the final simultaneous election round. As we have shown, the complete rent will be dissipated in this final round, so that the continuation utility after a first-round loss will be as bad as giving up once the third round is reached, namely zero. This ensures that the candidate who loses the first election will give up, securing that the front runner does not have to spend any resources in the second and third district. When both candidates follow this line of reasoning, they will know that, effectively, the only contested election is the first one, and that its winner will be the winner of the whole race. Consequently, they will spend $\alpha/4$ in the first district, and the loser of that first election will immediately give up. Thus, a mixed structure that starts with few sequential elections followed by a final, simultaneous round leaves even more expected rent to the candidates than a completely sequential structure.

In view of this result, it is a bit ironic that Super-Tuesday was introduced in 1988 by Southern states with the "hope that by holding their votes on the same day, they would increase the influence of the South in selecting presidential candidates and downplay the importance of the earlier New Hampshire primaries".[14] In our model, influence is actually shifted *to* the earlier primary states and *away* from the states that participate in Super-Tuesday.

In general, which temporal organization of the primaries would be optimal, if the objective is to minimize expenditures? For the case of many districts, it appears that the organizational form just presented is actually optimal; after all, the effective fight is reduced to the very first district, and it seems plausible that any other organization that induces a contest in more districts should be more expensive. Our numerical results are consistent with this conjecture, however, we do not have a formal proof that rent dissipation is always lower in a one-district game than in a J-district game, for $J > 1$.[15]

For small J, the scheme outlined above does not work, because the lagging candidate would not give up before Super-Tuesday. For a given set of parameters, the optimal organizational form can be determined numerically; however, no general pattern emerges.

---

successful candidate who withdraws "graciously", while they punish candidates who keep fighting without much hope for a final victory. The rewards could be either financial, or informal in the form of goodwill towards a future candidacy. With this incentive system in place, continuation of a campaign has an implicit price for the first-round loser, and once his continuation utility in the campaign race drops below a certain threshold, he will give up.

[14] BBC, http://news.bbc.co.uk/1/hi/in_depth/americas/2000/us_elections/glossary/q-s/652376.stm

[15] One difficulty in proving this conjecture is that it is in general *not true* that rent dissipation in J_0 districts is lower than rent dissipation in J_1 districts, if $J_0 < J_1$. See Table 1 for counterexamples.

For $\alpha = 0.5$, the optimal organization for J =3,5,7,9 and 11 is completely sequential. However, for $\alpha = 1$ and $J = 7$, the optimal organization consists of sequential elections in the first 4 districts, followed by a simultaneous election in the remaining districts. For $\alpha = 1$ and $J = 9$, the optimal organization starts with 2 sequential single district elections, followed by three multi-district elections in 2, 3 and 2 districts, respectively. The expected rent for a candidate in this regime is 0.2378, up from 0.1839 in the completely sequential organization.[16]

4.2 Asymmetric Candidates

Until now, we have assumed that both candidates are symmetric with respect to their ability to transform campaign expenditures into electoral victories. In this section, we analyze the case that one candidate is a stronger campaigner and wins, ceteris paribus, more often. In this setting, an objective for a good primary system is also to select the stronger candidate with a high probability.

There are many possibilities to model asymmetry between the two candidates. Probably the simplest one is as follows:[17] There are ψJ districts in which candidate 1 will win irrespective of campaign expenditures, and these districts are known, so that neither candidate will spend anything there. In the $(1 - \psi)J$ other districts, the election technology is the same as in our basic model. The parameter ψ measures the advantage for candidate 1 and lies between 0 (the symmetric case studied in the basic model) and 1/2.[18]

We will restrict ourselves to the asymptotic case of $J \to \infty$, which increases the tractability. However, the general effects we identify will also be present for a small number of districts.

We will first analyze this setting for simultaneous and then for sequential primaries.

Simultaneous elections. There are ψJ uncontested districts, in which candidate 1 will win irrespective of the spending levels and where, consequently, both candidates choose a spending level of 0. Of the remaining $(1 - \psi)J$ contested districts, candidate 1 has to win (slightly more than) a share of $\frac{1}{2} - \frac{\psi}{2(1-\psi)}$, in order to guarantee that he wins in a majority of all districts.

Assume that both candidates play (asymmetric) uniform strategies in the contested

[16] The complete numerical results for all cases mentioned can be obtained from the authors upon request.
[17] We will discuss different approaches to model asymmetry between candidates below.
[18] If $\psi > 1/2$, candidate 1 wins irrespective of the outcome in the contested districts.

districts.[19] Then, candidate 1 will win the nomination, if

$$\frac{x^\alpha}{x^\alpha + y^\alpha} > \frac{1}{2} - \frac{\psi}{2(1-\psi)} \tag{9}$$

which can be simplified to $x > (1-2\psi)^{1/\alpha} y$. If the inequality sign in (9) is reversed, candidate 2 wins. From the intuition in the symmetric case, it is quite clear that the equilibrium for J very large will be in mixed strategies. In fact, this is an asymmetric all pay auction.

Let $X = (1-\psi)Jx$ and $Y = (1-\psi)Jy$ be the total expenditures of candidate 1 and 2, respectively. Let the players' equilibrium strategies be given by the distributions $\Phi_1(X)$ and $\Phi_2(Y)$, respectively. The same arguments as in symmetric all pay auctions imply that players' equilibrium strategies cannot have atoms—except possibly at 0 for a candidate with an ex ante expected payoff of 0—, and that Φ_1 and Φ_2 must be strictly increasing on $[0, (1-2\psi)^{1/\alpha}]$ and $[0, 1]$, respectively. Player 2's expected payoff is

$$\Phi_1((1-2\psi)^{1/\alpha} Y) - Y.$$

Differentiation with respect to Y and setting the result equal to zero, because player 2 must be indifferent between all his strategies, yields $\phi_1 = (1-2\psi)^{-1/\alpha}$ on the interval $[0, (1-2\psi)^{1/\alpha}]$.[20] Player 1's expected payoff is

$$\Phi_2((1-2\psi)^{-1/\alpha} X) - X.$$

Applying the same steps as above, we get $\phi_2(Y) = (1-2\psi)^{1/\alpha} < 1$ on $(0, 1]$, and hence there must be an atom on 0, so that $\Phi_2(Y) = (1-(1-2\psi)^{1/\alpha}) + (1-2\psi)^{1/\alpha} Y$ for $Y \in [0, 1]$.

Since player 2 is willing to play an atom on 0, where he is certain to lose, his ex ante rent must be 0. If player 1 chooses $X = 0$, he wins if and only if player 2 chooses $Y = 0$. This happens with probability $1-(1-2\psi)^{1/\alpha}$. Therefore, $1-(1-2\psi)^{1/\alpha}$ is player 1's expected rent in equilibrium. Expected total campaign expenditures are equal to the prize minus both candidates' expected rents, hence equal to $(1-2\psi)^{1/\alpha}$.

