Journal of Theoretical Politics 17(1): 107–136
DOI: 10.1177/0951629805047800

INFORMATIONAL PARTY PRIMARIES AND STRATEGIC AMBIGUITY

Adam Meirowitz

ABSTRACT

While scholars have thoroughly explored the logic of two candidate electoral competition, much less has been accomplished in gaining an understanding of the role of party primaries. This paper presents an incomplete information model of primary and general elections and argues that party primaries do more than select party candidates. Party primaries serve an informational function. In an environment where candidates are uncertain about the preferences of voters, selection of desirable policy platforms is a risky, if not difficult, undertaking. Primary elections offer voters an early opportunity to signal their preferences to candidates. Before primary elections, the candidates, aware that information about voter preferences is forthcoming, have an incentive to remain ambiguous about their policy platforms. Early commitment makes them vulnerable to better informed candidates that they might face in the general election. The fully revealing equilibrium of the game yields a joint explanation of the role of party primaries and candidate ambiguity. Primaries aggregate information about voter preferences and candidate ambiguity has an option value.

KEY WORDS • elections • information transmission • political parties • primaries

1. Introduction

All elected US presidents with the exception of George Washington have entered office through a two-step electoral process. The first step is competition for a major party nomination and the second is a general election. The first step has been the subject of periodic reform and has exhibited radical variation. The general trend of this reform has been to move nomination power out of the hands of party elites and into the hands of the party membership or electorate (Davis, 1980; Lengle, 1981; Bartels, 1988; Polsby

The author is grateful for conversations with David Baron, Jonathan Bendor, Daniel Diermier, John Duggan, Yossie Feinberg, John Ferejohn, Mark Fey, Tim Groseclose, Keith Krehbiel and Alan Wiseman. I appreciate assistance from Natasha Zharinova. In addition, the paper has benefited from the comments of two reviewers and the editor of the *Journal*.

and Wildavsky, 1988; Schantz, 1996).[1] The distributive or policy consequences of this reform are not well understood. If primaries serve to polarize candidates by making them more responsive to the party-faithful voters, then the effects of this reform could be minor. A simple theory might predict that the party leadership has preferences similar to the median party voter. In this case, it might not matter whether party nominees are chosen by the party leadership or the party voters. An alternative story, suggested by the model we analyze, is that the reform enhanced informational efficiency in the face of uncertainty about preferences as it allows candidates to learn about the preferences of many primary voters instead of just the party leadership.

Two stylized facts about American elections are that (1) they exhibit a trend of uncertainty reduction and (2), as Downs put it, 'candidates becloud their policies in a fog of ambiguity' (Downs, 1957; Shepsle, 1972; Page, 1976). One manifestation of the first point is that public opinion polls become less volatile and more accurate as election day closes and there seems to be a sharp jump around the party primaries (Bartels, 1988; Gelman and King, 1993; Campbell, 2000). Bartels (1988) for example observes that Jimmy Carter in 1976 and Gary Hart in 1984 'succeeded in breaking out of the pack' while being known for vagueness and lack of substance. Moreover, Bartels notes that Hart's campaign exhibited a dramatic shift becoming more focused on substance around the end of the primary season. Page (1978) cites Barry Goldwater as an example of a noteworthy *insurgent exception* to the rule of ambiguous candidates. He credits Goldwater's tendency to make vocal uncompromising stands as his eventual undoing.

> The very issue stands which brought forth his core of activists are likely to alienate much of the general public. The candidate comes under intense pressure from opponents, party regulars, and money givers to 'move toward the center' and adopt more orthodox stances. . . . In trying to move toward the center, however insurgents risk alienating their original supporters; they have difficulty persuading the public that they have 'really' changed; and they are bombarded with charges of instability, weakness, and dishonesty, which damage their personal images. Recent insurgent candidates have not been notably successful in presidential elections. (Page, 1978: 119)

This paper develops an informational model of primary and general elections consistent with these features. The explanation involves candidates learning about voter preferences from the party affiliation of voters in the primaries. In equilibrium, primary candidates avoid committing to policy platforms until after the primaries. Before the primary, voters know little

1. For example, in the first half of the 19th century, presidential nominations were made by state legislatures and caucuses. It was not until the early part of the 20th century that the nomination process moved from state conventions or caucuses to primaries.

about the candidate platforms and after the primary they know much more. Additionally, after the primary candidates are better informed about voter preferences.

While party primaries are conjectured to be an important feature of the American electoral landscape, a theoretical understanding of primaries has been sought by comparatively few scholars. In contrast, two-candidate electoral competition (without primaries) has been extensively studied by theorists (e.g. Downs, 1957; Ledyard, 1984; Calvert, 1985; McKelvey and Ordeshook, 1985, 1986a, b; Banks, 1990; Chappell, 1994; Martinelli, 2002).[2] Some scholars have argued that primaries increase interparty or intraparty competition.[3] Arguments of this type suggest that primaries act as a mechanism to select party representatives in the general election and may also have some polarizing influence on the policy platforms of party candidates. Some empirical scholars have argued that primary competition is not, however, over policy (e.g. Gopoian, 1982; Norrander, 1986). This brings into question the validity of an explanation of primaries as a mechanism for competition over policy.[4] This paper posits a substantially different role for primaries. Primaries offer voters an opportunity to signal information about their preferences to candidates before candidates commit to policy platforms in the general election. In equilibrium, however, no competition over policy occurs in the primary election as primary candidates choose not to distinguish themselves. That is, the primaries are a forum for aggregating information about voter preferences rather than an arena for party competition over policy. In this setting, we find strong incentives for candidates to remain ambiguous throughout the primary, committing to policy only late in the election cycle.

Loosely speaking, the game considered has the following structure: There is a large population of voters with policy preferences and four candidates (two running for each of two party nominations). The policy space is unidimensional and voters have symmetric single-peaked preferences. Each voter knows only her own ideal point and all agents (voters and candidates) have prior beliefs over a set of possible distributions of voter ideal points. Candidates have preferences over policy and holding office. At the beginning

2. The comparison between theories of electoral competition and primaries may seem slightly overstated here, if one considers scholarly work on primaries such as Aldrich (1980), Bartels (1988), and Geer and Shere (1992). But these works do not develop a clear policy prediction based on a model of the process. This is in contrast to the cited works on elections.

3. Campbell (2000) reviews the work on campaign competition that makes this point. Somewhat more formal formulations of competition-based theories appear in Geer and Shere (1992) and Aldrich (1980).

4. The possibility that primaries are a forum for competition over non-policy dimensions, such as personality or competence, remains.

of the game, primary candidates simultaneously either commit to a policy point for the remainder of the game or remain completely ambiguous. Following the announcement of these primary stances, voters first affiliate with a party and then after all of the affiliation decisions are made, they vote in their party's primary.[5] The two primary winners (determined by plurality rule in each primary) then face each other in a general election. If a primary winner was ambiguous in the primary, she selects a definite policy position for the general election. If a primary winner selected a policy point in the primary, she is committed and cannot change her stance. The voters then cast their ballots and the winning candidate (determined by plurality rule) enacts her announced policy position.

The equilibrium has the following features:

(i) Endogenous party affiliation by voters conveys information to candidates about the preferences of voters.
(ii) Primary candidates choose to be ambiguous because (a) they know they will be better informed about voter preferences when selecting a policy platform in the general election and (b) commitment to a policy in the primary makes them vulnerable to a challenger in the general election who can choose a policy position based on the information revealed in the primary.
(iii) Primary voters reward candidates who are ambiguous because these candidates will pull the final policy as close to the party's median voter as possible in a general election.
(iv) Candidates learn from party affiliation, and primary winners compete in a general election characterized by 'little' uncertainty regarding voter preferences and hence announce the same centrally located platform.

The study of ambiguity by candidates is quite extensive (Downs, 1957; Zeckhauser, 1969; Fishburn, 1972; Shepsle, 1972; Page, 1976; McKelvey, 1980; Bernhardt and Ingberman, 1985; Snyder, 1990, 1992; McCarty and Rothenburg, 1996; Aragonès and Postlewaite, 2002; Kroszner and Stratmann, 1999).[6] Glazer (1990) was one of the first to argue that ambiguity could be a strategic response to uncertainty about the voter preferences. Alternatively, Alesina and Cukierman (1990) and Aragonès and Neeman

5. With this assumption the model can be interpreted as an election with closed primaries in which voters determine which party to register with while candidates are campaigning or open primaries in which voters receive good forecasts about the affiliation decisions.

6. Kroszner and Stratmann present an excellent review of the theoretical and empirical literature on this subject. Page (1976) is very precise in discussing the task of modeling this phenomenon and such examples as Austen-Smith (1983), Alesina and Cukierman (1990), and Aragonès and Postlewaite (2002) make it clear that ambiguity is a puzzle not yet explained by spatial models.

(2000) analyze models of elections in which ambiguity is motivated by the desire to remain unconstrained in the future. In these models, candidates exogenously prefer to be less constrained in office, whereas voters prefer precise stances. The equilibrium levels of ambiguity involve a trade-off between these opposing forces. In contrast, in the current paper, there is no exogenous value to ambiguity for the candidates and, in equilibrium, even risk averse voters prefer candidates who are ambiguous early in the election cycle.

In the model presented here, candidates choose to be ambiguous for reasons that are different from, albeit related to, those of Glazer, Alesina and Cukierman, and Aragonès and Neeman.[7] Here, primary candidates are uncertain about voter preferences and anticipate that, in the future, they will become better informed, and this induces strategic ambiguity early in the election cycle. We call this the *option value* explanation; i.e. since candidates may obtain future information about voter preferences, the ability to remain ambiguous or avoid taking a stance early has a positive option value in equilibrium. Interestingly, in the equilibrium, risk-averse voters also value the ambiguity of primary challengers from their party. With a minimal level of policy motivation and polarized candidate preferences, unconstrained candidates are reasonably good agents for the party consisting of those voters who support them. In other words, in equilibrium, part of the reason primary challengers remain ambiguous is because taking a stance will result in punishment by the primary voters and the loss of the primary. The equilibrium also exhibits the intuitive property that commitment by one candidate in the primary makes her vulnerable to an unconstrained challenger she might face in the general election.

