Web Supplement relating to “Candidate Strategies in Two-Stage Elections Beginning with a Primary,” which is forthcoming in the American Journal of Political Science, 2008

Derivations, Proofs of Theorems 1-4 including Extensions to Multicandidate Primary Elections, and a Comparison of Candidate Strategies for Expressive versus Strategic Primary Voting

James Adams and Samuel Merrill, III

Table of Contents

Proofs of Theorems 1 and 2 Pages 2-5

Derivations of Probabilities of Election for any Number of Candidates Pages 6-8

Proofs of Theorems 3 and 4 Pages 9-12

A Comparison of Candidate Strategies for Expressive versus Strategic Primary Voting Pages 13-28
Proofs of Theorems 1 and 2

The proofs of Theorems 1-2 presented here apply to the model outlined in “Candidate Strategies in Two-Stage Elections Beginning with a Primary.” See the paper for details of the two-stage election model, and also for the notation we employ below.

Proof of Theorem 1

Theorem 1. If there are two Democratic candidates $D_1$ and $D_2$ and one Republican candidate $R_1$, and if the policy distance component of the voter’s utility is concave and peaks at the voter’s ideal point, then

(a) there exists a Nash equilibrium in candidates’ office-seeking strategies,

(b) if $V_{D1} = V_{D2}$, then there is exactly one paired Nash equilibrium, i.e., for which $d_1 = d_2$.

Proof. We show that $P(D_1) = P(D_1; d_1, d_2, r_1)$ (for fixed $d_2$ and $r_1$) and $P(D_2) = P(D_2; d_1, d_2, r_1)$ (for fixed $d_1$ and $r_1$) are strictly quasi-concave (single-peaked)$^1$ on the (convex and compact) interval $[m_D, m_G]$ and that they peak in this interval. It then follows from Theorem 5.3 in McCarty and Meriowitz (2007: 108) that there exists a set of party strategies that constitute a Nash equilibrium. This will prove part (a). To see that $P(D_1)$ (and similarly $P(D_2)$), is strictly quasi-concave, note first that:

$$ P(D_1) = \frac{1}{1 + e^{[M_D(D_2) - M_D(D_1)]} + e^{[M_G(R_1) - M_G(D_1)]}}. $$  (WS1)

The first derivative of $P(D_1)$ with respect to $d_1$ is thus:

---

$^1$ If $U$ is continuous on a closed bounded interval $I$, then $U$ is strictly quasi-concave (or, equivalently, single-peaked) if $U$ has a unique local maximum on $I$ (see Roemer, 2001: 18). In particular, if $U$ is strictly quasi-concave, there exists $x_0 \in I$ such that $U(x_0) > U(x)$ for all $x \in I, x \neq x_0$. Note that if a continuous function is concave and peaks at $x_0$, then it is strictly quasi-concave.
\[
\frac{\partial P(D_1)}{\partial d_1} = \frac{g'(d_1, m_D) e^{[M_D(D_2) - M_D(D_1)]}}{\left(1 + e^{[M_D(D_2) - M_D(D_1)]} + e^{[M_G(R_1) - M_G(D_1)]}\right)^2} + \frac{g'(d_1, m_G) e^{[M_G(R_1) - M_G(D_1)]}}{\left(1 + e^{[M_D(D_2) - M_D(D_1)]} + e^{[M_G(R_1) - M_G(D_1)]}\right)^2}
\]