The equilibrium probability that player 1 wins the nomination is

$$Prob(Y < X[1-2\psi]^{-1/\alpha}) = \int_0^{(1-2\psi)^{1/\alpha}} \int_0^{x(1-2\psi)^{-1/\alpha}} \phi_2(y) dy \phi_1(x) dx = 1 - \frac{(1-2\psi)^{1/\alpha}}{2} \tag{10}$$

[19]Using the same procedure as in section 3.1, it is straightforward to show that it is optimal to play a uniform strategy, if the opponent plays a uniform strategy, and hence that a uniform equilibrium exists.

[20]Evidently, it does not make sense for player 1 to choose to spend more than the upper end of this interval, because this choice guarantees a win against all individually rational choices of player 2 (who does not spend more than 1 in any equilibrium).

This probability is greater than 1/2 and increasing in ψ, but it is always smaller than 1, even with a very large number of electoral districts.

Sequential elections. Now consider a sequential organization of primaries in the same setting. Since there are ψJ districts in which candidate 1 will win anyway, only election outcomes in contested districts provide further information. Without loss of generality, we can assume that the first ψJ elections are those in which candidate 1 wins anyway, and the remaining $(1-\psi)J$ elections are the contested ones.[21]

Hence, the initial value of the game for candidate 1 must be equal to the continuation value $v_{\psi J,\psi J}$ in the symmetric game. With the relative advantage ψ fixed and $J \to \infty$, candidate 1's absolute advantage gets very large. As we have shown in sections 3.3 and 3.4, a candidate who leads the race by a sufficiently high absolute margin receives a continuation payoff of approximately 1. Since the continuation utility is equal to the probability of winning the prize, minus the expected expenditures, this also shows that candidate 1 wins the nomination with probability close to 1. We summarize the results of this section in the following proposition:

Proposition 6. *Suppose candidate 1 wins in $(1-\psi)J$ districts uncontested, while the election technology in the remaining districts is the same as described in section 2.*

1. *If elections are held simultaneously, candidate 1's ex ante utility goes to $1-(1-2\psi)^{1/\alpha}$, and candidate 2's ex ante utility goes to 0, as $J \to \infty$. Furthermore, the probability that candidate 1 wins is $1-\frac{(1-2\psi)^{1/\alpha}}{2} < 1$.*

2. *If elections are held sequentially, candidate 1's ex ante utility goes to 1 and candidate 2's ex ante utility goes to 0 as J increases. Furthermore, the probability that candidate 1 wins goes to 1.*

To sum up, in both regimes, asymmetry between candidates reduces (expected) rent dissipation. This is a result similar to Che and Gale (2000a, 2000b). However, in our model the exogenous asymmetry is reinforced in the sequential organization through its endogenous asymmetry creating preemption effect, and so the effect of asymmetry is more forceful in a sequential organization.

Different ways of modelling asymmetry. As mentioned, there are other ways in which candidates could be asymmetric. The first one is that, in each district, the proba-

[21] No change arises, if the uncontested "elections" take place at any later time, because candidates know from the very beginning how these uncontested elections are going to turn out.

bility that candidate 1 wins is

$$\tilde{f}(x,y) = \psi + (1-\psi)f(x,y) = \psi + \frac{x^\alpha}{x^\alpha + y^\alpha}. \tag{11}$$

An interpretation of this function is that, in each district, there is a chance of $1-\psi$ that campaign expenditures are decisive (as in the basic model), and a chance of ψ that candidate 1 will win. This model of asymmetry leads to almost exactly the same results as the one analyzed above, provided that J is large. The reason is that, due to the law of large numbers, the number of districts that will be won anyway by candidate 1 (i.e., irrespective of spending) is very close to ψJ with very high probability, so the situation is very similar to the one analyzed above. The only difference is that there are no uncontested districts, as it is never known ex ante which district will be won anyway by candidate 1.

A second possible asymmetry is that the prize is worth more to candidate 1 than to candidate 2. Suppose, for example, that candidate 1 is more likely to win the general election than candidate 2 (and this is common knowledge among the candidates). If candidates care about their eventual probability of becoming president, then the prize of winning the nomination is worth more to candidate 1. Also in this case, numerical analysis has shown that the same qualitative results obtain: In a sequential organization the candidate with the higher valuation wins with a probability close to 1, while in a simultaneous organization, his winning probability is bounded away from 1.[22]

5 Discussion

Our advertising model of primary elections suggests that political parties and their candidates may have a preference for the sequential organization of primary elections, as it induces lower expected expenditures and higher expected rents than a simultaneous structure, which treats all primary states symmetrically. This cost advantage is created by a "New Hampshire Effect" that puts a disproportionate share of importance to early primary elections: While candidates compete fiercely in the first district, the asymmetry that results from the outcome of the first election generates a momentum effect which helps to reduce campaign expenditures in later states significantly. Therefore, our model also sheds light on a number of frequent empirical observations regarding the nomination process, such as the existence of campaign momentum and the allocation of a large share of campaign funds to early primary states. We can explain these facts as a direct result of the equilibrium strategies chosen by the candidates. In particular, we do not require "psychological momentum" on the part of the voters to generate these effects.

[22]Numerical results are available from the authors upon request.

Over the years, states have tended to move their primaries up to earlier dates. In 1976, the percentage of all primary voters who had voted by the end of March was 22.6%. In 1992, this number had increased almost twofold, to 42.5%. Most recently, in 2000, California moved its primary elections from May to March. A common explanation for this competition among states for early primary dates is that each state would like to see those candidates being nominated which best represent the preferences of its residents. (Think, for example, about the ethanol subsidy that might not exist if Iowa was not an early caucus state.) This explanation, of course, requires that early primary outcomes are in fact influential for the future direction of the race. Whether or not this can be the case in a theoretical model critically depends on whether candidates are included as decision making agents or not. In a model that abstracts from candidate behavior—specifically, in Dekel and Piccione's (2000) work on voting as an information transmission channel—the equilibria of sequential elections are equivalent to those of simultaneous elections. Consequently, the position of any state in a sequence of elections is irrelevant. Our model of candidate behavior, on the other hand, suggests that there exists a motivation for moving up state primary dates, as the temporal position of elections does matter.

One important consequence of these temporal shifts becomes clear in our model: If many states compete for primary dates that are early in the election year, the sequential primary system will eventually be transformed into a simultaneous one. A sequential organization basically requires that candidates can observe the outcome of an election before they commit resources to further campaigns. If elections are separated by short time intervals only, it seems unlikely that this is still a practical possibility. It can be expected that the disadvantages of simultaneous primaries, such as high rent dissipation rates and poor selection properties, would transpire in a formally sequential, but temporally dense environment, too.