In Section 2, we present the model, first defining the game form and then the relevant equilibrium concepts. In Section 3, we present and justify the results. We begin with several intermediate results characterizing equilibrium play at various histories in the game and then present the main proposition characterizing an equilibrium in which candidates are ambiguous in the primary and party affiliation allows the general election candidates to infer the distribution of voter ideal points. Proofs appear in the Appendix. In Section 4, we conclude with a narrow discussion of various extensions to the model and a broad discussion of possible interpretations of the characterized equilibrium.

7. Following the arguments in Page, we depart from modeling ambiguous stances as probability distributions. Instead we consider ambiguity as a condition of being less constrained. In the model, we consider a simple strategy space where candidates can either commit to a policy (a point in the unidimensional policy space) or remain ambiguous allowing themselves to commit to any policy in the future. The results can be generalized to a model where primary candidates can commit to any subset of the policy space.

2. The Model

A few words about the modeling choices are in order. We formulate a model involving four candidates (two from each of two parties) and a continuum of voters. Candidate policy preferences are common knowledge and voter preferences are private information, with candidates (and voters) facing uncertainty about the distribution of voter ideal points. The assumption of four candidates represents the simplest specification that allows for non-trivial primaries in each party. The assumption that there is a continuum of voters (as opposed to a large but finite population) dramatically simplifies the analysis when candidates update beliefs about population preferences and adequately represents voters as very ‘small’. The model considers only a single primary for each party, thereby precluding any results about momentum. This simplification and our focus on common knowledge of candidate preferences and imperfect information about voter preferences allows us to focus on a feature of primaries that tends to be ignored – namely the potential for institutions to provide opportunities for elites to learn about voter preferences.

Specific functional forms for utility functions and distributions are not specified, because the analysis is not complicated by this form of generality. While the informational structure of the model is reasonably parsimonious, we introduce a number of measure-theoretic concepts. They are used in defining the game form and strategies when the measure of certain sets may not be conveniently expressible in terms of density or distribution functions. Finally, to focus the analysis on questions of information aggregation, we assume that primary candidates of the same party are identical. The most pronounced effect of this assumption is that primary voting turns out to be uninteresting as, in equilibrium, primary candidates of the same party are indistinguishable. Analysis of primary voting behavior off the equilibrium path does focus on primaries in which the candidates are distinguished, because of actions taken in the primary. While this assumption is clearly not descriptive of actual elections, it allows us to focus on the trade-offs between clarity and ambiguity for candidates during the primary and for voters when comparing candidates that have taken different primary actions.

2.1. The Game

A continuum of voters $V := [0, 1]$ have policy preferences over the policy space $X := [0, 1]$. The utility function of voter $v \in V$ is

$$u_v(x) := \psi((x - y_v)^2) \tag{1}$$

where $x, y_v \in X$ and $\psi : [0, 1] \to \mathbb{R}$ is a strictly decreasing continuous function. The ideal points y_v of voters are distributed according to a continuous distribution function $F(x, \theta)$. When it simplifies the exposition, we represent this distribution with the corresponding atomless probability measure $\mu(\cdot, \theta)$ on $(X, \Im)$, where $\Im$ denotes the Borel sigma algebra of the support X. The Borel sigma algebra is a large collection of subsets of the unit interval. The parameter $\theta \in \Theta := [0, 1]$ is the unknown median of the distribution function, and the distribution and measure are related by the following identity, $F(x, \theta) := \mu([0, x), \theta)$. We assume that for any set $A \in \Im$, the function $\mu(A, \cdot)$ is continuous in θ on (0, 1), and that for every θ, $F(\cdot, \theta)$ has full support on X. We require that if $\theta < \theta'$ then for all $z \in (0, 1)$, $F(z, \theta) > F(z, \theta')$. That is, if $\theta < \theta'$ then $F(\cdot, \theta)$ strictly first-order stochastically dominates $F(\cdot, \theta')$.[8] We assume that the common prior on θ is $G(\cdot)$, a continuous distribution function on support [0, 1]. So $\mu(\cdot, \theta)$ is a device to count or, more precisely, measure voters given the parameter θ, whereas $G(\cdot)$ captures the underlying uncertainty about θ. In other words, for a fixed parameter θ the exact number of voters with ideal points in set $A \subset X$ (more precisely $A \in \Im$) is just $\mu(A, \theta)$. The uncertainty in the model is captured by the fact that agents do not know θ at the beginning of the game. Each voter knows her own ideal point and all players have beliefs about the location of θ characterized by the prior $G(\theta)$. The beliefs on θ induce beliefs about the probability measure $\mu(A, \theta)$ for any $A \in \Im$.

This prior, $G(\cdot)$, is assumed to capture the common beliefs of each voter v conditional on the voter's ideal point y_v. In other words, all voters know their own ideal point and have common beliefs about θ. This assumption is different than assuming that agents each have a common prior and then learn their ideal point. The latter framework requires that each voter then calculates a Bayesian posterior on θ given the ideal point she observes. Under such an assumption, the proof of Lemma 4 does not go through. The substantive bite of this formalization is that it requires two agents with different preferences (say an extreme liberal and an extreme conservative) to have the same beliefs about the location of the median voter.[9]

There are four candidates $C := \{l_1, l_2, r_1, r_2\}$ with the interpretation that l_1 and l_2 (r_1 and r_2) are in the L (R) party primary. The final outcome of the general election is a policy x and an office-holder $w \in C$. Specifically, the

8. This ordering is satisfied by many commonly used parametric families, like the family of normal distributions with unit variance.

9. While not innocuous, this assumption does seem reasonable. Voters bring preferences and beliefs to the election. The most reasonable substantive argument for common beliefs would be that voters enter the election with a long history of common exposure to the media and, therefore, have similar information. Voters have heterogenous preferences and common beliefs. This is the assumption that we make.

enacted policy is the stance taken by the candidate that wins the final election; i.e. candidates can commit. Candidates have utility functions over the final policy and winner of the form:

$$u_c(x, w) := (1 - \kappa)1_{\{w=c\}} + \kappa\gamma((x - y_c)^2). \tag{2}$$

where

$$1_{\{w=c\}} = \begin{cases} 1 & \text{if } c \text{ wins} \\ 0 & \text{otherwise} \end{cases}$$

the candidates' ideal points are $y_c = 0$ for $c \in \{l_1, l_2\}$, $y_c = 1$ for $c \in \{r_1, r_2\}$, the policy importance coefficient, κ, satisfies $0 < \kappa < 1$, and the function $\gamma : [0, 1] \to \mathbb{R}$ is strictly decreasing and continuous. Substantively, candidates care about policy and holding office. They prefer enacting policies closer to 0 (1) if they are either l_1 or l_2 (r_1 or r_2).[10] We shall use the notation x_c^t to denote the stance taken by candidate $c \in C$ in the primary ($t = 1$) or general election ($t = 2$).

The sequence of moves in the game is as follows:

Step 1. Primary candidates take a policy stance by selecting a primary platform $x_c^1 \in \{X \cup \emptyset\}$. If $x_c^1 = \emptyset$, then the candidate is said to be *ambiguous*; otherwise she is said to be *clear*.

Step 2. Voters observing the primary stances and knowing only their own ideal point y_v and the common belief $G(\theta)$ select a party to affiliate with $a_v \in \{L, R\}$.

Step 3. Following this affiliation, each voter updates her beliefs about θ based on the observed random quantity $siz(L) := \mu(\{y \in X : a_v(y) = L\}, \theta)$. This information can be thought of as the percentage of voters that have registered with party L.[11] Each party then has a primary election with voters simultaneously casting votes b_v^1 over the set $\{l_1, l_2\}$ if $a_v = L$ and $\{r_1, r_2\}$ if $a_v = R$. Since the affiliation and voting actions may be responsive to a voter's ideal point, we use the notation $a_v(y)$ and $b_v^1(y)$ to denote the affiliation and ballot choices, respectively, of voter v as a function of her ideal point

10. The assumption that there are only two types of candidates is not important. What matters is that l and r candidates are on opposite sides of the median. Since we have assumed that the distribution of the median has full support on X, there can only be two types of candidates. In a model where the median has the support of θ, a proper subset of X, there could be four types of candidates.

11. Empirical counterparts to the percentage of voters affiliating with each party are voter registration data when they are publicly available or media estimates of this information which tends to surface in campaign coverage.

$y \in X$.[12] We define the quantities $v(i, z) := \mu(\{y \in X : b_v^1(y) = z_i\}, \theta)$ for $z \in \{l, r\}$ and $i \in \{1, 2\}$. These quantities are random because they depend on the random variable θ. The primary winners are selected by plurality rule with

$$l := \begin{cases} l_1 \text{ with probability 1 if } v(1, l) > v(2, l) \\ l_2 \text{ with probability 1 if } v(1, l) < v(2, l) \\ \text{lottery with probability } \frac{1}{2} \text{ for each candidate otherwise} \end{cases} \quad (3)$$

$$r := \begin{cases} r_1 \text{ with probability 1 if } v(1, r) > v(2, r) \\ r_2 \text{ with probability 1 if } v(1, r) < v(2, r) \\ \text{lottery with probability } \frac{1}{2} \text{ for each candidate otherwise} \end{cases}$$

Step 4. Then the general election candidates l and r update their beliefs about θ based on the information $\{siz(L), v(1, l), v(2, l), v(1, r), v(2, r)\}$ and simultaneously announce policies, x_c^2, subject to the constraint that $x_c^2 \in B(x_c^1)$ where

$$B(x_c^1) = \begin{cases} X & \text{if } x_c^1 = \emptyset \\ \{x_c^1\} & \text{if } x_c^1 \in X \end{cases} \quad (4)$$

Step 5. After observing policy stances (x_l^2, x_r^2) voters simultaneously cast ballots, $b_v^2(y_v)$, over $\{l, r\}$ and the winning candidate $w \in \{l, r\}$ is decided by plurality rule. That is,[13]

$$w = \begin{cases} l \text{ with probability 1 if } \mu(\{y \in X : b_v^2(y) = l\}, \theta) > \frac{1}{2} \\ r \text{ with probability 1 if } \mu(\{y \in X : b_v^2(y) = r\}, \theta) > \frac{1}{2} \\ l \text{ with probability 1 if } \mu(\{y \in X : b_v^2(y) = l\}, \theta) = \frac{1}{2} \\ \quad \text{and } x_l^1 = \emptyset \neq x_r^1 \\ l \text{ with probability 0 if } \mu(\{y \in X : b_v^2(y) = l\}, \theta) = \frac{1}{2} \\ \quad \text{and } x_r^1 = \emptyset \neq x_l^1 \\ \text{lottery with probability } \frac{1}{2} \text{ for each candidate otherwise} \end{cases} \quad (5)$$

The winning candidate, w, enacts policy x_w^2.

