

Caucuses and Primaries under Proportional Representation

David Epstein Massimo Morelli Sharyn O'Halloran*

June 3, 2008

Abstract

We study a multi-district competition between two candidates when the winner is determined by the majority of total delegates obtained in the various States. The assignment of delegates in each State is proportional to the percentage of votes obtained by the two candidates, like in the Democratic party in the US, and each state can hold a high-cost election (a caucus) or a low-cost election (a primary). We examine the impact of higher variance of intensity of support on a candidate's spending patterns and electoral success in such a system and show that the theoretical predictions match patterns of votes received by Clinton and Obama in the 2008 Democratic primaries.

*Paper prepared for presentation at the Workshop on the Political Economy of Democracy, Barcelona, June 5-7, 2008. Preliminary draft; comments welcome.

1 Introduction

On March 4, 2008, the state of Texas held both a primary and a caucus to choose delegates for the Democratic National Convention, a unique arrangement known as the “Texas Two-Step.” Despite being held on the same day, in the same place, open to the same set of voters, and with the same two major candidates contesting the vote,¹ these two elections produced different results: Hillary Clinton won the primary with 50.9% of the vote to Barack Obama’s 47.4%, while Obama won the caucuses with 56.2% of the vote to Clinton’s 43.7%.² The turnout in these two events also differed widely: 2,868,454 voters cast ballots in the primary, over two and a half times as many as the estimated 1.1 million who participated in the caucus.³

Rarely do real world events present researchers with such a perfect natural experiment of competing electoral institutions, and rarer still do the contrasting institutions yield different results. Nor was this result exceptional; through May 20, Obama had won 14 out of the 16 caucus states, while Clinton had won 20 out of 34 primaries, with both candidates receiving almost identical numbers of total votes. Obama’s relative success in caucuses was explained by most commentators as due to his aggressive grass-roots organizing strategy⁴, but it clearly also has much to do with his ability to mobilize a number of supporters with very intense preferences for him relative to other candidates. This intensity was of significant importance in winning caucuses, where voters must go through a process taking two, three, four or more hours, as opposed to the (relatively) quick process of casting a vote in a

¹Edwards, Richardson, Biden and Dodd were still on the ballot as well, but they received a combined 1.72% of the vote in the primary and no delegates in the caucuses.

²All election results in the paper are taken from the New York Times website: <http://politics.nytimes.com/election-guide/2008/results/>.

³Exact turnout results for the caucuses are not available.

⁴See for instance http://www.boston.com/news/nation/articles/2008/05/04/small_state_plan_pays_dividends_for_obama/.

primary.

In this paper we explore in detail the implications of the observation that, even though each voter can only cast one vote, intensity of preferences matters when voting is costly. We use a simple voting model to derive a curvilinear relationship between the cost of voting and electoral outcomes between two candidates, each of whom has the same total support and the same intensity of support, but one candidate has a higher variance in intensity than does the other. In particular, the high-variance candidate will do best in contests with the highest and lowest costs of participation, while his opponent will do best in the intermediate-costs contests. We also show circumstances in which the candidate with the higher variance in support has an advantage over his opponent, being able to win the nomination while spending less money, despite having the same total support as his opponent. We then test the implications of the model on the results of the 2008 democratic presidential nominating race between Clinton and Obama, and find support for the curvilinear relationship derived from the theoretical model.

Our work fits into the broader discussion of the impact of electoral institutions on voting outcomes. This literature can be broadly divided into two categories: those studies that emphasize the impact of institutions on aggregating voters' preferences – presidential vs. parliamentary systems, proportional representation vs. plurality winners, candidate slates, and so on⁵ – and those studies that examine the impact of institutions on the composition of the electorate itself, including work on extending the franchise, on raising or lowering the cost of voting, and on raising the cost of non-voting, i.e., compulsory voting systems.⁶ Our work falls in the latter category, as the

⁵See for instance Shugart and Carey (1992) and Lijphart (1999)

⁶The literature on costly voting is voluminous; see Feddersen 2004 for a summary. The more specific question of the political consequences of expanding or contracting the electorate is also receiving renewed attention; see Lizzeri and Persico (2004). And a recent study of compulsory voting

costs of participation are significantly higher in caucuses as opposed to primaries, and this affects the size and composition of the electorate state by state. Rather than ask whether voting is “rational,” though, or how rational voters strategically cast their ballots in the face of costs to participation, we ask how differences in the intensity of preferences across voters and costs of voting across states or districts determines the winner of a multi-district simultaneous or sequential election, with the idea that these considerations may matter more than campaign strategies, fund raising, and/or election timing per se.