The problem of inefficient competition for early dates has spurred a discussion on alternative primary designs in the organization of the secretaries of state of the 50 U.S. states. One particular proposal is a system of regional primaries, in which the caucus and primary dates in Iowa and New Hampshire would remain unchanged; the other states are divided into four groups (Northeast, South, Midwest and West). All states in a group hold their primaries at the same day, starting with the first group one month after New Hampshire. The next group follows after an interval of one month, and so on. Which group goes first would rotate in a 16 year cycle. The stated reason for the cycle is fairness: Since the region with the first group primary is perceived as decisive, every state would enjoy this advantage once during the cycle. This reform proposal bears resemblance to the mixed system discussed in section 4.1, in which two sequential elections are followed by simultaneous elections in a group consisting of all remaining districts. We discussed there when such a system is optimal with respect to the equilibrium campaign levels.

Like every model, ours had to abstract from a number of issues that are certainly important in reality. In analyzing and explaining candidate behavior, we had to abstract from voting as a strategic decision. Our reduced form approach to capture the effects of campaign effort choice by the candidates prevents us from studying those informational aspects of elections that concern voters' uncertainty about candidates' qualifications. Other informational issues, however, concern the possibility of the candidates' uncertainty about each other's characteristics. For example, in the case of asymmetric campaign strength, candidates in our model are informed of the value of the parameter ψ. If they only possess a prior belief about this value, but do not know its realization, a sequential arrangement of elections would give rise to learning on the part of the candidates. In a simultaneous organization, on the other hand, no learning would take place. It is unclear how this kind of uncertainty would affect our results.

One simplification in our model, whose relaxation, we believe, would not alter our results in any fundamental way, is the assumption that all electoral districts are of equal size. The fact that New Hampshire, as the first primary state, is small compared to later states still has interesting implications. Because of the extensive campaign effort that arises in early primaries, placing small states at the beginning of the sequence appears to provide further evidence that the observed organization is indeed chosen in order to keep overall campaign costs as small as possible. Interestingly, an argument which is sometimes put forward by supporters of the existing primary system is that, given the perceived importance of early election outcomes, starting with a small state allows candidates with smaller budgets to make up for their lack of financial resources by increased personal efforts. Such a substitution would no longer be feasible in a larger state, regional group, or a one-day primary system.

An interesting, but rather challenging, extension of our model would be to consider what happens if the candidates face additional hard budget constraints (they maximize the same objective function as in our model, but cannot spend more than b_i, say).[23] Our

[23]The reason why the sequential case is difficult to analyze with this extension is that the number of victories up to a certain district does not uniquely capture the state of the game, but rather the sequence, in which victories were achieved, matters. Consider the 3 district case. With symmetric candidates, both of them will spend the same in the first district, and in the second district, they will presumably spend different amounts. If the budget constraint is binding in the third district (at least for the candidate who has less money – otherwise, the constraint does not matter at all), then candidates will spend different

result that total expected campaign spending is lower in a sequential structure than a simultaneous one is still very likely to hold in this new setting. More interesting, will expenditures be lower or higher, if players face additional hard budget constraints? The answer to this question might not be as obvious and clearcut as it appears. Che and Gale (2000) have shown that spending limits in lobbying models might actually increase the players' equilibrium spending if their valuations are asymmetric. Also, if players are asymmetric with respect to their hard budget constraint (and otherwise equal), will the candidate who has the advantage win "almost always" provided that there are many districts, as in section 4.2? This is an interesting question, because primary candidates may differ substantially in their spending possibilities. For example, in the 2000 Republican nomination race, John McCain was victorious in New Hampshire, but eventually lost the race to George W. Bush; this turn of events can, to some extent, be accounted for by McCain's lack of sufficient resources to fund a prolonged campaign against his wealthier opponent. These questions are left for further research.

6 Appendix: Proofs

6.1 Proof of Proposition 1

(a) Existence of a pure strategy SUCE has been shown in the main text; here, we show uniqueness. Consider an equilibrium with spending profiles (x, y); these profiles need not be uniform. Fix some district j and let the spending profile in the other districts be denoted x_{-j}, y_{-j}. Given (x_{-j}, y_{-j}), let $\widetilde{P}_j$ be the probability that district j is pivotal. Note that this probability is the same for both candidates.

Since (x, y) is an equilibrium, it must be true that x_j maximizes

$$\frac{x_j^\alpha}{x_j^\alpha + y_j^\alpha}\widetilde{P}_j - x_j,$$

which yields the first-order condition

$$\frac{\alpha x_j^{\alpha-1} y_j^\alpha}{\left(x_j^\alpha + y_j^\alpha\right)^2}\widetilde{P}_j - 1 = 0. \tag{12}$$

A similar first-order condition is obtained for candidate 2:

$$\frac{\alpha y_j^{\alpha-1} x_j^\alpha}{\left(x_j^\alpha + y_j^\alpha\right)^2}\widetilde{P}_j - 1 = 0. \tag{13}$$

amounts in the third district. Hence, it matters (for the spending in the third district) whether candidate 1 or 2 won the first district.

To satisfy (12) and (13) simultaneously, we need $x_j = y_j$, so that district j is won with equal probability by either candidate, regardless of the spending profile in the other districts. Since this reasoning can be applied to all districts, each district is won with probability 1/2, and consequently

$$\widetilde{P}_j = \widetilde{P} = \binom{J-1}{J^*-1}\left(\frac{1}{2}\right)^{J-1} \quad \forall j.$$

Equations (12) and (13) can now be solved uniquely for

$$x_j = y_j = \frac{1}{4}\alpha\binom{J-1}{J^*-1}\widetilde{P},$$

which coincides with (4).

(b) The argument why a pure strategy equilibrium does not exist for $J > K(\alpha)$ is mainly given in the main text. It remains to be shown formally that the right hand side of inequality (5),

$$\frac{2^J}{J\binom{J-1}{(J-1)/2}} = \frac{2^J\left(\frac{J-1}{2}!\right)^2}{J!} \tag{14}$$

is decreasing in J and goes to 0 for $J \to \infty$. Going from J to $J+2$ multiplies the numerator of the right hand side of (14) by $4\left(\frac{J+1}{2}\right)^2 = (J+1)^2$, and multiplies the denominator by $(J+1)(J+2)$. Hence, the value of the fraction decreases.