12. We assume that primary voters know the affiliation results when they cast primary votes. The alternative assumption, that primary voters do not know the affiliation results, leads to a more complicated analysis.

13. The resolution of ties in (5) may seem ad hoc. But the resolution of an open set problem in some general elections requires that when candidate c has committed in the primary and candidate $-c$ has not, $-c$ will win with probability 1 if she makes the median voter indifferent between the two candidates. An alternative approach would be to look for ε-equilibria general election stances that allow candidates to select policies that are arbitrarily close to optimal. In this case, an ε-equilibria stance of the second mover would be arbitrarily close to the move characterized in the equilibrium to the game with this tie-breaking rule.

The game is one of perfect recall and complete (but imperfect) information. To simplify future expressions, we introduce the following summaries of histories which track information that agents can use to form beliefs. A history h^1 is a vector of primary platforms $\mathbf{x}^1 = (x^1_{l_1}, x^1_{l_2}, x^1_{r_1}, x^1_{r_2})$. A history h^2 is the measure of voters that affiliated with party L and a history h^1. Formally, a generic h^2 is a pair $(siz(L), h^1)$. A history h^3 is a vector of general election stances $\mathbf{x}^2 = (x^2_l, x^2_r)$, primary voting results $(v(1, l), v(1, r))$, and a history h^2. When a history h^3 is observable, the values $v(2, l)$ and $v(2, r)$ are also known. This information is redundant and, therefore, suppressed.[14]

2.2 Equilibrium Concepts

As summarized in the introduction, we characterize an equilibrium in which primary candidates are ambiguous and the primary winners learn the location of the median voter, θ, from the observed $siz(L)$ and knowledge of the equilibrium strategies of voters. In this section, we present the relevant concepts to define a notion of equilibrium which requires that agents play simultaneous best responses and process information in a reasonable or efficient manner. There are two complications in the game that require the use of non-standard concepts. The first is that the game form has simultaneous moves and following a profile of simultaneous stances, $\mathbf{x}^1$, candidates need to draw inferences about the uncertain parameter θ. Since existing concepts do a poor job at pinning down the beliefs of agents at histories that are off the equilibrium path, we impose a few additional restrictions to ensure that an equilibrium initial stance, $\mathbf{x}^1$, is supported by strategies in which candidates anticipate drawing reasonable inferences from the observed $siz(L)$ off the path. We do not pin down beliefs at histories that require more than one candidate deviation from the equilibrium initial stance, $\mathbf{x}^1$, as the play at these histories has no influence on the desirability of deviations from the equilibrium path of play.

The second issue is that since there is a continuum of voters, a voter's actions are payoff inconsequential. We address this problem by requiring voters to play strategies that are weakly undominated and simultaneous best responses in a family of games that differ from the current game only in that the agent has control of a set of voters that has positive measure. In the remainder of this section, we formalize these concepts.

14. We suppress the history $(\mathbf{x}^1, siz(L), v(1, l), v(2, l))$ which contains information visible after the primary because in the equilibria that we consider candidates need only use the information available in h^2 to learn θ. To save on notation, we use the histories h^2 to track the information that candidates in the general election and voters in primaries use when forming their posteriors on θ.

2.2.1 Strategies. We will focus only on strategies that are symmetric in the sense that all voters with the same ideal point use the same affiliation and voting strategies. This assumption allows us to drop the v subscript in the term y and functions $a(y), b^1(y), b^2(y)$. Since this is a game with multiple stages, strategies are functions of the histories defined earlier. Thus, $a(y, h^1)$ denotes the affiliation choice of a voter with ideal point y having observed history h^1. Similarly, we use the notation $b^1(y, h^2), b^2(y, h^3)$. The general election stances follow a history h^2 and are denoted $x_c^2(h^2)$. In order to resolve open set problems when voters are indifferent between their pure strategies, it is convenient to allow voters to mix in the affiliation and voting stages. With a slight abuse of notation, we denote lotteries over the available pure actions as functions of the ideal point, given the relevant histories, as $\tilde{a}(\cdot, h^1) : X \to [0, 1]$, $\tilde{b}^1(\cdot, h^1) : X \to [0, 1]$, $\tilde{b}^2(\cdot, h^3) : X \to [0, 1]$. The interpretation is that $\tilde{a}(y, h^1)$ is the probability that a player with ideal point y that has observed history h^1 will affiliate R. Similarly, if a voter with ideal point y has affiliated L (R), then $\tilde{b}^1(y, h^1)$ denotes the probability she will vote l_2 (r_2) in the primary, and $\tilde{b}^2(y, h^3)$ is the probability she will vote for $l(r)$ in the general election.[15]

Combining these modifications allows us to define a symmetric strategy profile.

Definition 1: *A symmetric strategy profile σ to the game is a vector of vectors and mappings*

$$\sigma = \left\langle x^1, \tilde{a}(\cdot, h^1), \tilde{b}^1(\cdot, h^2), \{x_c(h^2)\}_{c \in C}, \tilde{b}^2(\cdot, h^3)\right\rangle$$

The simplest equilibria, in which candidates learn voter preferences, involve learning voter preferences from just the affiliation choices. Accordingly, we focus on equilibria where, in updating their beliefs about θ, general election candidates condition on only the affiliation choices ($siz(L)$) as opposed to the primary voting choices. We term as fully-revealing for h^1 affiliation strategies in which for a given h^1 voters with ideal points less than some number π_{h^1} affiliate L and voters with ideal points higher than π_{h^1} affiliate R. In the next subsection, we will demonstrate why the term fully-revealing is appropriate. We focus on equilibria of this form because they are the simplest equilibria in which candidates learn θ. We define a fully revealing profile in the following manner.

15. When a non-null set of voters mix, we are agnostic as to whether a law of large numbers holds for the continuum (Judd, 1985). Specifically, if it does, then when voters mix a tie can occur with positive probability. In this case, Equations 3 and 5 impose tie-breaking rules. If it does not hold, then even if all voters mix with non-degenerate probabilities, there will still be one candidate who gets the highest measure of votes. We do, however, assume a construction in which the events $\{m$ percent of voters vote for $c\}$ are measurable. Results by Judd justify this assumption.

DEFINITION 2: *A symmetric strategy profile σ is fully-revealing, if for all h^1, which differ from $\mathbf{x}^1$ in σ by at most one coordinate, σ is fully-revealing for h^1. That is, affiliation is of the following form for some $\pi_{h^1} \in (0, 1)$:*

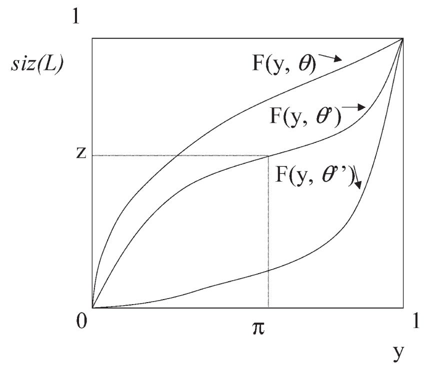
$$\tilde{a}(\cdot, h^1) = \begin{cases} 1 & \text{if } y > \pi_{h^1} \\ 0 & \text{if } y < \pi_{h^1} \\ \frac{1}{2} & \text{otherwise} \end{cases}$$

In other words, a symmetric strategy profile σ is fully-revealing if, on the path, affiliation is of this cutpoint form and for any path following a deviation by only one primary candidate from the stances that σ calls for, affiliation strategies are also of this simple cutpoint form. It should be noted that π_{h^1} need not be the same for each h^1. However, in the equilibrium characterized π_{h^1} will be the same for all h^1 in which at most one primary candidate is clear.

2.2.2 Beliefs. In this section, we limit the possible posterior beliefs that agents can have about θ. Since candidates that take stances in the general election have observed h^2, it is necessary to define posterior beliefs on θ conditional on a history h^2. To ensure that an equilibrium with ambiguity in the primary is supported by strategies that are best responses by agents with reasonable beliefs, we will constrain beliefs at histories in which at most one primary candidate is clear. Given a fully-revealing symmetric strategy σ, in which $\mathbf{x}^1 = (\emptyset, \emptyset, \emptyset, \emptyset)$, at any history h^1 that differs from $\mathbf{x}^1$ by one coordinate, affiliation is of the cutpoint form in Definition 2. This means that at any such history h^2, agents know that $siz(L) = F(\pi_{\mathbf{x}^1}, \theta)$. Recall that we have assumed that, for all $z \in (0, 1)$, $F(z, \theta) \neq F(z, \theta')$. This means that at any history h^2 in which at most one primary candidate is not ambiguous and for any observed $siz(L) \in (0, 1)$, there is exactly one θ which would induce the strategy $\tilde{a}(\cdot, h^1)$ to yield the observed $siz(L)$. Specifically, this θ solves the equation $siz(L) = F(\pi_{\mathbf{x}^1}, \theta)$.