Our work also relates to studies of the primary system, most of which have been empirical, including classic works by Polsby (1983) and Bartels (1988). The former describes in detail the rise of the primary system as opposed to the former nominating system in which deals were worked out in “smoke-filled rooms,” arguing that the switch probably hurt the Democratic party, on balance. The latter describes the role of momentum in the primary system; those candidates who do well in early primaries are apt to do well later too, and small fluctuations in the early vote can have major consequences for the entire nominating process.

Issues of momentum and information aggregation in sequential elections are now becoming quite well studied in the theoretical literature. Starting with Feddersen and Pesendorfer (1996), game theorists have made significant contributions to our understanding of how informational cascades affect vote choice. Recent work in this tradition includes Fey (2000), Callander (2007), Iaryczower (2007), and Ali and Kartik (2007). Empirical studies of sequences of costly voting include early work by Lohmann (1994) and a more recent experimental study by Battaglini, et. al. (2007). These studies have done much to illuminate the role of strategic voting in

 (Helmke and Meguid, 2008) also shows consistent patterns of enfranchisement in democracies.

aggregating information from sequential contests. Unlike this literature, though, we focus on the impact of different voting institutions on primary outcomes, and on the ability of primaries to produce a general election winner. Sequential elections are important for us not because they make the $n + 1$ st voter more confident of which is the better candidate, but because they allow a candidate with heterogeneous supporters, competing in contests with differential costs of voting, to edge nearer to the nomination.

2 Model

Consider two exogenous candidates, i, j who run for election (or nomination). The winner is the one who obtains the majority of total delegates from n States, of equal size, and the assignment of delegates to each candidate is proportional to the percentage of votes received, State by State. In other words, the total delegates supporting candidate i are a fraction $\sum v_k^i/n$ of the total available delegates (ignoring the integer problem), where $v_k^i \in [0, 1]$ denotes the percentage of votes obtained by candidate i in State k , $k = 1, 2, \dots, n$. Normalize the size of the electorate in each of the States to 1, so that v_k^i is both a percentage and the number of votes, so we will use these terms interchangeably.

Before the campaign, each voter l is characterized by a parameter $d_i(l) \in \mathcal{R}$, which denotes the expected utility difference for l if i wins against j (so that if it is positive it means that l prefers i to win). Since we want to highlight the role of the distribution of preference intensities, we assume that the two candidates are “equal” on all other fronts, and also in terms of “total” intensity. In other words, assume

$$\int_0^1 d_i^k(l) dl = 0 \quad \forall k$$

and

$$\#\{l : d_i(l) > 0\} = \#\{l : d_i(l) < 0\}.$$

The positive d 's thus have the same mean as the absolute value of the mean of negative d 's, but the variance for positive numbers is higher. For simplicity, assume that the supporters of candidate i are divided in two groups, the high intensity and low intensity ones, whereas the supporters of candidate j have all the same intensity, such that the distribution of intensities for i is a mean preserving spread: $d_i^h > |d_j| > d_i^l$ where for some $\alpha \in [0, 1]$

$$\alpha d_i^h + (1 - \alpha) d_i^l = |d_j|. \text{ } ^7$$

Each candidate has by assumption unlimited funds for the campaign⁸ and here is the simplest possible way to describe what the campaign does in this benchmark model: each candidate decides what voters to “target,” under full information about their d type. Targeting one voter means spending one dollar talking to him or her, and what this targeted contact does is raise his or her absolute value of the initial d by a fixed amount g , so that if voter l with $d_i(l) > 0$ is targeted by i his posterior perception of utility differential is $d_i(l) + g$. No more than one dollar can be spent on each supporter. We will later extend the analysis to negative campaigns, but for now we assume that candidate i can only target effectively (that is, add the g to their excitement) for the voters who are initially inclined towards him or her.