To calculate an approximation of $K(\alpha)$, use Sterling's approximation formula ($n! \approx (n/e)^n\sqrt{2\pi n}$) to obtain

$$\frac{2^J\left[\left(\frac{J-1}{2e}\right)^{\frac{J-1}{2}}\sqrt{2\pi\frac{J-1}{2}}\right]^2}{\left(\frac{J}{e}\right)^J\sqrt{2\pi J}} = \frac{e^{-(J-1)}(J-1)^J\sqrt{2\pi}}{J^Je^{-J}\sqrt{J}} = \sqrt{2\pi}e\left(\frac{J-1}{J}\right)^J\frac{1}{\sqrt{J}} \approx \sqrt{\frac{2\pi}{J}}$$

which goes to zero for $J \to \infty$. Solving for J yields the approximation $K(\alpha) \approx \frac{2\pi}{\alpha^2}$ mentioned in the text.

We now show that a mixed strategy SUCE exists. Consider first the game that arises when we restrict players to use only uniform strategies. This game has a symmetric Nash equilibrium (hence a SUCE), because v_1 and v_2 satisfy the sufficient conditions for equilibrium existence in discontinuous games in Dasgupta and Maskin (1986). It remains to be shown that this SUCE of the restricted game is also an equilibrium of the original game, in which players are free to choose non-uniform strategies as well. Lemma 1 below shows that, if player 1 (resp., 2) plays a uniform strategy, then any non-uniform strategy is dominated for player 2 (resp., 1). Hence, if there is no profitable deviation from the equilibrium candidate using a uniform strategy, there also cannot be a profitable deviation using non-uniform strategies; hence, the SUCE is also an equilibrium of the original game.

Lemma 1. *Suppose player 2 plays a uniform strategy. Then, any non-uniform strategy $x = (x_1, x_2, \ldots, x_J)$ is dominated for player 1 by the uniform strategy with the same total expenditure,* $X = (\frac{\sum x_i}{J}, \frac{\sum x_i}{J}, \ldots, \frac{\sum x_i}{J})$.

Proof. Fix player 2's strategy to be a uniform campaign strategy with $y_j = y$ for all j; y can be drawn from a random distribution. Suppose that player 1 plays, with positive probability, some pure strategy with $x_k > x_l$ for some districts k and l. Consider a deviation which leaves expenditures in all other districts unchanged and equates the campaign levels in k and l so that total expenditures do not change:

$$\widetilde{x}_k = \widetilde{x}_l = (x_k + x_l)/2 \;\equiv\; \widetilde{x}.$$

Let Q_n, $n \in \{0, 1, 2\}$, denote the probability that player 1 wins exactly n districts among the two districts k and l, when using strategy x. Similarly, let $\widetilde{Q}_n$ denote the probability of winning exactly n districts among the two districts k and l, when using strategy $\widetilde{x}$. Finally, let P_n, $n \in \{0, \ldots, J-2\}$, be the probability of winning *exactly* n out of the remaining $J-2$ districts.

Player 1's gain from changing to the new strategy is

$$\Delta E u_1 = E\left\{(\widetilde{Q}_1 + \widetilde{Q}_2 - Q_1 - Q_2)P_{J^*-1} + (\widetilde{Q}_2 - Q_2)P_{J^*-2}\right\},$$

where the expectation is taken with respect to y. Since P_{J^*-1} and P_{J^*-2} do not change when switching from strategy x to $\widetilde{x}$, a sufficient condition for $\Delta E u_1$ to be positive is that both $E\left\{(\widetilde{Q}_1 + \widetilde{Q}_2 - Q_1 - Q_2)\right\} \geq 0$ and $E\left\{\widetilde{Q}_2 - Q_2\right\} \geq 0$.

Since x^α is concave due to $\alpha \in (0, 1]$, we have $\widetilde{x}^\alpha - x_l^\alpha > x_k^\alpha - \widetilde{x}^\alpha$ and $(\widetilde{x}^\alpha)^2 > x_k^\alpha x_l^\alpha$. Therefore,

$$E(\widetilde{Q}_1 + \widetilde{Q}_2 - Q_1 - Q_2) = E\left(y^{2\alpha}\frac{y^\alpha(\widetilde{x}^\alpha - x_k^\alpha + \widetilde{x}^\alpha - x_l^\alpha) + (\widetilde{x}^\alpha)^2 - x_k^\alpha x_l^\alpha}{(\widetilde{x}^\alpha + y^\alpha)(\widetilde{x}^\alpha + y^\alpha)(x_k^\alpha + y^\alpha)(x_l^\alpha + y^\alpha)}\right) > 0$$

and

$$E(\widetilde{Q}_2 - Q_2) = E\left(y^\alpha \widetilde{x}^\alpha \frac{x_k^\alpha(\widetilde{x}_l^\alpha - x_l^\alpha) - x_l^\alpha(x_k^\alpha - \widetilde{x}_k^\alpha)}{(\widetilde{x}_k^\alpha + y^\alpha)(\widetilde{x}_l^\alpha + y^\alpha)(x_k^\alpha + y^\alpha)(x_l^\alpha + y^\alpha)}\right) > 0.$$

This shows that strategy $\widetilde{x}$ dominates strategy x.

To prove that the uniform strategy X dominates x, consider the following algorithm: Step 1: Start with x, and select the two districts with the greatest and the smallest spending (say, k and l). Step 2: Replace the spending levels in these districts by their mean, $(x_k + x_l)/2$; by the arguments presented above, this strategy will be better for player 1 than the initial strategy. Step 3: If the new strategy profile $(\widetilde{x})$ is not yet uniform, go back to Step 1 and repeat the procedure for the two districts with the greatest and the smallest spending level under $\widetilde{x}$.

Clearly, this algorithm converges to the uniform strategy X, and since utility increases with every cycle, it is proved that the uniform strategy X dominates x. □

(c) To prove that the rent is completely dissipated in the mixed strategy SUCE, we need the following three lemmata.

Lemma 2. *If* Λ *is a symmetric UCE in* G_J^{sim}*, then* $\inf supp(\Lambda) > 0$*.*

Proof. First, observe that 0 cannot be played with positive probability, because otherwise, a player could increase his expected payoff strictly by shifting weight from 0 to a sufficiently small, but positive number a. Second, if 0 is not played with positive probability, then

$$\left.\frac{\partial Eu\left(x \mid \Lambda\right)}{\partial x}\right|_{x=0} = -J < 0,$$

so that, by continuity, positive campaign levels very close to zero are dominated by a zero bid. Hence, $\inf supp(\Lambda) > 0$. $\square$

Lemma 3. *The function* F*, as defined in (2) satisfies the following monotone likelihood ratio property:*

$$\frac{F\left(x', y'\right)}{F\left(x, y'\right)} > \frac{F\left(x', y\right)}{F\left(x, y\right)}$$

for all $x' > x$ *and* $y' > y$*.*

Proof. We have

$$\frac{F\left(x', y\right)}{F\left(x, y\right)} = \frac{F(x,y) + \int_x^{x'} F_x(t,y)dt}{F(x,y)} = 1 + \int_x^{x'} \frac{F_x(t,y)}{F(x,y)}dt,$$