Figure 1 illustrates how the assumption that $F(z, \theta)$ is ordered by first-order stochastic dominance in θ results in this fact. The figure exhibits three different distributions $F(\cdot, \theta)$, $F(\cdot, \theta')$, $F(\cdot, \theta'')$. Given the equilibrium cutpoint, $\pi_{\mathbf{x}^1}$, only one of the distribution functions is consistent with the observed $siz(L)$. Since there is a one-to-one mapping between the distribution functions and the parameter θ, this means that there is only one possible θ consistent with the observed $siz(L)$ and equilibrium cutpoint for history h^2. Accordingly, the posterior belief of a candidate at such a history h^2 should assign probability 1 to the θ that solves $siz(L) = F(\pi_{\mathbf{x}^1}, \theta)$.

We use this logic to define the types of beliefs that agents can have about θ based on observing a history h^2 and knowledge that players are using a particular profile σ.



siz(L) = F(π, θ')

If π is used and *siz(L)*=*z* is observed, then the median ideal point is θ'.

Figure 1. Inferring the Median

Definition 3: *A belief is a distribution function on* Θ *conditional on history* h^2. *Given a fully-revealing symmetric strategy profile* σ, *a belief* $\rho(\theta \mid h^2, \sigma)$ *is consistent if following any history* $h^2 = (\mathbf{x}^1, siz(L))$ *in which at most one primary candidate is clear*

$$\rho(\theta \mid h^2, \sigma) = \begin{cases} 1 & \text{if } \theta \geq \theta^*(\mathbf{x}^1, siz(L)) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where $\theta^*(\mathbf{x}^1, siz(L))$ solves $siz(L) = F(\pi_{\mathbf{x}^1}, \theta^*(\mathbf{x}^1, siz(L)))$.

We reiterate that when σ is fully-revealing and at most one primary candidate is not ambiguous, $\rho(\cdot \mid h^2, \sigma)$ is a step function taking value 0 for θ less than the true value, and value 1 for θ greater than or equal to the true value.[16] While it is not immediately clear that (6) corresponds to Bayes' rule, this result can be attained as follows. We know that the conditional expectation $E(1_{\{\theta=\theta'\}} \mid \{siz(L) = \mu(\{y : a(y, h^1) = L\}, \theta)\})$ is equivalent to the posterior probability that $\theta = \theta'$ conditional on history h^2 and knowledge that strategy σ is played. This conditional expectation is defined by the condition[17]

16. We will sometimes refer to beliefs that are concentrated at a point, meaning beliefs of this form.

17. This condition is the definition of a version of conditional probability (Durret, 1996: 219).

$$\int_{\{\theta: siz(L)=\mu(\{y:a(y,h^1)=L\},\theta)\}} E(1_{\{\theta=\theta'\}} \mid \{siz(L)=\mu(\{y:a(y,h^1)=L\},\theta)\})dG(\theta)$$

$$= \int_{\{\theta'': siz(L)=\mu(\{y:a(y,h^1)=L\},\theta'')\}} E(1_{\{\theta=\theta'\}})dG(\theta) \qquad (7)$$

But we have already argued that the set $\{\theta'' : siz(L) = \mu(\{y : a(y, h^1) = L\}, \theta'')\}$ is a singleton when σ is fully-revealing. This means that the right-hand side of (7) is 1 for $\theta' = \theta^*(x^1, siz(L))$ and 0 everywhere else. Thus, the only posterior belief conditional on a history h^2 and knowledge of the equilibrium strategies σ that is a version of conditional probability and, therefore, satisfies Bayes' rule, is of the form in Definition 3. Definition 3 says nothing about beliefs at histories h^2 which have more than one clear primary candidate, or in which strategies are not symmetric and fully-revealing.

2.2.3 Equilibrium. Equilibrium concepts like perfect Bayesian equilibrium impose no limitations on the beliefs that agents have at histories h^2 that occur with probability zero in equilibrium. While we will not pin down beliefs everywhere, we will ensure that the equilibrium attained is supported by primary candidates anticipating reasonable beliefs should they deviate from the equilibrium primary stances. It is for this reason that Definition 3 limits the types of beliefs that agents have at histories where one primary candidate is clear.

Definition 4: *A fully-revealing equilibrium (FRE) is a symmetric strategy profile and belief* $(\sigma, \rho(\cdot \mid \cdot, \sigma))$ *such that: (i)* $\rho(\cdot \mid \cdot, \sigma)$ *is consistent given* σ*; (ii) at each history, with an initial stance* $\mathbf{x}^1$ *differing from the equilibrium stance by at most one coordinate, the candidate actions called for by* σ *are simultaneous best responses when expected utilities are calculated with* $G(\theta)$ *at an initial history and* $\rho(\cdot \mid \cdot, \sigma)$ *otherwise; (iii) at each history, with an initial stance* $\mathbf{x}^1$ *differing from the equilibrium stance by at most one coordinate, the voter actions called for by* σ *are simultaneous best responses given the prior* $G(\theta)$ *for affiliation and the belief* $\rho(\cdot \mid \cdot, \sigma)$ *at all subsequent information sets; (iv) the strategy profile* σ *is fully-revealing.*

Several comments are in order. First, we have not defined what beliefs are if primary voting is responsive to voter ideal points and, therefore, conveys information. This is not problematic because (by Condition (iv)) if an FRE exists then the affiliation choices along the path (and off the path by one deviation in initial primary platforms) must provide all information about θ. In other words, both on the path of an FRE and for a single candidate primary platform deviation, beliefs will be pinned down by Definition 3 in a way that is reasonable. Second, it should be noted that Definition 3 is stronger than the notion of consistency normally used in a perfect Bayesian equilibrium. We are making a selection here that requires that in choosing primary stances candidates evaluate deviations from a conjectured profile by considering their payoff in a history where θ is known by the general election – a conjecture that would be self-fulfilling if this deviation occurred. Third, we are not pinning down best responses (or requiring their existence) at information sets that are irrelevant in determining if an agent has an incentive to deviate from the conjectured profile of equilibrium actions. Specifically, in supporting an equilibrium where all candidates are ambiguous in the primary, the desirability of a deviation by one primary candidate does not hinge on knowing what would happen in a general election following primary deviation by two, three or four candidates.[18]

2.2.4 Atom proofness. The fact that a voter's actions are never payoff-consequential because there is a continuum of voters introduces a complication, as weak dominance is not a satisfactory means to select from a possibly large set of equilibria. In the current model, every policy is supportable as the final policy in an FRE. That is, because no voter has influence, FRE neither pins down the way voters vote nor how they affiliate. Contemporary models of politics with a continuum of voters (Banks and Duggan, 2000) assume that in a voting stage voters vote sincerely. This restriction seems reasonable and often serves to induce results that are consistent with similar models having a large but finite population. In the current paper, voters take actions on which strategic players condition their posterior beliefs. As such, it is not so clear what action to attribute to one of an uncountable infinity of agents. We are not aware of other papers that address this issue.

Our resolution of this problem is to require agents to take actions that would be best responses if they were able to control the actions of a small (but positive measure) set of their neighbors. Let $\sigma_y^{(h)}$ denote the action that strategy profile σ calls for a voter with ideal point $y \in X$ to take at history h. We now define a game which is slightly different from the original game.

18. While, in principle, an FRE need not be a perfect Bayesian Equilibrium, there are FRE which are perfect Bayesian equilibria. Given strategies and beliefs satisfying the FRE conditions we can impose beliefs that are concentrated at histories that have at least two candidates unambiguous (unreasonable but satisfying Bayes' rule which is undefined). Given these beliefs, simultaneous best responses will exist at all subsequent histories.

DEFINITION 5: *We refer to the* (y, ε)*-relative game (for* $\varepsilon > 0$*) as the extensive form game which differs from the original game only in that* $\mu(\cdot, \theta)$ *is replaced with the following measure*

$$\mu_\varepsilon^y(A, \theta) = \begin{cases} \mu(A, \theta) & \text{if } A \cap [y - \varepsilon, y + \varepsilon] = \emptyset \\ \mu([y - \varepsilon, y + \varepsilon], \theta) + \mu(A \backslash \{[y - \varepsilon, y + \varepsilon]\}, \theta) & \text{if } y \in A \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The measure $\mu_\varepsilon^y(A, \theta)$ in (8) has an atom at y with measure $\mu([y - \varepsilon, y + \varepsilon], \theta)$. The (y, ε)-relative game differs from the standard game only in that now a voter with ideal point y has control of the actions that his close neighbors cast, i.e. an agent with ideal point y is an atom. Given this concept we define the refinement.

DEFINITION 6: *An FRE* $(\sigma, \rho(\cdot \mid \cdot, \sigma))$ *is atom proof (denoted APFRE) if for almost everywhere* $y \in X$ *for some* $\delta > 0 :$ *(1) at every* h*, the action* $\sigma_y^{(h)}$ *is weakly undominated in all* (y, ε)*-relative games with* $\varepsilon \in (0, \delta]$ *and (2) at every* h *that occurs on the path of* σ *or any* h^1 *that differs from the one called for by* σ *in only one coordinate of* x^1, $\sigma_y^{(h)}$ *is a simultaneous best response to every remaining voter playing as specified in* σ *in all* (y, ε)*-relative games for* $\varepsilon \in (0, \delta]$.

It is not difficult to see (and we establish in Lemma 1) that, in any APFRE,

$$\tilde{b}^2(y, h^3) = \begin{cases} 0 & \text{if } |x_l - y| < |x_r - y| \\ 1 & \text{if } |x_l - y| > |x_r - y| \end{cases} \tag{9}$$

since there are only two candidates. That is, atom proofness requires voters to vote sincerely at a general election voting stage. So that unless a voter is indifferent between the two candidates, in an APFRE she will vote for the candidate with the preferred stance. This is true because in any (y, ε)-relative game weak dominance requires voting as if pivotal.