The final two assumptions concern cost of voting and voting behavior: States

⁷For instance, in a simple spatial model with linear utilities and candidate platforms $x_i = 0$ and $x_j = 1$, the d_i^h voters could be a point mass of weight $1/3$ with ideal points $y_i^h = .1$, the d_i^l voters a point mass of weight $1/6$ with ideal points $y_i^l = .4$, and the d_j voters a point mass of weight $1/2$ with ideal points $y_j = .8$

⁸As we will see, the result does not depend on this simplifying assumption.

have voting costs C_k , where to assure the possibility of positive turnout we assume that $C_k < d_i^h + g$ for all k . Voters vote sincerely and turn out *if and only if* their after-campaign utility differential is greater than the cost of voting in their State.⁹

Proposition 1. *There exist voting costs \bar{C} and $\underline{C} < \bar{C}$ such that candidate i wins in all states for which $C_k > \bar{C}$, candidate j wins when $\bar{C} > C_k > \underline{C}$, and the election is a tie when $C_k < \underline{C}$.*

Proof. Take \bar{C} to be $d_j + g$. Then for all $C_k > \bar{C}$, candidate j cannot convince her voters to vote, so turnout in her favor is zero. Candidate i , on the other hand, can convince his high-intensity supporters to turn out, possibly by campaigning, since by assumption $C_k < d_i^h + g$.

Now take $\underline{C} = d_i^l + g < \bar{C}$. For all $C_k \in (\underline{C}, \bar{C})$, Candidate j can now entice her supporters to vote, possibly by campaigning, so she gets full turnout, while Candidate i only has his high-intensity supporters turn out. Thus j wins all elections in this range.

Finally, for $C_k < \underline{C}$, both candidates can generate full turnout, so all voters cast ballots and the election is a tie. □

The proposition establishes a basic curvilinear relation between the cost of voting and the electoral outcome: for high-cost states, the candidate with the greater number of intense supporters will do well, as more voters will, for instance, spend three, four, or more hours at a caucus to vote. In the middle range, Candidate j does better, as her voters will all pay moderate costs to cast their ballots. And at the low end of cost all voters go to the polls; as we assume that the overall support for the candidates is equal, these elections end in ties.

⁹Note that here the cost of voting is a common cardinal measure of for example time needed to go to vote and participate in the election, since the subjective and personal or emotional things can all be lumped in the d parameters.

Table 1 illustrates the various possible combinations of voting, expenditures, and outcomes. It is constructed by noting that if a voter with intensity d_1 will vote for her candidate with no campaign expenditures, then another voter with intensity $d_2 > d_1$ will vote as well.

Case	Who Votes	i Pays	j Pays	Winner
1	d_i^h	0	0	i
2	d_i^h	d_i^h	0	i
3	d_i^h, d_j	0	0	j
4	d_i^h, d_j	0	d_j	j
5	d_i^h, d_j	d_i^h	d_j	j
6	d_i^h, d_j, d_i^l	0	0	tie
7	d_i^h, d_j, d_i^l	d_i^l	0	tie
8	d_i^h, d_j, d_i^l	d_i^l	d_j	tie
9	d_i^h, d_j, d_i^l	d_i^h, d_i^l	d_j	tie

Table 1: Possible Combinations of Voters and Campaign Expenditures

The table reveals some interesting regularities regarding spending and voting patterns. First, in equilibrium, notice that the winning candidate spends weakly more funds getting their supporters to turn out than does the losing candidate. This is true even though the cost of advertising is zero, so the losing candidate is refraining from spending only because he knows that it will not change the outcome.

Also, whenever the high-intensity candidate wins an election and spends money, he spends on his highest intensity supporters. This echoes the conventional wisdom in the primaries literature (e.g., Polsby 1983) that successful candidates “mobilize a faction,” rather than appeal to a broad segment of the party faithful.