Similarly, $\frac{F(x',y')}{F(x,y')} = 1 + \int_x^{x'} \frac{F_x(t,y')}{F(x,y')}dt$. Thus, it is sufficient to show that F_x/F increases in y, or equivalently

$$F_{xy}F - F_x F_y > 0. \tag{15}$$

Differentiating F with respect to x, we obtain

$$
\begin{aligned}
F_x &= \sum_{k=J^*}^{J}\binom{J}{k}\left[\alpha k\frac{x^{\alpha k-1}y^{\alpha(J-k)}}{(x^\alpha+y^\alpha)^J}-J\alpha x^{\alpha-1}\frac{x^{\alpha k}y^{\alpha(J-k)}}{(x^\alpha+y^\alpha)^{J+1}}\right] \\
&= \frac{\alpha}{(x^\alpha+y^\alpha)^{J+1}}\sum_{k=J^*}^{J}\binom{J}{k}\left[kx^{\alpha k-1}y^{\alpha(J-k+1)}-(J-k)\,x^{\alpha(k+1)-1}y^{\alpha(J-k)}\right] \\
&= \frac{\alpha}{(x^\alpha+y^\alpha)^{J+1}}\left[\binom{J}{J^*}J^*x^{\alpha J^*-1}y^{\alpha(J-J^*+1)}-\binom{J}{J^*}(J-J^*)\,x^{\alpha(J^*+1)-1}y^{\alpha(J-J^*)}\right. \\
&\quad +\binom{J}{J^*+1}(J^*+1)\,x^{\alpha(J^*+1)-1}y^{\alpha(J-J^*)}-\binom{J}{J^*+1}(J-J^*-1)\,x^{\alpha(J^*+2)-1}y^{\alpha(J-J^*-1)} \\
&\quad \vdots \\
&\quad \left.+\binom{J}{J}Jx^{\alpha J-1}y^{\alpha J^*}-0\right] \\
&= \frac{\alpha J^*}{(x^\alpha+y^\alpha)^{J+1}}\binom{J}{J^*}x^{\alpha J^*-1}y^{\alpha J^*}. \qquad (16)
\end{aligned}
$$

(The second and third term in the summation cancel out, the fourth and fifth, etc.). In a similar way, we get

$$F_y=-\frac{\alpha J^*}{(x^\alpha+y^\alpha)^{J+1}}\binom{J}{J^*}x^{\alpha J^*}y^{\alpha J^*-1}. \qquad (17)$$

Differentiating (16) with respect to y yields

$$F_{xy}=\frac{(\alpha J^*)^2}{(x^\alpha+y^\alpha)^{J+2}}\binom{J}{J^*}(xy)^{\alpha J^*-1}(x^\alpha-y^\alpha). \qquad (18)$$

Using (16)–(18), we can rewrite (15) as

$$
\begin{aligned}
F_{xy}F-F_xF_y = {} & \frac{(\alpha J^*)^2}{(x^\alpha+y^\alpha)^{J+2}}\binom{J}{J^*}(xy)^{\alpha J^*-1}(x^\alpha-y^\alpha)\cdot\sum_{k=J*}^{J}\binom{J}{k}\frac{x^{\alpha k}y^{\alpha(J-k)}}{(x^\alpha+y^\alpha)^J} \\
&+\frac{\alpha J^*}{(x^\alpha+y^\alpha)^{J+1}}\binom{J}{J^*}x^{\alpha J^*-1}y^{\alpha J^*}\cdot\frac{\alpha J^*}{(x^\alpha+y^\alpha)^{J+1}}\binom{J}{J^*}x^{\alpha J^*}y^{\alpha J^*-1}.
\end{aligned}
$$

Collecting common terms, this expression reduces to

$$
\begin{aligned}
& \frac{(\alpha J^*)^2}{(x^\alpha+y^\alpha)^{2(J+1)}}\binom{J}{J^*}(xy)^{\alpha J^*-1}\left[(x^\alpha-y^\alpha)\cdot\sum_{k=J*}^{J}\binom{J}{k}x^{\alpha k}y^{\alpha(J-k)}+\binom{J}{J^*}x^{\alpha J^*}y^{\alpha J^*}\right] \\
= {} & \frac{(\alpha J^*)^2}{(x^\alpha+y^\alpha)^{2(J+1)}}\binom{J}{J^*}(xy)^{\alpha J^*-1}x^{\alpha(J+1)}>0. \qquad \square
\end{aligned}
$$

Lemma 4. *Suppose $J > K(\alpha)$. For each $x > 0$ there exists $\vartheta(x) > 1$ such that*

$$\frac{F(x', x)}{F(x, x)} > \frac{x'}{x},$$

for all $x' \in (x, \vartheta(x)x)$. Furthermore, let Λ be a symmetric UCE. There exists a constant $\bar{\vartheta} > 1$ such that $\vartheta(x) \geq \bar{\vartheta}$ for all $x \in supp(\Lambda)$.

Proof. To prove the first statement, rewrite the inequality in the Lemma as

$$\frac{F(x', x)}{x'} > \frac{F(x, x)}{x}.$$

We will prove that

$$\left. \frac{\partial}{\partial x} \frac{F(x, y)}{x} \right|_{y=x} > 0. \tag{19}$$

It is easily verified that the left-hand side of (19) is continuous in x for all $x > 0$. The result then follows. To prove (19), we show that $xF_x(x, x) - F(x, x) > 0$. Using (16) and $F(x, x) = \frac{1}{2}$, we get

$$\begin{aligned} xF_x(x, x) - F(x, x) &= \frac{\alpha J^*}{(2x^\alpha)^{J+1}} \binom{J}{J^*} x^{2\alpha J^*} - \frac{1}{2} \\ &= 2^{-(J+1)} \frac{\alpha J!}{(J^* - 1)!^2} - \frac{1}{2} > 0, \end{aligned}$$

which is true, because $\frac{\alpha J!}{(J^*-1)!^2} > 2^J$ by Lemma 1 for $J > K(\alpha)$.

To prove the second statement, recall that $\inf supp(\Lambda) > 0$ by Lemma 1, and $x \notin supp(\Lambda)$ if $x > 1$. Hence, there exists a compact set $W \subset (0, 1]$ such that $supp(\Lambda) \subseteq W$, which bounds $\left. \frac{\partial}{\partial x} \frac{F(x,y)}{x} \right|_{y=x}$ away from zero on $supp(\Lambda)$. This implies that $\vartheta(x)$ can be chosen to be bounded away from 1 for all $x \in supp(\Lambda)$ $\square$

Proof. We can now prove that the complete rent is dissipated in the mixed strategy SUCE, if $J > K(\alpha)$. Let $\bar{x} = \inf\{\Lambda(x) > 0\} > 0$. Assume that Λ has a density λ at $\bar{x}$, i.e. $Prob(x = \bar{x}) = 0$. This is just an assumption for convenience of notation; we will indicate in footnotes how to adjust the proof if Λ has an atom at $\bar{x}$.