It is less obvious how APFRE restricts the possible party affiliation and primary voting strategies. Intuitively, APFRE requires that voters take actions that are consistent with the way they would like to move the beliefs of the receiver. More precisely, atom proofness requires that voters who would like candidates to believe the measure of voters in party L (R) is larger than what they anticipate it being, will join party L (R). To clarify the role of (y, ε)-relative games in the analysis, we note that these games do not occur but rather serve as a fiction that players imagine they may be playing. As such, the modeled notion of rationality is that when a player is small (one of a continuum of players), she selects actions that are best responses (and weakly undominated) to a neighborhood of games in which she controls the moves of her neighbors and is, therefore, in some sense big. The fact that we look at small ε corresponds to looking at (y, ε)-relative games where the voter is big, but not too big.[19]

3. Results

We now establish the existence of an APFRE in which all primary candidates are ambiguous, party affiliation allows the general election candidates to infer the true θ and the general election candidates announce platforms that correspond to θ.

3.1 Intermediate Results

In this section, we present several lemmas that characterize the play of agents at relevant histories. We first establish several results that help pin down the simultaneous best responses in general elections.

LEMMA 1: *In any APFRE general election voting satisfies* (9).

The intuition behind this result was discussed in the previous section. Weak dominance in (y, ε)-relative games corresponds to voting for the candidate with the preferred platform. We will call voting satisfying Equation 9 sincere. Note that (9) is not a complete description of voting, as it does not state how voters behave when candidates are located at the same policy. The resolution of this form of indifference is clarified in results later. We now characterize the simultaneous best response stances of candidates l and r at possible h^2 histories.

Given two general election platforms, the winner w is defined by (5). We use the notation $1_{\{w=c\}}(x_l, x_r, \theta)$ to denote the indicator function taking value 1 if following x_l, x_r and given θ the winning candidate is c and 0 otherwise. The term θ appears as an argument because the winner depends on θ. At a history h^2, candidate announcements must be best responses to each other. Given a history h^2 and posterior distribution $\rho(\cdot, h^2)$ on Θ, the candidates' simultaneous problems are

19. The reader should note that APFRE only pins down the behavior of a set of voters with full measure. This allows us not to worry about how a player with an ideal point that corresponds to the discontinuity in $\tilde{a}(\cdot, h^1)$ would play in (y, ε)-relative games. For an ideal point other than this point of discontinuity, the best responses are eventually constant in ε as $\varepsilon \downarrow 0$.

$$x_l^2 \in \arg\max_{x \in B(x_l^1)} \left\{ \begin{array}{c} \int_\Theta((1-\kappa)+\kappa\gamma((x)^2))1_{\{w=l\}}(x, x_r^2, \theta)d\rho(\theta, h^2) \\ +\int_\Theta \kappa\gamma((x_r^2)^2)1_{\{w=r\}}(x, x_r^2, \theta)d\rho(\theta, h^2) \end{array} \right\} \tag{10}$$

$$x_r^2 \in \arg\max_{x \in B(x_r^1)} \left\{ \begin{array}{c} \int_\Theta((1-\kappa)-\kappa\gamma((1-x)^2))1_{\{w=r\}}(x_l^2, x, \theta)d\rho(\theta, h^2) \\ +\int_\Theta \kappa\gamma((1-x_l^2)^2)1_{\{w=l\}}(x_l^2, x, \theta)d\rho(\theta, h^2) \end{array} \right\} \tag{11}$$

The following lemmas establish the equilibrium play of two candidates l and r with common beliefs $\rho(\cdot, h^2)$.

Lemma 2: *If $\rho(\cdot, h^2)$ is concentrated at $\hat{\theta}$, voting is sincere, and $x_l^1 = x_r^1 = \emptyset$, the unique simultaneous best responses of the candidates are $(\hat{\theta}, \hat{\theta})$. This profile is supported with voters resolving a tie with a fair coin. The winner is l with probability $\frac{1}{2}$ and r with probability $\frac{1}{2}$.*

Lemma 2 is just a restatement of the median voter theorem. In the case in which both candidates are constrained, the choices at history h^2 are trivial. In the case of one constrained and one unconstrained candidate, the problem is similar to the models considered by Romer and Rosenthal (1978) and Bernhardt and Ingberman (1985).

Lemma 3: *If $\rho(\cdot, h^2)$ is concentrated at $\hat{\theta}$, voting is sincere, and $x_c^1 = \emptyset$, $x_{-c}^1 = x$, then the unique simultaneous best responses of the candidates (x_l^2, x_r^2) are* (i) $(x, 2\hat{\theta} - x)$ *if* $x < \hat{\theta}$ *and* (x, x) *if* $x \geq \hat{\theta}$, *for* $-c = l$ *and* (ii) $(2\hat{\theta} - x, x)$ *if* $x > \hat{\theta}$ *and* (x, x) *if* $x \leq \hat{\theta}$, *for* $-c = r$. *Candidate c wins the election with probability 1.*

The intuition behind Lemma 3 is that when candidates know the median voter's ideal point, and one candidate is constrained to a policy, the other unconstrained candidate will enact the policy closest to her ideal point that makes the median voter indifferent between the two candidates.

These lemmas establish the unique simultaneous best responses at the general election stage when beliefs are concentrated. Since APFRE only pins down play at histories that are close to the equilibrium path, we will only characterize play at histories on the path and off the path following any unilateral deviation. As long as we can show that beliefs are concentrated following a primary when at most one candidate is not ambiguous it will not be necessary to characterize the general election play of candidates when beliefs are not concentrated.[20]

3.2 The Existence of APFRE

In this section we establish the main result, the existence of an APFRE in which all primary candidates are ambiguous. The argument hinges on two key points. The first point is that since $\mu(\cdot, \theta)$ first-order stochastically dominates $\mu(\cdot, \theta')$ for $\theta > \theta'$, when affiliation is of the cutpoint form, candidates knowing the cutpoint and observing $siz(L)$ can infer θ. The second point is that since the general election candidates will have learned θ from affiliation a primary candidate that constrains herself is vulnerable in the general election.

We verify that following an initial stance vector $\mathbf{x}^1$, which has either all primary candidates ambiguous or just one primary candidate clear, simultaneous best responses at histories h^1 and h^2 exist that satisfy the conditions imposed by APFRE. Once this fact is established (Lemma 4), we show that taking one such set of simultaneous best responses to histories h^1, h^2, h^3 and adding $\mathbf{x}^1 = (\emptyset, \emptyset, \emptyset, \emptyset)$ is an APFRE (Proposition 1). Establishing this last fact hinges on showing that given the actions that would occur at an h^1 in which exactly one candidate is clear in the primary the unilateral deviation from $(\emptyset, \emptyset, \emptyset, \emptyset)$ is not desirable.

Without loss of generality we write the best responses only for the case of candidate l_1 being clear.

LEMMA 4: *Given any initial stance profile* $\mathbf{x}^1 \in \{(\emptyset, \emptyset, \emptyset, \emptyset), (x, \emptyset, \emptyset, \emptyset)\}$ *(for* $x \in X$*), simultaneous best responses at histories* h^1 *and* h^2 *that satisfy the conditions imposed by APFRE exist, and the best responses are of the form:*

$$\tilde{a}(y, h^1) = \begin{cases} 1 & \text{if } y > \pi \\ 0 & \text{if } y < \pi \\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{12}$$

20. Under the candidate preference assumption pure strategy simultaneous best responses for two candidates facing uncertainty about θ may not exist (Ball, 1999). By not requiring the existence of simultaneous best responses at every possible type of election, the equilibrium concept used allows us to establish an equilibrium in which θ is learned prior to the general election on the path, and following any deviation by a single agent.

$$\tilde{b}^1(y, h^2) = \begin{cases} 1 & \text{if } y > \pi \text{ and } \Pi(y, h^2) < 0 \\ 0 & \text{if } y > \pi \text{ and } \Pi(y, h^2) > 0 \\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{13}$$

for some $\pi \in (0, 1)$[21] *where*

$$\Pi(y, (x', \emptyset, \emptyset, \emptyset), siz(L)) := \int \psi((\theta - y)^2) d\rho(\theta, h^2) - (1 - \rho(x', h^2)) \int \psi((2\theta - x' - y)^2) d\rho(\theta, h^2) - \rho(x', h^2)\psi((x' - y)^2) \tag{14}$$

and

$$\Pi(y, (\emptyset, \emptyset, \emptyset, \emptyset), siz(L)) := 0. \tag{15}$$

Much of the intuition behind Lemma 4 hinges on observing that voters individually have an incentive to use affiliation strategies of the form in (12). Given that at most one primary candidate is clear, primary voters anticipate that on the path of play the final policy will correspond to the candidates' perception of the median voter's ideal point, θ. This is true because two unconstrained general election candidates that believe they know the location of θ with probability 1 will announce policies that correspond to the median voter's ideal point (Lemma 2). Additionally, since knowledge of the equilibrium cutpoint π and the observed $siz(L)$ allow the candidates to infer θ (as follows from the argument following Definition 3 and Figure 1), the general election candidates will announce positions that correspond to the actual median voter's ideal point. Given this and the fact that the beliefs candidates form are responsive to $siz(L)$, a voter deciding which way to affiliate must balance the possible gains and losses from increasing or decreasing $siz(L)$. The equilibrium cutpoint value π has the property that voters with ideal points less (greater) than π would prefer in expectation to mislead candidate beliefs in the downward (upward) direction. In a (y, ε) -relative game, with ε sufficiently small, a voter with ideal point $y < (>)\pi$ wishes (in the sense of expected utility) to decrease (increase) the candidates estimate of θ by at least a little. She therefore wants to increase (decrease) $siz(L)$ by at least ε. This corresponds to a strict incentive to affiliate L (R), since this action increases (decreases) $siz(L)$.