(WS2)

where \(g'(d_1, m_D) = \frac{\partial [g(d_1, m_D)]}{\partial d_1}\), and \(g'(d_1, m_G) = \frac{\partial [g(d_1, m_G)]}{\partial d_1}\). Now, \(P(D_1)\) reaches a local maximum if and only if \(\frac{\partial P(D_1)}{\partial d_1} = 0\) (and \(\frac{\partial^2 P(D_1)}{\partial d_1^2} < 0\)). Hence if we can show that the numerator in equation WS2 equals zero for at most one value of \(d_1\), this will establish that there exists at most one critical point for \(P(D_1)\) in the interval \([m_D, m_G]\), i.e. that \(P(D_1)\) is strictly quasi-concave on \([m_D, m_G]\). Denoting the numerator in equation WS2 by \(T(d_1)\), we have:

\[
\frac{\partial [T(d_1)]}{\partial d_1} = \left[ g''(d_1, m_D) - [g'(d_1, m_D)]^2 \right] \exp[M_G(R_1) - M_G(D_1)]
+ \left[ g''(d_1, m_G) - [g'(d_1, m_G)]^2 \right] \exp[M_D(D_2) - M_D(D_1)]
\]

(WS3)

where \(g''(d_1, m_D) = \frac{\partial^2 [g(d_1, m_D)]}{\partial d_1^2}\), and \(g''(d_1, m_G) = \frac{\partial^2 [g(d_1, m_G)]}{\partial d_1^2}\).

Recall that we have specified that the voters’ policy loss function \(g\) is concave and peaks at the voter \(i\)’s ideal point, so that \(g(x, m_D)\) peaks at \(m_D\) and \(g(x, m_G)\) peaks at \(m_G\). Now, it is clear from equation WS3 that given that voters have concave policy loss functions – i.e. given that

\(g''(d_1, m_D) \leq 0\) and \(g''(d_1, m_G) \leq 0\) for all values of \(d_1\) – then \(\frac{\partial [T(d_1)]}{\partial d_1} < 0\) for all values of \(d_1\).

Thus, there is at most a single value of \(d_1\), say \(d_1^*\), for which \(T(d_1) = 0\), which implies in turn that \(P(D_1)\) has at most a single critical point, at \(d_1^*\). To see that \(\frac{\partial^2 P(D_1)}{\partial d_1^2}\) is negative when evaluated at \(d_1 = d_1^*\), note that \(\frac{\partial P(D_1)}{\partial d_1}\) is of the form \(\frac{T(d_1)}{q(d_1)}\), so that, evaluated at \(d_1^*\),
\[
\frac{\partial^2 P(D_i)}{\partial d_i^2} = \frac{q(d_i)T'(d_i) - T(d_i)q'(d_i)}{q(d_i)^2}, \text{ because } T(d_i^*) = 0. \quad \text{But } \frac{q(d_i)T'(d_i)}{q(d_i)^2} \text{ is negative because } q(d_i) > 0 \text{ and } T'(d_i) < 0. \quad \text{It follows that } P(D_i) \text{ has at most a single local maximum, i.e. that } P(D_i) \text{ is strictly quasi-concave. This completes the proof of part (a).}
\]

To prove part (b), assume further that \(d_1 = d_2\) and \(V_{D_1} = V_{D_2}\), so that \(M_D(D_2) = M_D(D_1) = 0\). Hence the expression in equation WS3 simplifies to

\[
\frac{\partial[T(d_i)]}{\partial d_i} = \left[ g''(d_1, m_G) - [g'(d_1, m_G)]^2 \right] \left( \exp[M_G(R_1) - M_G(D_1)] \right) + \left[ g''(d_1, m_D) - [g'(d_1, m_D)]^2 \right].
\]

Now, it is clear from the above equation that when voters have concave policy loss functions – then \(\frac{\partial[T(d_i)]}{\partial d_i} < 0\) for all values of \(d_1\) – i.e. as we shift the pairing \(d_1=d_2\) rightward in the interval \([m_D, m_G]\), the value of the numerator \(T(d_i)\) in the expression for \(\frac{\partial P(D_i)}{\partial d_i}\) given above declines.

Thus, there is at most a single value of \(d_1\) for which \(d_1=d_2\) and \(T(d_1) = 0\), which implies that there is at most a single value of \(d_1\) for which \(d_1=d_2\) and \(\frac{\partial P(D_i)}{\partial d_i} = 0\).