3 Two-State Example

To illustrate the advantage that the high-variance candidate may have in a system of primaries and caucuses, assume for the moment that the election is contested in two States, which have costs of voting $C_1 > C_2$ (representing the fact that State 1 has Caucus elections and State 2 has a primary).

Proposition 2. *There exist generic parameter values under which candidate i wins the election (1) spending less money in the campaign; and (2) focusing his campaign in the State with higher cost of voting.*

Proof. The proof is by example. Suppose

$$d_i^h + g > C_1 > d_j + g > C_2 > d_i^l + g.$$

In this example candidate j can mobilize all her supporters in one State with the lower cost, but not in the other, while candidate i can effectively mobilize only the high intensity supporters but in both States. This contrast can be sufficient to generate the result: clearly i wins the State with high cost and candidate j wins the other State, but if $d_i^h > C_2$, candidate i loses in State 2 by “less” than he wins by in State 1, hence he wins the elections even though he doesn’t really have to spend money in State 2, and when $2d_i^h < d_j$, he will win the primaries while losing the popular vote. □

Remark 1. *In the example above the asymmetric equilibrium in the campaign game is clearly unique and does not depend on the order in which the States are addressed nor on any sophisticated strategy.*

Remark 2. *If the two States had equal cost of voting, namely both primaries or both caucuses, then the equilibrium would have to be symmetric, and the candidate with lower variance of intensities in her support could have an advantage if the cost of voting remains too high to mobilize the least intense supporters of the high variance candidate.*

Thus the high-variance candidate relies on the possibility of out-mobilizing his opponent in caucus states; otherwise he would have no chance to win the election. This insight can be extended to a simple model of strategic candidate locations, where one candidate strategically separates himself from his opponent in order to create a cadre of high-intensity supporters.

Take a simple spatial model where voter ideal points are uniformly distributed in the $[0, 1]$ interval, voters have linear utilities, and the policy platform x_j of candidate j is fixed at $x_j > 1/2$.¹⁰ Again assume a two-state competition with voting costs $C_1 > C_2$, ignore for the moment the possibility of advertising, and allow candidate i to pick any policy platform within the space.

Then traditional (costless) voting theory would predict candidate i would locate at x_j , or $x_j - \epsilon$ for some arbitrarily small ϵ , winning the election. In the presence of costly voting, though, such a strategy would leave the utility difference between the candidates at zero (or ϵ), and hence turnout would be zero as well.

As long as $C_1 < 2x_j - 1$, candidate i can win the election by locating at $x_i = x_j - C_1$. Then voters with ideal points between 0 and x_i have utility difference $d_i = C_1$, as do voters between x_j and 1, so they all vote in State 1's caucus. But by construction $x_1 > 1 - x_j$, so candidate i wins the caucus state, and he will win the primary state by the same margin as well. So candidates will strategically position

¹⁰Candidate j may, for instance, have a longer track record of votes which defines her policy positions to a greater extent than her opponents'.

themselves, separating from their opponents just to the point that their most fervent supporters will turn out in the states with the highest cost of voting.

4 Evidence From the 2008 Democratic Primaries

The theoretical model above predicts that, under certain circumstances, primary returns for the candidate with a higher variance in the intensity of his support will show a curvilinear relation relative to the costs of voting. In particular, the high-variance candidate should do best when the cost of voting is highest or lowest, and his opponent should do best in states with intermediate costs of voting.

We now examine whether these patterns are apparent in the 2008 Democratic primary election contest between Barack Obama and Hillary Clinton. To a first approximation, the requirements of the model are met: both have nearly equal levels of support in the public at large; as of this draft, Clinton led 50.1% to 49.9% in the popular vote. And Obama's followers for much of the primary season included both a set of ardent supporters, giving rise to the "Obama-mania" phenomenon, and independents, with perhaps a lower level of attachment to any Democratic candidate.

Some limitations are also apparent. First, we have no way to directly measure the intensity of support of Obama followers as opposed to Clinton's, and the latter certainly has her share of voters highly dedicated to her cause. Second, many primaries were contested with other candidates in the field, particularly John Edwards, who gathered a significant number of votes. How the Edwards voters would have cast their ballots were he not in the race is difficult to determine. And third, data availability limits us to testing only the predictions of relative voter support across states, rather than the predictions made on spending patterns.