21. The characterization of π appears in the proof of the result.

At the level of primary voting, the intuition is as follows. Primary voters facing two ambiguous candidates are indifferent between the outcome of their primary (as $\Pi(y, (\emptyset, \emptyset, \emptyset, \emptyset)) = 0$) and, therefore, can toss a coin. Primary voters facing one constrained and one ambiguous candidate anticipate that if the ambiguous candidate wins the primary the final outcome will correspond to the ideal point of the median voter (Lemma 2) and if the constrained candidate (say l_1) wins the primary the final outcome will correspond to $\max\{x_l^1, 2\theta - x_l^1\}$ (Lemma 3). The formula in (14) is the difference in expected utility between the final policies θ and $\max\{x_l^1, 2\theta - x_l^1\}$. While clear candidates may receive a positive measure of primary votes, the facts that $\max\{x_l^1, 2\theta - x_l^1\} \geq \theta$ and the median voter in party L has an ideal point that is less than θ imply that the clear candidate will lose the primary. A last complication is to verify that no voter that is supposed to affiliate with R would prefer to affiliate L and boost l_1's chances of winning the primary. This possibility is ruled out by the fact that for sufficiently small ε in any (y, ε)-relative game such a deviation will not change the probability that l_1 wins the primary (as $v(1, l) < v(2, l)$ so that a small change in $v(2, l)$ is of no consequence), but such a deviation will move $siz(L)$ in the wrong direction (which does move the final policy in the wrong direction with probability 1).

Inspection of (12) and (13) indicates that primary candidates have no incentive to unilaterally deviate from ambiguity. The relevant conclusion is that if one candidate is clear and the other candidates are ambiguous, then the clear candidate loses the primary with probability 1. If all primary candidates are ambiguous, then the probability that any $c \in C$ wins the primary is $\frac{1}{2}$, the probability that c wins the general election is $\frac{1}{4}$ and the final policy is θ with probability 1. Thus, given that three primary candidates are ambiguous the fourth primary candidate has a strict incentive to also be ambiguous. This logic leads to the main result.

Proposition 1: *There exists an APFRE in which the path of play is*

$$x^1 = (\emptyset, \emptyset, \emptyset, \emptyset)$$

$$\tilde{a}(y, h^1) = \begin{cases} 1 & \text{if } y > \pi \\ 0 & \text{if } y < \pi \\ \frac{1}{2} & \text{otherwise} \end{cases} \tag{13}$$

$$\tilde{b}^1(y, h^1) = \tfrac{1}{2}$$

$$x_l = x_r = \theta$$

$$\tilde{b}^2(y, h^1) = \tfrac{1}{2}$$

for some $\pi \in (0, 1)$.

Several points should be noted about the equilibrium path.

(i) All primary candidates are ambiguous.
(ii) Party affiliation follows a cutpoint strategy and fully reveals θ.
(iii) If one primary has precise candidates and the other has ambiguous candidates, primary voting will be sincere.
(iv) In the general election the party candidates announce platforms corresponding to the actual median voter's ideal point.

The ambiguity in the primary results for two reasons. First, candidates anticipate that in the future they will be in a better position to select platforms, so that unilaterally committing makes a candidate vulnerable. Second, the primary median voters anticipate that if a candidate can do so and still win she will move policy toward the party median's ideal point. As such, the median party voter has no incentive to reward a primary candidate that commits because such a candidate is vulnerable and offers nothing over a candidate from the party who is unconstrained.

4. Discussion

We first consider narrow questions about extensions and then consider broad questions of interpretation. In terms of preferences, the assumption that candidates value office and policy and are homogeneous with respect to the weight they put on each of these components can be relaxed. The critical feature is that at histories where θ has been learned candidate preferences induce the same type of play that occurs in the characterized equilibrium. The symmetry assumption regarding voter preferences can be relaxed without any qualitative changes. Introduction of heterogeneous asymmetry is problematic. In this case, θ is not a sufficient statistic for candidates in general elections involving one ambiguous and one constrained candidate. Analysis when voter preferences do not admit a core for every θ – either through relaxing the single-peakedness assumption or increasing the dimensionality of the policy space – is quite difficult. Little can be said about the extent to which the intuition of the model generalizes in these ways.

While the model does not impose many restrictions on the distribution of ideal points and the prior beliefs that players have over θ, it does require that jointly these primitives satisfy the monotonicity condition: if $\theta < \theta'$ then $F(\cdot, \theta)$ strictly first-order stochastically dominates $F(\cdot, \theta')$. If this condition is not satisfied, then affiliation alone cannot fully reveal θ. In this case, the characterized equilibria cannot be attained and it is not clear what the resulting equilibria are like.

In the present game, there are two open questions. Are there also equilibria in which no primary candidates are ambiguous? Are there equilibria in which all primary candidates are ambiguous but no information is conveyed by affiliation and primary voting? It is not likely that the former attain. If following a profile in which three candidates commit and one is ambiguous, information about θ is revealed by affiliation and voting, there will be a strong incentive for a candidate to be ambiguous when three candidates are clear. The difficulty lies in demonstrating that it is not possible to support strategies in which no information about θ is learned following such a primary profile. Pooling equilibria (with mixed affiliation strategies) in which all candidates are ambiguous exist. If candidates anticipate that voters of each type will affiliate with each party with equal probability, then such affiliation strategies are optimal. Moreover, any candidate that unilaterally commits will either be vulnerable (if θ is learned following such a deviation), or will be no more attractive than an uncommitted candidate (if θ is not learned). Accordingly, it is possible to construct voting strategies that support this type of equilibrium.

Broadly speaking, we address two seemingly unrelated phenomena, candidate ambiguity and the effect of party primaries in a unidimensional model of electoral politics in which agents have imperfect information about the preferences of the electorate. The equilibrium explanation is that affiliation with parties conveys information to general election candidates and this affiliation takes place in the primaries. Ambiguity then is a best response during the primary, as it has an option value. The informational content of primaries in the model is explained as follows. In equilibrium, voters sort into parties according to their preferences. Voters with ideal points to the left (right) of some cutpoint affiliate with party L (R). General election candidates that observe the number of L and R party affiliators can infer the distribution of voter ideal points and, thus, the location of the median voter's ideal point. Given the fact that primary candidates anticipate learning about voter preferences after the primary, they have an incentive to remain ambiguous and, therefore, unconstrained in the primary.

While the equilibrium involves no policy competition in the primary, the paper should not be construed as an argument that primaries are electorally uninteresting. The model assumes that candidates in the same party can only differ by the policy platforms they eventually choose. In such a world, we find that competition over policy will wait until the general election. This does not mean that there are not strategic incentives for fierce competition over other candidate characteristics – which are outside the scope of this model. Accordingly, evidence of non-trivial primaries does not directly refute the validity or value of the analysis. An explanation of primary competition which is in the spirit of the model can be offered. Candidates of the same party compete quite fiercely over issues which either (1) will not turn out to be very salient in the contest between the parties or (2) are such that the candidates are already well identified with policy points.[22] In either of these cases the cost of early commitment is negligible. In contrast, strong refutation of the theory would be offered by evidence of intense primary competition and narrowing of positions on issues that turn out to be critical in the general election.

The equilibrium illustrates how the electoral predictions of the Downsian model may be preserved with policy-motivated candidates and a high degree of uncertainty about the policy preferences of the electorate. The inclusion of a primary (with endogenous voter affiliation) and the ability of candidates to remain ambiguous in the primary result in a game form in which full information aggregation is possible; hence the outcome mimics the simple Downsian model with complete and perfect information. Thus, in moving from a simple primary-free theory of elections to a theory with primaries the critical difference is information aggregation. Without primaries the electoral outcome is only centrist in expectation. With primaries the candidates select platforms that are actually centrist.[23]

The explanation that affiliation with parties in the primary serves an informational role is interesting in light of the previously discussed history of democratization of the party nomination procedures. This democratization may be desirable from an information efficiency perspective as it increases the ability of primaries to convey information about voter preferences. While it is surely unwarranted to pose this rationale as the dominant explanation for the change in nomination procedures, the plausibility of this explanation, either as a cause or effect of the reform, is raised by the analysis. This suggests that in comparing elections with nomination processes that have varying degrees of democratization a critical distinction may be the degree of informational efficiency. Finally, while the model departs from the reality of primary and general elections in several key ways, the analysis focuses on a feature of politics that is often overlooked – the presence of institutional features that help candidates or elites learn about voter preferences, and the effect that these features have on the actions of elites and the masses.

22. Issues that are often termed valence dimensions like personality or war record fall into this second category. George W. Bush was severely constrained in competition with John McCain over desirable war records.

23. In the contexts of repeated elections and polls, Pikety (2000), Shotts (2000), Meirowitz (2001), and Razin (2003) also analyze models in which early voter behavior transmits information.

REFERENCES

Aldrich, John H. (1980) 'A Dynamic Model of Presidential Nomination Campaigns', *American Political Science Review* 74: 651–69.

Alesina, Alberto and Alex Cukierman (1990) 'The Politics of Ambiguity', *Quarterly Journal of Economics* 4: 829–50.

Aragonès, Enriqueta and Zvika Neeman (2000) 'Strategic Ambiguity in Electoral Competition', *Journal of Theoretical Politics* 12(2): 183–204.

Aragonès, Enriqueta and Andrew Postlewaite (2002) 'Ambiguity in Election Games', *Review of Economic Design* 7(3): 233–55.

Austen-Smith, David (1983) 'The Spatial Theory of Electoral Competition: Instability, Institutions, and Information', *Environment and Planning C: Government and Policy* 1: 439–59.

Ball, Richard (1999) 'Discontinuity and Non-Existence of Equilibrium in the Probabalistic Spatial Voting Model', *Social Choice and Welfare* 16: 533–55.

Banks, Jeffrey S. (1990) 'A Model of Electoral Competition with Incomplete Information', *Journal of Economic Theory* 50(2): 309–25.