Using equation WS2, we note that \(T(m_D) > 0\) [because \(g'(m_D, m_D) = 0\) and \(g'(m_D, m_G) > 0\)] and \(T(m_G) < 0\) [because \(g'(m_G, m_D) < 0\) and \(g'(m_G, m_G) = 0\)]. Hence,

\[
\frac{\partial P(D_i)}{\partial d_1} \bigg|_{d_1=m_D} > 0 \quad \text{and} \quad \frac{\partial P(D_i)}{\partial d_1} \bigg|_{d_i=m_G} < 0. \quad \text{By the intermediate value theorem, there exists a point between } m_D \text{ and } m_G \text{ for which } \frac{\partial P(D_i)}{\partial d_1} = 0, \text{ and by previous argument, this point is unique. This completes the proof of part (b) of Theorem 1.}
Proof of Theorem 2

**Theorem 2.** Suppose there is one candidate $R_1$ and two Democratic candidates $D_1$ and $D_2$, whose equilibrium positions $d_1^*$ and $d_2^*$ lie strictly between $m_D$ and $m_G$. Then a unilateral increase (decrease) in $V_{R1}$ gives both Democratic candidates incentives to shift unilaterally toward (away from) $m_G$.

**Proof.** Suppose that the three candidates are at equilibrium positions $d_1^*$, $d_2^*$, and $r_1^*$ for pre-campaign valence values $V_{D_1}, V_{D_2}$, and $V_{R1}$. If the pre-campaign valence image for the Republican increases to $V_{R1} + \alpha$, where $\alpha > 0$, then it suffices to show that for this larger pre-campaign valence image, $\frac{\partial P(D_1)}{\partial d_1}$, evaluated at $d_1^*$, $d_2^*$, and $r_1^*$, is positive. But it follows from equation WS2 that $\frac{\partial P(D_1)}{\partial d_1}$ has the same sign as the numerator of the expression obtained by applying the quotient rule, i.e.,

$$
g'(d_1, m_D) \exp[M_D(D_2) - M_D(D_1)] + g'(d_1, m_G) \exp[M_G(R_1) + \alpha - M_G(D_1)]
$$

evaluated at $d_1^*$, $d_2^*$, and $r_1^*$. In turn, this expression is greater than the identical expression without $\alpha$, i.e.,

$$
g'(d_1, m_D) \exp[M_D(D_2) - M_D(D_1)] + g'(d_1, m_G) \exp[M_G(R_1) - M_G(D_1)]
$$

because $g'(d_1, m_G) < 0$, since $g$ is concave and peaks at $m_G$. But the latter expression is zero, because $d_1^*$, $d_2^*$, and $r_1^*$ constitute a Nash equilibrium for $V_{D_1}, V_{D_2}$, and $V_{R1}$. This completes the proof of Theorem 2.
Derivation of the probabilities of winning office in a model where both parties hold primaries with any number of candidates

The derivations apply to the general scenario where both parties hold primaries with any number of candidates. We assume that in each party primary the winning candidate is the one preferred by the median primary election voter. Let $D_j$, $j = 1, \ldots, n$ denote the candidates in the Democratic primary and $R_k$, $k = 1, \ldots, m$ be the candidates in the Republican primary. Then the probability that candidate $D_1$ wins both the primary and the general election is the sum over $k$, $k = 1, \ldots, m$, of the probabilities that candidate $D_1$ wins the Democratic primary and the general election with $R_k$ as the general election opponent. Under the assumption that in each party primary the winning candidate is the one preferred by the median primary election voter, for fixed values of the campaign valence images $\varepsilon_{D_1}$ and $\varepsilon_{R_1}$, the probability $P(D_1, R_1)$ that $D_1$ wins the primary and general election with $R_1$ as the general election opponent is given by

$$P\left[ M_D(D_1) + \varepsilon_{D_1} > M_D(D_j) + \varepsilon_{D_j}, j = 2, \ldots, n; M_R(R_1) + \varepsilon_{R_1} > M_R(R_k) + \varepsilon_{R_k}, k = 2, \ldots, m; \right. \left. and M_G(D_1) + \varepsilon_{D_1} > M_G(R_1) + \varepsilon_{R_1} \right]$$