The data used for this section comprise, first, the rules by which each state held

its primary and/or caucus for the Democratic presidential nomination in 2008.¹¹ The contests differed by their method (primary or caucus), their eligibility requirements, and the minimum number of days prior to the contest that voters could register to participate. Both the method and registration rules are straightforward; the definitions of the eligibility categories are as follows:

- Open: An open contest is one in which any registered voter may participate.
- Closed: A closed contest is one in which a voter may only vote if he is registered with that political party (you must be a registered Democrat to vote in the Democratic contest and a registered Republican to vote in the Republican contest.)
- Modified: A modified contest is one in which voters may participate in if they are either registered with that party, or are a registered voter with no party affiliation (e.g. you may vote in your party's contest, or either party's contest if you are independent.)
- Affiliation Change: In a modified election, if you are a registered voter but not registered with a party, your party affiliation must be changed to the party of the primary you vote in.

From these categories, we create a variable "Costrank" indicating in an ordinal manner how costly each contest is to participate in, from voters' point of view. We used the following rules:

1. All caucuses were deemed to be more costly than all primaries.

¹¹These rules can be found on the web at <http://www.thegreenpapers.com> and are summarized in the appendix.

2. Within each category, closed contests were deemed to be the most costly, as they allow only party members to vote.
3. Within each category, modified contests with affiliation change came next, followed by modified contests without affiliation change, under the view that some psychic costs are involved for independents to declare themselves members of one party or the other.
4. Finally, within each category, came open contests.

The resulting set of Costrank levels is given in Table 2. As defined, of course, the distance between adjacent categories is equal, which is certainly a simplification of reality. In particular, the potential gulf between caucus and primary states is not captured directly in our Costrank variable.

Costrank	Method	Eligibility	Δ Aff?
1	Primary	Open	
2	Primary	Modified	N
3	Primary	Modified	Y
4	Primary	Closed	
5	Caucus	Open	
6	Caucus	Modified	N
7	Caucus	Modified	Y
8	Caucus	Closed	

Table 2: Definition of “Cost Rank” categories.

For each contest we also calculated the percent of the two-candidate Clinton vote as the number of votes for Clinton divided by the sum of the Clinton and Obama votes. Summary statistics for all variables are given in Table 3, where the eligibility variable is coded as 1 for open, 2 for modified and 3 for closed.

Our model predicts that the Clinton vote will be curvilinear in the cost of voting

Table 3: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Clinton Vote	0.45	0.14	0.08	0.73	48
Caucus State	0.27	0.45	0	1	48
Eligibility	2.1	0.88	1	3	48
Days Prior	22.33	10.3	0	38	46
Costrank	4.04	2.57	1	8	48

in various primary and caucus contests. As a first look at the data, Figure 1 shows a scatterplot and quadratic fit of Clinton’s results. As predicted, Clinton does better in states with intermediate costs and worse at either end, although especially so in closed caucuses.

We now subject the data to a simple regression analysis. Were one to naively approach the data as a set of independent variables, one might run a regression similar to that in Model 1 of Table 4, treating each aspect of the caucus or primary contest separately. As shown, only the caucus vs. primary variable is significant here; the more specific type of contest (open, modified or closed) has no extra impact.

But our model cautions us to look for nonlinear impacts of these variables. For instance, for a caucus state, making the primary open rather than closed lowers the cost of participation and should thus favor Clinton. But in a primary state, lowering the cost of participation should help Obama. These considerations are wrapped up in our Costrank variable, and Model 2 in the table looks for direct and quadratic effects of Costrank alone.