Since the player is willing to play $\bar{x}$ in equilibrium, it must yield utility u:

$$\int_{\bar{x}}^{1} F(\bar{x}, x) d\Lambda(x) - \bar{x} = u \tag{20}$$

Splitting the integral on the left hand side, we get[24]

$$\lambda(\bar{x})\varepsilon F(\bar{x}, \bar{x}) + \int_{\bar{x}+\varepsilon}^{1} F(\bar{x}, x) d\Lambda(x) - \bar{x} + O(\varepsilon^2) = u$$

[24] If Λ has an atom of size λ_0 at $\bar{x}$, we would have to replace $\lambda(\bar{x})\varepsilon$ by λ_0 in the following formula.

where $O(\varepsilon^2)$ is a second order term ignored in the following.[25] Hence, we can solve for

$$\lambda(\bar{x})\varepsilon = \frac{u + \bar{x} - \int_{\bar{x}+\varepsilon}^{1} F(\bar{x}, x)d\Lambda(x)}{F(\bar{x}, \bar{x})} \tag{21}$$

Playing $\bar{x} + \epsilon$ cannot give a candidate a higher utility than u, so we have (again splitting the integral and ignoring second order effects)

$$\lambda(\bar{x})\varepsilon F(\bar{x} + \varepsilon, \bar{x}) + \int_{\bar{x}+\varepsilon}^{1} F(\bar{x} + \varepsilon, x)d\Lambda(x) - (\bar{x} + \varepsilon) \leq u \tag{22}$$

Substituting from (21) and rearranging slightly, we have

$$\int_{\bar{x}+\varepsilon}^{1} \left[F(\bar{x} + \varepsilon, x) - \frac{F(\bar{x} + \varepsilon, \bar{x})}{F(\bar{x}, \bar{x})} F(\bar{x}, x) \right] d\Lambda(x) \leq u + \bar{x} + \varepsilon - \frac{F(\bar{x} + \varepsilon, \bar{x})}{F(\bar{x}, \bar{x})}(u + \bar{x}) \tag{23}$$

The integrand on the left hand side is positive by Lemma 3, and the right hand side is (using Lemma 4) smaller than $-\frac{\varepsilon u}{\bar{x}}$, which is negative and a first order term if $u > 0$. This shows that (23) cannot hold for positive u, and so $u = 0$. $\square$

6.2 Proof of Proposition 2

The main arguments are in the text. It remains to be shown that the following is true.

Lemma 5. *(i)* $v_{j,k} \geq v_{j,k-1}$

(ii) $v_{j,k} \geq v_{j+1,k}$

Proof. Let $\Delta_{j+1,k} \equiv v_{j+1,k+1} - v_{j+1,k}$. Write the continuation value at state (j, k) as

$$v_{j,k} = \max_{x \geq 0} \left(v_{j+1,k} + f(x, y)\Delta_{j+1,k} - x \right), \tag{24}$$

letting $v_{j,k} = 1$ for all $j \geq J^*$ and $k \geq J^*$, and $v_{j,k} = 0$ for all $j \geq J^*$ and $k \leq j - J^*$.

We will prove (i) by induction on j. Obviously (i) is true for $j = J + 1$ and all k, as $v_{J+1,k} = 1$ for $k \geq J^*$, and $v_{J+1,k} = 0$ for $k < J^*$. Assuming (i) is true for some $j + 1$, we have $\Delta_{j+1,k} \geq 0$. We will now show that (i) also holds for j.

Since $x = 0$ is a feasible choice on the right hand side of (24),

$$v_{j,k} \geq v_{j+1,k} + f(0, y)\Delta_{j+1,k} \geq v_{j+1,k}. \tag{25}$$

Next, observe that for all $x \geq 0$ we have

$$v_{j+1,k-1} + f(x, y)\Delta_{j+1,k-1} - x \leq v_{j+1,k-1} + \Delta_{j+1,k-1} = v_{j+1,k}.$$

[25] When we choose ε sufficiently small, the first order effects derived in the following will dominate any second order effect.

For the maximum of the left-hand side taken over $x \geq 0$, it must therefore be true that

$$v_{j,k-1} \leq v_{j+1,k} \tag{26}$$

Combining (25) and (26), we obtain that $v_{j,k} \geq v_{j+1,k} \geq v_{j,k-1}$, so that $v_{j,k} \geq v_{j,k-1}$, proving (i) for j. This completes the proof by induction for (i), and (ii) then follows immediately from inequality (25). □

6.3 Proof of Proposition 4

Proof. 1. For given b, the continuation values are linked by

$$w_i^b = w_{i-1}^b + (w_{i+1}^b - w_{i-1}^b)\frac{(x_i^b)^\alpha}{(x_i^b)^\alpha + (y_i^b)^\alpha} - x_i^b \text{ for all } i \geq 0, \tag{27}$$

an analogous condition for the runner-up

$$w_{-i}^b = w_{-(i+1)}^b + (w_{-(i-1)}^b - w_{-(i+1)}^b)\frac{(y_i^b)^\alpha}{(x_i^b)^\alpha + (y_i^b)^\alpha} - y_i^b \text{ for all } i \geq 0, \tag{28}$$

and the terminal conditions $w_b^b = 1$ and $w_{-b}^b = 0$. From the two first order conditions on stage i, we have

$$\frac{x_i^b}{y_i^b} = \frac{w_{i+1}^b - w_{i-1}^b}{w_{-(i-1)}^b - w_{-(i+1)}}. \tag{29}$$

Now suppose that $\{w_i^{b_0}\}_{i=-b_0..b_0}$ are the continuation values for a b_0-advantage game, and let $b_1 < b_0$. We claim that $w_i^{b_1} = \gamma + \delta w_i^{b_0}$, where γ and δ are determined by $w_{b_1}^{b_1} = \gamma + \delta w_{b_1}^{b_0} = 1$ and $w_{-b_1}^{b_1} = \gamma + \delta w_{-b_1}^{b_0} = 0$. It is easy to see that $\frac{x_i^{b_1}}{y_i^{b_1}} = \frac{\delta x_i^{b_0}}{\delta y_i^{b_0}}$ will be unchanged by this linear transformation, and then (27) and (28) continue to hold for $w_i^{b_1} = \gamma + \delta w_i^{b_0}$.