Banks, Jeffrey S. and John Duggan (2000) 'A Multidimensional Model of Repeated Elections', typescript, California Institute of Technology.

Bartels, Larry M. (1988) *Presidential Primaries and the Dynamics of Public Choice*. Princeton, NJ: Princeton University Press.

Bernhardt, Daniel and Daniel Ingberman (1985) 'Candidate Reputation and the Incumbency Effect', *Journal of Public Economics* 26: 47–67.

Calvert, Randal (1985) 'Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence', *American Journal of Political Science* 28(1): 127–46.

Campbell, James E. (2000) *The American Campaign*. College Station: Texas A&M University Press.

Chappell, Henry W. Jr (1994) 'Campaign Advertising and Political Ambiguity', *Public Choice* 79(3–4): 281–303.

Davis, James W. (1980) *Presidential Primaries*. Connecticut: Greenwood Press.

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

Durrett, Richard (1996) *Probability: Theory and Examples*, 2nd edn. Belmont: Duxbury.

Fishburn, Peter C. (1972) 'Lotteries and Social Choice', *Journal of Economic Theory* 5: 189–207.

Geer, John G. and Mark E. Shere (1992) 'Party Competition and the Prisoner's Dilemma: An Argument for the Direct Primary', *Journal of Politics* 54: 741–61.

Gelman, Andrew and Gary King (1993) 'Why are American Presidential Election Campaign Polls So Variable When Votes Are So Predictable?', *British Journal of Political Science* 23: 409–51.

Glazer, Amihai (1990) 'The Strategy of Candidate Ambiguity', *American Political Science Review* 84: 237–41.

Gopoian, J. David (1982) 'Issue Preference and Candidate Choice in Presidential Primaries', *American Journal of Political Science* 26: 523–46.

Judd, Kenneth (1985) 'The Law of Large Numbers with a Continuum of IID Random Variables', *Journal of Economic Theory* 35: 19–25.

Kroszner, Randall and Thomas Stratmann (1999) 'Does Political Ambiguity Pay? Corporate Campaign Contributions and the Rewards to Legislator Reputation', typescript, University of Chicago.

Ledyard, John (1984) 'The Pure Theory of Large Two-Candidate Elections', *Public Choice* 44: 7–41.

Lengle, James (1981) *Representation and Presidential Primaries*. Connecticut: Greenwood Press.

Martinelli, César (2002) 'Simple Plurality versus Plurality Runoff with Privately Informed Voters', *Social Choice and Welfare* 19(4): 901–19.

McCarty, Nolan and Lawrence Rothenberg (1996) 'Commitment and the Campaign Contribution Contract', *American Journal of Political Science* 40: 872–904.

McKelvey, Richard (1980) 'Ambiguity in Spatial Models of Policy Formation', *Public Choice* 35: 385–402.

McKelvey, Richard and Peter C. Ordeshook (1985) 'Elections with Limited Information: A Fulfilled Expectations Model using Contemporaneous Poll and Endorsement Data as Information Sources', *Journal of Economic Theory* 36: 55–85.

McKelvey, Richard and Peter C. Ordeshook (1986a) 'Sequential Elections with Limited Information', *American Journal of Political Science* 29(4): 480–512.

McKelvey, Richard and Peter C. Ordeshook (1986b) 'Information, Electoral Equilibria, and the Democratic Ideal', *Journal of Politics* 48(4): 909-37.

Meirowitz, Adam (forthcoming) 'Polling Games and Information Transmission', *Games, and Economic Behavior*.

Norrander, Barbara (1986) 'Correlates of Vote Choice in the 1980 Presidential Primaries', *Journal of Politics* 48: 186–66.

Page, Benhamin (1976) 'The Theory of Political Ambiguity', *American Political Science Review* 70(3): 742–52.

Page, Benjamin (1978) *Choices and Echoes in Presidential Elections: Rational Man and Electoral Democracy*. London: The University of Chicago Press.

Pikety, Thomas (2000) 'Voting as Communicating', forthcoming in *Review of Economic Studies* 67: 169–92.

Polsby, Nelson W. and Aaron Wildavsky (1988) *Presidential Elections: Contemporary Strategies of American Electoral Politics*. New York: The Free Press.

Razin, Ronny (2003) 'Signaling and Election Motivations in a Voting Model with Common Values and Responsive Candidates', *Econometrica* 71(Jul.): 1083–119.

Romer, Tom and Howard Rosenthal (1978) 'Political Resource Allocation, Controlled Agendas, and the Status Quo', *Public Choice* 33: 27–43.

Schantz, Harvey L. (1996) *American Presidential Elections: Process, Policy, and Political Change*. Albany, NY: State University of New York Press.

Shepsle, Kenneth (1972) 'The Strategy of Ambiguity: Uncertainty and Electoral Competition', *American Political Science Review* 66: 555–68.

Shotts, Kenneth (2000) 'A Signalling Model of Repeated Elections', typescript, Northwestern University.

Snyder, James M. Jr (1990) 'Campaign Contributions as Investments: The US House of Representatives, 1980–1986', *Journal of Political Economy* 98: 1195–227.

Snyder, James M. Jr (1992) 'Long-term Investing in Politicians or, Give Early, Give Often', *Journal of Law and Economics* 35: 15–43.

Zeckhauser, Richard (1969) 'Majority Rule with Lotteries on Alternatives', *Quarterly Journal of Economics* 83: 696–703.

APPENDIX

We now prove the lemmas and proposition.

Proof of Lemma 1. In the text.

Proof of Lemma 2. Assume the hypothesis. Voting strategies that have candidates l and r win with probability $\frac{1}{2}$ if they have the same stances satisfy (9). To see that $x_c^2 = \hat{\theta}$ is a best response to $x_{-c}^2 = \hat{\theta}$ given this type of voting, note that for any deviation by l

she will lose with probability 1 and the final policy is unchanged. Thus, the deviation is undesirable. To see that this is the only pair of simultaneous best responses, we consider two cases. (i) Assume $x_c^2 = x_{-c}^2 < \hat{\theta}$. Without loss of generality, assume that under the voting strategy c wins with probability less than 1. Since a deviation by c to $x_c^2 + \varepsilon$ ($\varepsilon > 0$) will result in victory with probability 1, the fact that $\gamma(\cdot)$ is continuous and X is compact and, thus, $\gamma(\cdot)$ is uniformly continuous implies that for sufficiently small ε this deviation is desirable even if $y_c = 0$. If $y_c = 1$, then the deviation is clearly desirable. The case of $x_c^2 = x_{-c}^2 > \hat{\theta}$ is analogous. (ii) Assume $x_c^2 \neq x_{-c}^2$. Either $\hat{\theta} \in \{x_c^2, x_{-c}^2\}$ or $\hat{\theta} \notin \{x_c^2, x_{-c}^2\}$. If the former holds then, without loss of generality, we have $\hat{\theta} = x_c^2$. So, by (9), c wins the general election. But a deviation by c to $\lambda y_c + (1-\lambda)\hat{\theta}$ will result in c winning for $\lambda > 0$ sufficiently small, and move policy towards y_c and, thus, increase c's payoff. Thus, the deviation is desirable. If the latter holds then, without loss of generality, either $||x_c^2 - \hat{\theta}|| > ||x_{-c}^2 - \hat{\theta}||$ or there is equality. If there is inequality, then c wins with probability 0 and the final policy is x_{-c}^2, so a deviation to $(1-\lambda)x_{-c}^2 + \lambda\hat{\theta}$ will result in c winning and, for $\lambda > 0$ sufficiently small, increase c's payoff (again using the uniform continuity of $\gamma(\cdot)$). If there is equality, then at least one candidate (say $-c$) loses with positive probability and the deviation to $(1-\lambda)x_{-c}^2 + \lambda\hat{\theta}$ is preferred for sufficiently small $\lambda > 0$. □

Proof of Lemma 3. Assume the hypothesis. By assumption, $B_c(x_c^1) = \{x\}$, so it remains to show only that x is the unique best response to x for c and that this policy causes c to win with probability 1. Let $c = l$. Assume that

$$\tilde{b}^2(y, h^3) = \begin{cases} 0 & \text{if } |x_l - y| \leq |x_r - y| \\ 1 & \text{if } |x_l - y| > |x_r - y| \end{cases} \tag{17}$$

Since $0 < \kappa$, the optimal stance x_c for c wins and minimizes $|x_c - y_c|$. This occurs uniquely at the claimed solutions. To establish uniqueness note that for any sincere $\tilde{b}^2(y, h^3)$ that has $-c$ win with positive probability if $|x_l - x_c| = |x_l - x_c|$ no best response for c exists (open set problem). A similar argument establishes the claim for $c = r$. Thus, the claim is established. □

Proof of Lemma 4.[24] Applying the lemmata, it is sufficient to establish that (12) and (13) constitute weakly undominated strategies and mutual best responses at histories of the form $(\emptyset, \emptyset, \emptyset, \emptyset)$, $(x', \emptyset, \emptyset, \emptyset)$ for almost everywhere $y \in X$ for some $\delta > 0$, in the (y, ε)-relative games with $\varepsilon \in (0, \delta]$. (Note that for the histories $(\emptyset, x', \emptyset, \emptyset)$, $(\emptyset, \emptyset, x', \emptyset)$, $(\emptyset, \emptyset, \emptyset, x')$ similar arguments and expressions apply.) There are two cases:

Case 1. Assume $h^1 = (\emptyset, \emptyset, \emptyset, \emptyset)$. In this case, (15) applies and (13) reduces to

$$\tilde{b}^1(y, h^1) = \tfrac{1}{2} \tag{18}$$

24. Following the discussion regarding the assumption that voters have heterogenous preferences and common beliefs, we note that Step 2 of this argument breaks down if voters form distinct posteriors from $G(\cdot)$ and the revealed y_v.