As predicted, the quadratic term is negative and significant, while the linear term is insignificant. When one adds back in the caucus variable in Model 3, it is still significant and negative, indicating that, as hypothesized above, the linear formulation of Costrank downplays the significant difference between primary and caucus states. And Model 4 shows that adding the eligibility requirements back in

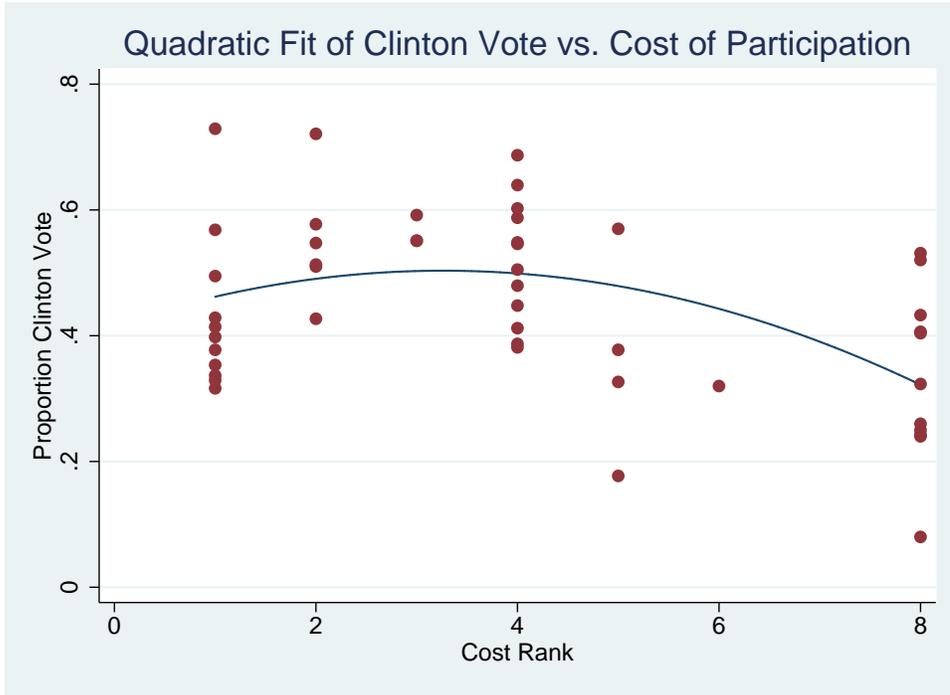
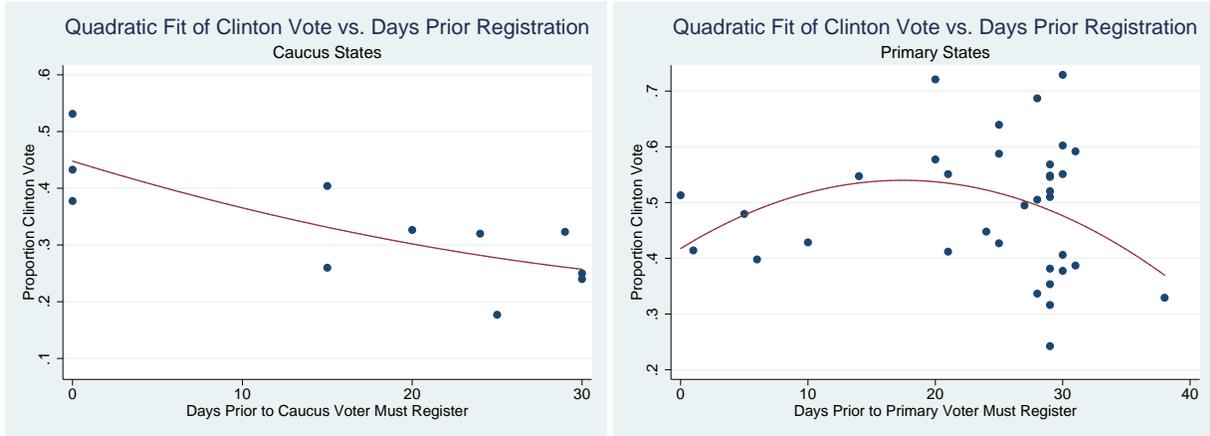


Figure 1: Relation between cost of participation and Clinton vote.

	PClinton (1)	reg2 (2)	reg3 (3)	reg4 (4)	reg5 (5)
Caucus State	-.186 (.041)***		-.131 (.053)**	-.123 (.060)**	
Eligibility	.030 (.019)			.009 (.029)	
Days Prior	-.003 (.002)				.014 (.008)*
Days Prior ²					-.0004 (.0002)*
Costrank		.053 (.031)*	.061 (.030)**	.054 (.037)	
Costrank ²		-.008 (.003)**	-.007 (.003)**	-.007 (.003)*	
Obs.	46	48	48	48	35
R ²	.338	.249	.339	.341	.086

Table 4: Regression Results.