2. For all i, the following inequality must hold

$$w_i + w_{-i} < w_{i+1} + w_{-(i+1)}, \tag{30}$$

because of the following consideration: $w_i + w_{-i}$ is equal to the prize, 1, minus the expected future expenditures by both candidates; the latter can be split into the expected expenditures that will be incurred until state $i+1$ is reached for the first time, plus the expected expenditures following that event. Similarly, $w_{i+1} + w_{-(i+1)}$ is equal to the prize minus the expected expenditures by both candidates following state $i+1$. Hence, the difference between $w_{i+1} + w_{-(i+1)}$ and $w_i + w_{-i}$ is equal to the expected expenditures to be made between the time when state i is reached and the time when state $i+1$ is reached for the first time, and is strictly positive. Equation (30) is equivalent to $w_{i+1} - w_{i-1} > w_{-(i-1)} - w_{-(i+1)}$. Together with (29), this implies $x_i > y_i$, and hence that the player who has an advantage will win the next district with a probability that is greater than 1/2.[26] $\square$

6.4 Proof of Proposition 5

For the proof of the first part, we need 2 lemmata:

Lemma 6. *In any b-advantage game, $w_{i+1} > 1 - 2w_{-i}$.*

Proof. If player 1 is, at present, i districts ahead, he maximizes

$$\max_{x_i} \frac{x_i}{x_i + y_i} w_{i+1} + \frac{y_i}{x_i + y_i} w_{i-1} + -x_i \tag{31}$$

Player 2 maximizes

$$\max_{y_i} \frac{y_i}{x_i + y_i} w_{-(i-1)} + \frac{x_i}{x_i + y_i} w_{-(i+1)} - y_i \tag{32}$$

Solving the two first order conditions, we get (29), and using this in the first order conditions yields

$$x_i = \frac{(w_{-(i-1)} - w_{-(i+1)})(w_{i+1} - w_{i-1})^2}{[(w_{i+1} - w_{i-1}) + (w_{-(i-1)} - w_{-(i+1)})]^2} \tag{33}$$

and

$$y_i = \frac{(w_{-(i-1)} - w_{-(i+1)})^2(w_{i+1} - w_{i-1})}{[(w_{i+1} - w_{i-1}) + (w_{-(i-1)} - w_{-(i+1)})]^2} \tag{34}$$

Substituting in the objective function of player 1, we find

$$w_i = \frac{w_{i+1} - w_{i-1}}{[(w_{-(i-1)} - w_{-(i+1)}) + (w_{i+1} - w_{i-1}))]} w_{i+1} + \frac{w_{-(i-1)} - w_{-(i+1)}}{[(w_{-(i-1)} - w_{-(i+1)}) + (w_{i+1} - w_{i-1}))]} w_{i-1} - \frac{(w_{-(i-1)} - w_{-(i+1)})(w_{i+1} - w_{i-1})^2}{[(w_{i+1} - w_{i-1}) + (w_{-(i-1)} - w_{-(i+1)})]^2} \tag{35}$$

Consequently, we can calculate the difference between w_{i+1} and w_i as

$$\begin{aligned} w_{i+1} - w_i &= \frac{(w_{i+1} - w_{i-1})(w_{-(i-1)} - w_{-(i+1)})}{[(w_{-(i-1)} - w_{-(i+1)}) + (w_{i+1} - w_{i-1}))]} + \frac{(w_{i+1} - w_{i-1})^2(w_{-(i-1)} - w_{-(i+1)})}{[(w_{-(i-1)} - w_{-(i+1)}) + (w_{i+1} - w_{i-1}))]^2} \\ &< 2(w_{-(i-1)} - w_{-(i+1)}) \end{aligned}$$

[26] While we use the parametric form (i.e., $f(x, y) = x^\alpha/(x^\alpha + y^\alpha)$ in (29)), it is easy to see that $w_{i+1} - w_{i-1} > w_{-(i-1)} - w_{-(i+1)}$ implies $x_i > y_i$ for any $f()$ that is symmetric and concave in x.

Hence,

$$\begin{aligned}
w_2 - w_1 &\leq 2(w_0 - w_{-1}) \\
w_3 - w_2 &\leq 2(w_{-1} - w_{-2}) \\
&\vdots \\
w_b - w_{b-1} &\leq 2(w_{-(b-2)} - w_{-(b-1)})
\end{aligned}$$

Summing up all inequalities starting from the ith one, we have $w_b - w_i = 1 - w_i < 2(w_{-(i-1)} - w_{-(b-1)}) < 2w_{-(i-1)}$, as claimed. □

Lemma 7. *For the 2-advantage game,* $w_0 = \frac{\sqrt{33}-5}{4}$, $w_{-1} = \frac{(\sqrt{33}-5)^3}{64}$ *and* $w_1 = \frac{4}{27}\frac{523503\sqrt{33}-3007199}{(9\sqrt{33}-49)^3}$

Proof. From (27), (28), (33) and (34), the following conditions must hold in the $b = 2$-advantage game: $w_{-2} = 0$, $w_2 = 1$ and

$$w_0 = \frac{1}{2}w_1 + \frac{1}{2}w_{-1} - \frac{w_1 - w_{-1}}{4} = \frac{1}{4}w_1 + \frac{3}{4}w_{-1} \tag{36}$$

$$\begin{aligned}
w_1 = \frac{w_2 - w_0}{w_2 - w_{-2}}w_2 + \frac{w_0 - w_{-2}}{w_2 - w_{-2}}w_0 - \frac{(w_2 - w_0)^2(w_0 - w_{-2})}{w_2 - w_{-2}} \\
= 1 - 2w_0 + 3(w_0)^2 - (w_0)^3
\end{aligned} \tag{37}$$

and

$$\begin{aligned}
w_{-1} = \frac{w_2 - w_0}{w_2 - w_{-2}}w_{-2} + \frac{w_0 - w_{-2}}{w_2 - w_{-2}}w_0 - \frac{(w_2 - w_0)(w_0 - w_{-2})^2}{w_2 - w_{-2}} \\
= (w_0)^3
\end{aligned} \tag{38}$$

These equations have the solution given above. □

Proof of the first part of Proposition 5: From Proposition 4, we know that for any b, w_i^b can be written $\gamma(b)w_i^2 + \delta(b)$, for any i; we write $\gamma(b)$ and $\delta(b)$ in order to show that γ and δ depend on b. Since all $w_i^b \geq 0$, $w_{-2}^b = \gamma(b)w_{-2}^2 + \delta(b) = \delta(b) \geq 0$, δ must be nonnegative. From Lemma 6 with $i = 1$, we have $w_2^b = \gamma(b)w_2^2 + \delta(b) \geq 1 - 2[\gamma(b)w_{-1}^2 + \delta(b)]$ and hence $\gamma\left(1 + 2\left(\frac{\sqrt{33}-5}{4}\right)^3\right) + 3\delta \geq 1$.