2 In the two-Democrat, two-Republican case, the assumption that in each party primary the winning candidate is the one preferred by the median primary election voter is the natural Downsian assumption; if a primary involves more than two candidates, this assumption is more restrictive. However this assumption may be a reasonable approximation for multicandidate primaries under either of the following conditions: 1) that the distribution of the primary electorate’s policy preferences is extremely compact, relative to the distribution of policy preferences in the general electorate; 2) that voters vote strategically in multicandidate primaries in the sense that they support their preferred candidate from among the set of candidates who have a realistic chance of winning the primary (not to be confused with strategic voting with an eye towards general election competitiveness).
where, for convenience, we have written

\[ W_{D_j} = M_D(D_j) - M_D(D_j), j = 2, \ldots, n, \]
\[ W_{R_k} = M_R(R_k) - M_R(R_k), k = 2, \ldots, m, \]
\[ W_G = M_G(D_1) - M_G(R_1). \]  

Because \( \varepsilon_{D_j}, j = 2, \ldots, n \) and \( \varepsilon_{R_k}, k = 2, \ldots, m \) are independent, it follows – again for fixed \( \varepsilon_{D_1} \) and \( \varepsilon_{R_1} \) – that

\[
P(D_1R_1) = \begin{cases} 
\prod_{j=2}^{n} F_{D_j}(\varepsilon_{D_1} + W_{D_j}) \prod_{k=2}^{m} F_{R_k}(\varepsilon_{R_1} + W_{R_k}), & \text{if } \varepsilon_{R_1} < \varepsilon_{D_1} + W_G, \\
0, & \text{otherwise}
\end{cases}
\]

where the symbol \( \prod_{j=2}^{n} (\cdot) \) denotes a product as \( j \) varies from 2 to \( n \). Now taking account that \( \varepsilon_{D_1} \) and \( \varepsilon_{R_1} \) may take on any values, we have

\[
P(D_1R_1) = \int_{-\infty}^{\infty} \prod_{j=2}^{n} F_{D_j}(\varepsilon_{D_1} + W_{D_j}) f_{D_1}(\varepsilon_{D_1}) \int_{-\infty}^{\varepsilon_{D_1} + W_G} \prod_{k=2}^{m} F_{R_k}(\varepsilon_{R_1} + W_{R_k}) f_{R_1}(\varepsilon_{R_1}) d\varepsilon_{R_1} d\varepsilon_{D_1}.
\]

Writing \( s = \varepsilon_{D_1} \) and \( t = \varepsilon_{R_1} \), we have

\[
P(D_1R_1) = \int_{-\infty}^{\infty} \exp[-e^{-s} \sum_{j=2}^{n} e^{-W_{D_j}}] \exp[-e^{-s}] e^{-s} \int_{-\infty}^{\infty} \exp[-e^{-t} \sum_{k=2}^{m} e^{-W_{R_k}}] \exp[-e^{-t}] e^{-t} dt ds
\]

\[
= \int_{-\infty}^{\infty} \exp[-e^{-s} \sum_{j=2}^{n} e^{-W_{D_j}}] \exp[-e^{-s}] e^{-s} \int_{-\infty}^{\infty} \exp[-e^{-t} \left( \sum_{k=2}^{m} e^{-W_{R_k}} + 1 \right)] e^{-t} dt ds
\]

\[
= \int_{-\infty}^{\infty} \exp[-e^{-s} \sum_{j=2}^{n} e^{-W_{D_j}}] \exp[-e^{-s}] e^{-s} \int_{-\infty}^{\infty} \exp[-u \left( \sum_{k=2}^{m} e^{-W_{R_k}} + 1 \right)] du ds \quad [u = e^{-t}]
\]
Similar formulas can be obtained for the probability $P(D_l R_i)$ that $D_1$ wins the Democratic primary and the general election against any Republican candidate $R_i$. Finally, the probability that $D_1$ wins office is the sum of the probabilities $P(D_l R_i), l = 1, \ldots, m$. The corresponding probabilities for the other candidates are similar. This completes the derivation for the $n$-Democrat, $m$-Republican case.
Proofs of Theorems 3 and 4