In three steps, we show that there exists some $\pi \in (0, 1)$ such that (12) and (18) are simultaneous best responses for almost everywhere y in any (y, ε)-relative game with $\varepsilon \in (0, \delta]$ for some δ.

Step 1: Assume that (12) is played. This implies that the final outcome will have x_w^2 equal to the correct θ with probability 1 (since following such a strategy the true θ is learned and by Lemma 2 this is the stance of the candidates) regardless of which candidates win the primaries. Thus, the ballot function in (18) is outcome inconsequential in any (y, ε)-relative game. Therefore, (18) is a best response.

Step 2: Assume that (18) is played. To establish that (12) is a simultaneous best response for almost everywhere y in any (y, ε)-relative game with $\varepsilon \in (0, \delta]$ for some δ, we observe that given such a profile candidates will learn the correct θ and the final policy will correspond with the true θ with probability 1. By $x^E(\theta, siz(L), \pi)$, denote the point of discontinuity in the candidate belief (6) when $siz(L)$ is observed and affiliation satisfies (12) with cutpoint π. Clearly $x^E(\theta, siz(L), \pi) = \theta^*(\pi, siz(L))$ which we have already argued exists and is a singleton. Moreover, $x^E(\theta, siz(L), \pi)$ is strictly decreasing in its second argument. For a voter with ideal point π to be indifferent between $a(y, h^1) \in \{L, R\}$ in (y, ε)-relative game with ε arbitrarily small, it must be the case that she has no ex ante incentive to increase or decrease $siz(l)$ from the random value $F(\pi, \theta)$. We first establish the existence of a π satisfying this indifference condition and then demonstrate that for a voter with $y < (>)\pi$ higher (lower) values of $F(\pi, \theta)$ are desired and thus for ε sufficiently small voting as in (12) is a best response. These three results are sufficient for affiliation to satisfy the requirements of APRFE. We first extend $x^E(\theta, F, \pi)$ to be defined on $F \in [-1, 2]$ in the following manner: $x^E(\theta, F, \pi) = x^E(\theta, 0, \pi)$ if $F < 0$ and $x^E(\theta, F, \pi) = x^E(\theta, 1, \pi)$ if $F > 1$ for all θ, π. We now note that the problem

$$s(\pi) := \arg \max_{s \in [-1,1]} \int \psi((x^E(\theta, F(\pi, \theta) + s, \pi) - \pi)^2) dG(\theta) \tag{19}$$

is the optimization of a continuous objective function over a compact support. By the Theorem of the Maximum, the solution $s(\pi)$ is an upper hemi continuous correspondence. We now demonstrate that $s(\pi)$ is single valued for $\pi \in [0, 1]$. Assume that, for some $\pi \in [0, 1]$, there are $s' < s''$ with $s', s'' \in s(\pi)$. By (19) this implies that

$$\int \begin{bmatrix} \psi((x^E(\theta, F(\pi, \theta) + s', \pi) - \pi)^2) - \\ \psi((x^E(\theta, F(\pi, \theta) + s'', \pi) - \pi)^2) \end{bmatrix} dG(\theta) = 0 \tag{20}$$

Now since $s' < s''$ we have $F(\pi, \theta) + s' < F(\pi, \theta) + s''$ for every $\theta \in [0, 1]$. But since $x^E(\theta, siz(L), \pi)$ is strictly decreasing in its second argument (if the argument is in $[0, 1]$) and $\psi(\cdot)$ is strictly decreasing $\psi((x^E(\theta, F(\pi, \theta) + s'', \pi) - \pi)^2) < \psi((x^E(\theta, F(\pi, \theta) + s', \pi) - \pi)^2)$ for every θ. Thus,

$$\begin{bmatrix} \psi((x^E(\theta, F(\pi, \theta) + S, \pi) - \pi)^2) - \\ \psi((x^E(\theta, F(\pi, \theta) + S, \pi) - \pi)^2) \end{bmatrix} \leq 0 \tag{21}$$

for every θ, with a strict inequality for a non-null set of θ's. But this contradicts (20). Thus, $s(\pi)$ is single-valued. Since $s(\pi)$ is upper hemi continuous, this implies that it is a continuous function. Now consider the mapping $\beta : [0, 1] \to [0, 1]$ defined as $\beta(\pi) := \max\{0, \min\{\pi + s(\pi), 1\}\}$. Since $s(\pi)$ is a continuous function $\beta(\pi)$ is a continuous function from a compact set into itself. By Brouwer's fixed point theorem, there exists a $\pi^* \in [0, 1]$ such that $\beta(\pi^*) = \pi^*$. But this means that $s(\pi^*) = 0$. Thus, if voters affiliate with cutpoint π^* then a voter with ideal point π^* will be indifferent between her affiliation choices. It remains to show that for a voter with $y < (>)\pi^*$ higher (lower) values of $F(\pi^*, \theta)$ are desired. This is equivalent to showing that $s(\pi) \leq (\geq)0$ given $\pi > (<)\pi^*$. At the value π^* (characterized earlier), we have

$$0 = \arg \max_{s \in [-1,1]} \int \psi((x^E(\theta, F(\pi^*, \theta) + s, \pi^*) - \pi^*)^2) dG(\theta) \tag{22}$$

But given consistent beliefs and candidate best responses $x^E(\theta, F(\pi^*, \theta), \pi^*) = \theta$ as such. Moreover, since $x^E(\theta, F(\pi^*, \theta), \pi^*)$ is strictly decreasing in its second argument, a sufficient condition for $s(\pi) \leq (\geq)0$ given $\pi > (<)\pi^*$ is that

$$\arg \max_{s \in [-1,1]} \int \psi((\theta + z - \pi^* + \delta)^2) dG(\theta) \leq \arg \max_{s \in [-1,1]} \int \psi((\theta + z - \pi^*)^2) dG(\theta) \tag{23}$$

for $\delta > 0$ and that the inequality is $\geq$ for $\delta < 0$, for δ having sufficiently small magnitude. But given $s(\pi^*) = 0$ and the strict monotonicity of $\psi(\cdot)$ and $x^E(\theta, F(\pi^*, \theta) + s, \pi^*)$, the value $z = 0$ solves the right-hand side of (23). But this means that $z = -\delta$ solves the left-hand side of (23). Thus the ordering is attained. Thus, we have established that affiliation of the form (12) is a simultaneous best response for almost everywhere y in any (y, ε)-relative game with $\varepsilon \in (0, \delta]$ for some δ.

Step 3: Consider a deviation in both affiliation and primary voting. Since the primary voting change is payoff in inconsequential the argument in Step 2 ensures that a simultaneous deviation in both voting and affiliation is not desirable for sufficiently small δ.

Case 2: Assume $h^1 = (x', \emptyset, \emptyset, \emptyset)$ for some $x' \in X$. We show that there exists some $\pi \in (0, 1)$ such that strategies of the form (12) and (13) are simultaneous best response for almost everywhere y in any (y, ε)-relative game with $\varepsilon \in (0, \delta]$ for some δ where $\Pi(y, h^1)$ is defined by (14). We accomplish this in three steps.

Step 1: Assume that (12) is played. Given this, general election candidates will know θ and by Lemmas 2 and 3 the final policy will be θ if l_2 wins the primary and $\max\{x, 2\theta - x\}$ if l_1 wins the primary. The primary ballots of voters that have affiliated R are inconsequential so (13) is optimal for them. For voters that have affiliated L under (12) (that is they have ideal point $y < \pi$) Equation 14 is the difference in expected utility between having l_1 and l_2 win the primary. Accordingly, (13) corresponds to weakly undominated and mutual best response ballots in (y, ε)-relative games.

Step 2: Assume that (13) is played. We first note that $\max\{x, 2\theta - x\} \geq \theta$. This implies that if affiliation satisfies (12) then by (13) $v(1, L) < v(2, L)$ since under (12) the median of L is less than or equal to θ with probability 1. Given this, to establish that (12) is a mutual best response it remains only to verify that no voter has an incentive to deviate from (12) to change $siz(L)$ when the final policy will correspond to the discontinuity in (6). But the logic of case 1 establishes this fact.

Step 3: To see that no voter has an incentive to deviate in affiliation and primary voting note that since the inequality $v(1, L) < v(2, L)$ is strict, if voters affiliate by (12) a deviation by a voter with $y > \pi$ to $a(y, h^1) = L$ and ballot $\tilde{b}(y, h^1) = 0$ will not change the probability that l_1 wins the primary in (y, ε)-relative games for ε sufficiently small as $v(1, L) + \varepsilon < v(2, L)$ holds for ε sufficiently small. Thus, the simultaneous deviation will only effect the outcome through inferences about θ and thus the argument of Case 2 applies. Combining the steps yields the result. □

Proof of Proposition 1. By Lemma 4, we know that responses to $\mathbf{x}^1 = (\emptyset, \emptyset, \emptyset, \emptyset)$ and any $\hat{\mathbf{x}}^1$ with a unilateral primary candidate deviation exist which satisfy the requirements of APRFE. Moreover, the beliefs are well defined. By Case 2 of the preceding proof, we know that the expected utility to each primary candidate under the profile $(\emptyset, \emptyset, \emptyset, \emptyset)$ is $(1 - \kappa)/4 + \kappa\gamma((\hat{\theta} - y_c)^2)$, whereas the expected utility to c if c unilaterally deviates to $x_c \neq \emptyset$ is $\kappa\gamma((\hat{\theta} - y_c)^2)$. Since $\kappa \in (0, 1)$ no unilateral deviation from $\mathbf{x}^1 = (\emptyset, \emptyset, \emptyset, \emptyset)$ is desirable and, thus, this is a profile of simultaneous best responses for primary candidates. Thus, the result is established. □

ADAM MEIROWITZ is an Assistant Professor of Politics at Princeton University. His research focuses on the use of game theory to study information aggregation and preference aggregation in political institutions. ADDRESS: Department of Politics, Corwin Hall, Princeton University, Princeton NJ 08544, USA [email: ameirowi@Princeton.edu].