Therefore, w_0^b must be at least as large as the solution of the following constrained optimization problem:

$$\begin{aligned}
\min_{\gamma,\delta} \gamma\frac{\sqrt{33}-5}{4} + \delta, \text{ subject to} \\
\delta \geq 0 \\
\gamma\left(1 + 2\left(\frac{\sqrt{33}-5}{4}\right)^3\right) + 3\delta \geq 1
\end{aligned}$$

This problem has the solution $\delta = 0, \gamma = \frac{1}{1+2\left(\frac{\sqrt{33}-5}{4}\right)^3}$, and hence a lower bound for w_0^b is

$$\frac{\frac{\sqrt{33}-5}{4}}{1+2\left(\frac{\sqrt{33}-5}{4}\right)^3} = \frac{2\sqrt{33}-10}{27\sqrt{33}-147} \approx 0.18377.$$

Remark: Going through the same steps, but using the 3-advantage game, one gets a slightly better approximation for w_0^b, 0.183847. This is already virtually indistinguishable from the numerical results.

2. Denote $\lim_{J\to\infty} v_{j,2k-j-1} = \nu_k$ and $\lim_{b\to\infty} w_k^b = w_k^\infty$. We claim that $\nu_k = w_k^\infty$.

Fix any b, and consider the continuation payoffs in the initial stages of a long original game, at those nodes where no candidate has achieved an advantage of more than b districts over his opponent. Clearly, these continuation payoffs can be interpreted as those of a modified b-advantage game with payoffs ν_b for the first candidate to reach an advantage of b districts, and ν_{-b} for the loser.

We know from the proof of proposition 4 that the continuation payoffs in all b advantage games are linear transformations of each other, so that we must have

$$\nu_k = \gamma + \delta w_k^\infty$$

The claim is that $\gamma = 0$ and $\delta = 1$. Note first that $\delta \geq 0$, since it must be at least weakly better to win more districts. We claim that $\gamma = 0$. Since $\lim_{k\to-\infty} \nu_k = 0$ (i.e., if the disadvantage becomes too large, the continuation utility goes to zero), and $\lim_{k\to-\infty} \nu_k = \gamma + \delta \cdot 0 = \gamma$, γ must be equal to zero.

Is it possible that $\delta \neq 1$? Note that, at each node, the expenditures in the limit of the original game are δ times the corresponding expenditures in the limit of the $b-$advantage game. Since the transition probabilities between two nodes are exactly the same for all b-advantage games, this implies that total expected spending in the limit of the original game is δ times the spending in the limit of the b-advantage games. If $\delta < 1$, then the candidates spend, in the limit of a very large primary game, in each district only δ times what they spend in the b-advantage game (for $b \to \infty$). Note that, since candidates are symmetric at the begin of the game,

$$\nu_0 = \frac{1 - \text{expected total campaign spending}}{2}.$$

Hence, if $\delta < 1$, then $\nu_0 > w_0^\infty$, since expected campaign spending is lower than in the limit of the $b-$advantage game. However, since also $\nu_0 = \delta w_0^\infty < w_0^\infty$, this leads to a contradiction. Similarly, assuming $\delta > 1$ also leads to a contradiction. $\square$.

References

[1] Adams, W., 1987: "As New Hampshire Goes...", in: Orren, G. and N. Polsby (eds.), "Media and Momentum," Chatham House, Chatham, NJ.

[2] Bartels, L., 1988: "Presidential Primaries and the Dynamics of Public Choice", Princeton University Press, Princeton, NJ.

[3] Baye, M., D. Kovenock, and C. De Vries, 1990: "The All Pay Auction with Complete Information", CentER Discussion Paper 9051.

[4] Baye, M., D. Kovenock, and C. De Vries, 1994: "The Solution to the Tullock Rent-Seeking Game when $R > 2$: Mixed Strategy Equilibria and Mean Dissipation Rates", *Public Choice*, **81**, 363–380.

[5] Bikhchandani, S., D. Hirshleifer, D., and I. Welch, 1992: "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades", *Journal of Political Economy*, **100**, 992–1026.

[6] Cook, R., 2000: "United States Presidential Primary Elections, 1968-1996: A Handbook of Election Statistics", CQ Press, Washington, D.C.

[7] Cook, R. and A. McGillivray, 1997: "U.S. Primary Elections 1995-1996", CQ Press, Washington, D.C.

[8] Che, Y.-K. and I. Gale, 2000a: "Caps on Political Lobbying", *American Economic Review*, **88**, 643–651.

[9] Che, Y.-K. and I. Gale, 2000b: "Difference-Form Contests and the Robustness of All-Pay Auctions", *Games and Economic Behavior*, **30**, 22–43.

[10] Dasgupta, P. and E. Maskin, 1986: "The Existence of Equilibrium in Discontinuous Games, I: Theory," *Review of Economic Studies*, **53**, 1–26.

[11] Dekel, E. and M. Piccione, 2000: "Sequential Voting Procedures in Symmetric Binary Elections," *Journal of Political Economy*, **108**, 34–55.

[12] Erikson, R. and T. Palfrey, 2000: "Equilibria in Campaign Spending Games: Theory and Data," *American Political Science Review*, **94**, 595–609.

[13] Heckelman, J., 2000: "Sequential Elections and Overlapping Terms: Voting for US senate", *Journal of Economic Behavior and Organization*, **42**, 97–108.

[14] Malbin, M., 1985: "You Get What You Pay for, but Is That What You Want?," in: G. Grassmuck (ed.), "Before Nomination: Our Primary Problems," American Enterprise Institute for Public Policy Research, Washington, D.C.

[15] McGillivray, A. and R. Scammon, 1994: "America at the Polls: A Handbook of American Presidential Election Statistics," Vol. 2, Congressional Quarterly Inc., Washington, D.C.

[16] Oak, M., 2001: "On the Role of the Primary System in Candidate Selection", Cornell University, mimeo.

[17] Robinson, M. and M. Sheehan, 1983: "Over the Wire and on TV," Russell Sage Foundation, New York.

[18] Snyder, J., 1989: "Election Goals and the Allocation of Campaign Resources," *Econometrica*, **57**, 630–660.

[19] Snyder, J., 1990: "Campaign Contributions as Investments: The U.S. House of Representatives, 1980-1986," *Journal of Political Economy*, **98**, 1195–1227.

[20] Swank, O., 2001: "A comment on sequential elections and overlapping terms: voting for US Senate", *Journal of Economic Behavior and Organization*, **46**, 437-440.

[21] Tullock, G., 1980: "Efficient Rent-Seeking," in: J. Buchanan, R. Tollison, and G. Tullock (eds.), "Towards a Theory of the Rent Seeking Society," 97–112, Texas A&M University Press.

[22] Witcover, J., 1977: "Marathon: The Pursuit of the Presidency 1972-1976," New American Library, New York.