Theorems 3 and 4 are proved for the multicandidate case, which includes the specialized scenario assumed in Theorems 3-4 in the paper “Candidate Strategies in Two-Stage Elections Beginning with a Primary,” in which both parties hold two-candidate primaries, as well as that of Theorems 1-2 in the same paper, in which one party holds a two-candidate primary while a single candidate runs unopposed in the other party’s primary.

Proof of Theorem 3

**Theorem 3.** If both parties hold primaries, if \( V_{D1} = V_{D2} = \ldots = V_{Dn} \) and \( V_{R1} = V_{R2} = \ldots = V_{Rm} \), and if in each party primary the winning candidate is the one preferred by the median primary election voter, then there exists an agglomerated equilibrium in candidates’ office-seeking strategies, i.e., one such that \( d_1^* = d_2^* = \ldots = d_n^* \) and \( r_1^* = r_2^* = \ldots = r_m^* \).

**Proof.** First, note that if an agglomerated equilibrium exists then it must be one for which the Democratic candidates locate in the policy interval \([m_D, m_G]\), while the Republicans locate in the policy interval \([m_G, m_R]\). We next show that, with the Republican candidates located at a common position in the interval \([m_G, m_R]\), there will be exactly one location \( d_1 = d_2 = \ldots = d_n \) in the interval \([m_D, m_G]\) that constitutes an equilibrium in the Democratic candidates’ strategies. If \( r_1 = r_2 = \ldots = r_m \) and \( V_{R1} = V_{R2} = \ldots = V_{Rm} \), it follows that \( M_R(R_1) = M_R(R_2) = \ldots = M_R(R_m) \). Thus, using equation WS5 and the analogous equations for \( P(D_1 R_k) \), \( k = 2, \ldots, m \), the probability that candidate \( D_1 \) wins office, given by \( P(D_1) = \sum_{k=1}^{m} P(D_1 R_k) \), simplifies to
\[ P(D_1) = \frac{1}{1 + \sum_{j=2}^{n} e^{-[M_{D}(D_1) - M_{D}(D_j)]} + me^{-[M_{G}(D_1) - M_{G}(R_i)]}}. \] (WS6)

The derivative of the above expression with respect to \( d_1 \) is

\[
\frac{\partial P(D_1)}{\partial d_1} = \frac{g'(d_1, m_D) \sum_{j=2}^{n} e^{[M_{D}(D_j) - M_{D}(D_1)]} + mg'(d_1, m_G) e^{[M_{G}(R_i) - M_{G}(D_1)]}}{\left( \sum_{j=1}^{n} e^{[M_{D}(D_j) - M_{D}(D_1)]} + me^{[M_{G}(R_i) - M_{G}(D_1)]} \right)^2}. \] (WS7)

From this point on, demonstrating that given \( r_1 = r_2 = \ldots = r_m \), there is exactly one value of \( d_1 \) in the policy interval \([m_D, m_G]\) for which \( d_1 = d_2 = \ldots = d_n, \frac{\partial P(D_1)}{\partial d_1} = 0 \), and \( \frac{\partial^2 P(D_1)}{\partial^2 d_1} \leq 0 \) follows the same logic that we used for the similar demonstration for Theorem 1. Also, as in that theorem, we can show that \( P(D_1) \) is single-peaked and peaks at the equilibrium point \( d_1 \). By a parallel set of arguments, when the candidates \( D_j \) are all at the same location in the policy interval \([m_D, m_G]\), then there must exist exactly one location for the Republican candidates \( r_1^* = r_2^* = \ldots = r_m^* \) in the interval \([m_G, m_R]\) that constitutes an equilibrium in the Republican candidates’ office-seeking strategies.