(a) Caucus States

(b) Primary States

Figure 2: Clinton Vote vs. Registration Requirements

yields no additional explanatory power.

Various specification tests were performed to examine these results. No one case was classified as an influential outlier, as would be indicated by a Cook's distance greater than 1. Neither did any state have particularly high leverage, and collinearity concerns entered only in the expected high correlation between the Costrank variable and its square.

We have not yet incorporated the days prior registration information to our analysis, as it is not clear how it relates to the elements of Costrank. However, Figure 2 shows a scatterplot of days prior and the Clinton vote for caucus and non-caucus states separately, with Clinton doing better in the low-cost caucus states and a curvilinear relation between cost and results in the primary states. Model 5 in the regression table, run for primary states only, confirms that this latter curvilinear relation is statistically significant, though not highly so.

5 Conclusion

This paper presented a model of primaries and caucuses with heterogeneous voters and different costs of voting across states. We derived a nonlinear relationship between costs of voting and election outcomes in an environment where two candidates have equal overall support, but one candidate has a higher variance of support than the other; more zealots, but more apathetic supporters as well. We showed that in a certain set of parameter values the high variance candidate will have an advantage over his rival, winning high-cost caucuses by less than he loses low-cost primaries. We analyzed patterns of electoral competition from the Clinton-Obama primary race in 2008 and found them to be consistent with the predictions of our model.

We found as well that candidates may strategically locate themselves in a policy space, differentiating themselves from their opponent, in order to galvanize their most intense supporters. The next question to ask is whether these candidates may have more difficulty appealing to the median national voter in what is, after all, a low-cost general election. This raises the questions regarding optimal electoral institutional design, and if differing the costs of participation at the electoral and primary stage makes sense. These issues are becoming increasingly relevant to real-world events, as the primary system is in a state of flux and may undergo significant revisions before the next presidential election cycle, and we leave these important questions to future work.

References

- [1] Ali S. Nageeb and Navin Kartik. 2007. "Social Learning in Elections." Manuscript, University of California, San Diego.
- [2] Bartels, L. M. (1988): *Presidential Primaries and the Dynamics of Public Choice*, Princeton, NJ: Princeton University Press.
- [3] Battaglini, M., R. Morton, and T. Palfrey (2007): "Efficiency, Equity, and Timing in Voting Mechanisms," *American Political Science Review*, 101, 409-424.
- [4] Borgers, T. [2004], "Costly Voting," *American Economic Review*, Vol. 94, No. 1, 57-66.
- [5] Callander, S. (2007): "Bandwagons and Momentum in Sequential Voting," *Review of Economic Studies*, 74, 653-684.
- [6] Feddersen, T. and W. Pesendorfer (1996): "The Swing Voter's Curse," *The American Economic Review*, 86, 408-424.
- [7] Feddersen. 2004. Rational choice theory and the paradox of not voting. *Journal of Economic Perspectives*. Volume 18, Number 1, Winter 2004 , pp. 99-112(14)
- [8] Fey, M. (2000): "Informational Cascades and Sequential Voting," University of Rochester Working Paper.
- [9] Helmke, Gretchen, and Bonnie M. Meguid. 2008. "Endogenous Institutions: The Origins of Compulsory Voting Laws." Manuscript, University of Rochester.
- [10] Iaryczower, M. (2007): "Strategic Voting in Sequential Committees," Caltech, working paper.

- [11] Lijphart, Arend. 1999. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. Yale University Press.
- [12] Lizzeri, Alessandro and Nicola Persico. 2004. "Why Did the Elites Extend the Suffrage? Democracy and the Scope of Government, With an Application to Britain's 'Age of Reform.'" *The Quarterly Journal of Economics* 119(2), May 2004, pp. 707 - 765.
- [13] Lohmann Susanne *The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989-91* *World Politics*, Vol. 47, No. 1 (Oct., 1994), pp. 42-101
- [14] Palfrey, T. and Levine, D. [2006], "The Paradox of Voter Participation? A Laboratory Study," *American Political Science Review* , forthcoming.
- [15] Palfrey, T. and Rosenthal, H. [1985], "Voter Participation and Strategic Uncertainty," *American Political Science Review*, Vol. 79, No. 1, 62-78.
- [16] Polsby Nelson W *Consequences of Party Reform*. By.. (New York: Oxford University Press, 1983.
- [17] Shugart, Matt and John Carey. 1992. *Presidents and Assemblies: Constitutional Design and Electoral Dynamics*. Cambridge University Press.

Appendix: Primary Contest Rules

State	Date	Eligibility	Δ Aff?	Method	Registration	Days Prior
AL	5-Feb	Open		Primary	26-Jan	10
AK	5-Feb	Closed		Caucus	6-Jan	30
AZ	5-Feb	Closed		Primary	7-Jan	29
AR	5-Feb	Open		Primary	6-Jan	30
CA	5-Feb	Modified	N	Primary	22-Jan	14
CO	5-Feb	Closed		Caucus	7-Jan	29
CT	5-Feb	Closed		Primary	31-Jan	5
DE	5-Feb	Closed		Primary	12-Jan	24
DC	12-Feb	Closed		Primary	14-Jan	29
FL	29-Jan	Closed		Primary	31-Dec	30
GA	5-Feb	Open		Primary	7-Jan	29
HI	19-Feb	Closed		Caucus	5-Jan	30
ID	5-Feb	Open		Caucus	2-May	25
IL	5-Feb	Open		Primary	8-Jan	28
IN	6-May	Modified	N	Primary	7-Apr	29
IA	3-Jan	Closed		Caucus	3-Jan	0
KS	5-Feb	Closed		Caucus	21-Jan	15
KY	20-May	Closed		Primary	22-Apr	28
LA	9-Feb	Closed		Primary	9-Jan	31
ME	10-Feb	Closed		Caucus	26-Jan	15
MD	12-Feb	Closed		Primary	14-Jan	29
MA	5-Feb	Modified	N	Primary	16-Jan	20
MI	15-Jan	Open		Primary	17-Dec	30
MN	5-Feb	Open		Caucus	16-Jan	20
MS	11-Mar	Open		Primary	10-Feb	30
MO	5-Feb	Open		Primary	9-Jan	27
MT	3-Jun	Open		Primary	5-May	29
NE	19-May	Modified	N	Caucus	25-Apr	24
NV	19-Jan	Modified	Y	Caucus	19-Jan	0

State	Date	Eligibility	Δ Aff?	Method	Registration	Days Prior
NH	8-Jan	Modified	N	Primary	8-Jan	0
NJ	5-Feb	Modified	Y	Primary	15-Jan	21
NM	5-Feb	Closed		Primary	8-Jan	28
NY	5-Feb	Closed		Primary	11-Jan	25
NC	6-May	Modified	N	Primary	11-Apr	25
ND	5-Feb	Open		Caucus	5-Feb	0
OH	4-Mar	Modified	Y	Primary	3-Feb	30
OK	5-Feb	Closed		Primary	11-Jan	25
OR	20-May	Closed		Primary	29-Apr	21
PA	22-Apr	Closed		Primary	24-Mar	29
RI	4-Mar	Modified	Y	Primary	2-Feb	31
SC	26-Jan	Open		Primary	20-Dec	38
SD	3-Jun	Closed		Primary	19-May	15
TN	5-Feb	Open		Primary	7-Jan	29
TX	4-Mar	Open		Mixed	4-Feb	29
UT	5-Feb	Modified		Primary	6-Jan	30
VT	4-Mar	Open		Primary	27-Feb	6
VA	12-Feb	Open		Mixed	12-Jan	29
WA	9-Feb	Modified	N	Caucus	4-Feb	5
WV	13-May	Modified	N	Primary	22-Apr	20
WI	19-Feb	Open		Primary	18-Feb	1
WY	8-Mar	Closed		Caucus	22-Feb	15