Next, define a function \( h_1 \) that assigns to each value \( r \) in the interval \([m_G, m_R]\) the value \( d \) in the interval \([m_D, m_G]\) that represents the equilibrium location for the common position of the Democratic candidates \( D_1, D_2, \ldots, D_n \) when the Republican candidates are all located at \( r \).

Denote by \( h_2 \) the corresponding function that assigns to each value \( d \) in the interval \([m_D, m_G]\) the

Note that, while we are shifting the Democratic candidates as a group, at the equilibrium we obtain, no Democratic candidate can improve his election probability by shifting unilaterally away from the common location.
value $r$ in the interval $[m_G, m_R]$ that represents the equilibrium for $R_1, \ldots, R_m$ when the Democratic candidates are all located at $d$. Since the function $h_1 \circ h_2$ maps the closed, bounded, and convex set $[m_G, m_R]$ into itself, if we can show that this function is continuous, then by the Brouwer Fixed Point Theorem (McCarty and Meriowitz, 2007: 124), $h_1 \circ h_2$ has a fixed point, i.e., there exists a point $r_0$ in $[m_G, m_R]$ such that $h_2(h_1(r_0)) = r_0$. It will then follow that setting $r_1 = r_2 = \ldots = r_m = r_0$ and $d_1 = d_2 = \ldots = d_n = h_1(r_0)$ constitutes a Nash equilibrium.

Because the composition of continuous functions is continuous, to show that $h_1 \circ h_2$ is continuous, it suffices to show that $h_1$ is continuous. To do this, define $p(d, r)$ as the probability that $D_1$ wins office given that the Republicans are all located at $r$ and the Democrats are all located at $d$. Fix a value of $r$, say $r_1$, and let the corresponding value at which the Democrats are all located be $d_1$. Specifically, we show that $h_1$ is continuous at $r_1$. Given $\varepsilon > 0$, because, as a function of $d$, $p(d, r_1)$ is single-peaked and peaks at $d_1$,

$$
\gamma = p(d_1, r_1) - \min[p(d_1, \varepsilon), p(d_1, \varepsilon)] > 0.
$$

Since also $p$ is continuous on the closed, bounded set $[m_D, m_G] \times [m_G, m_R]$, it is uniformly continuous. Hence there exists $\delta > 0$ such that if $|r - r_1| < \delta$, then $|p(d, r) - p(d, r_1)| < \gamma / 2$. Thus, for $|r - r_1| < \delta$, if $|d - d_1| \geq \varepsilon$, by the continuity of $p$ and the definition of $\gamma$,

$$
p(d, r) < p(d, r_1) + \gamma / 2 < p(d_1, r_1) + \gamma / 2 - \gamma = p(d_1, r_1) - \gamma / 2.
$$

But again by the continuity of $p$, $p(d_1, r) > p(d_1, r_1) - \gamma / 2$, so for any fixed $r$ with $|r - r_1| < \delta$, the maximum of $p(d, r)$ over $d$ must occur for $|d - d_1| < \varepsilon$, i.e., using our previous notation,

$$
|h_1(r) - h_1(r_1)| < \varepsilon.
$$

This completes the proof of Theorem 3.
Proof of Theorem 4

**Theorem 4.** Suppose that both parties hold primaries, that $V_{D1} = V_{D2} = ... = V_{Dn}$ and $V_{R1} = V_{R2} = ... = V_{Rm}$, that in each primary the winning candidate is the one preferred by the median primary election voter, and that the parties are located at agglomerated equilibrium configurations, with $d_1^* = d_2^* = ... = d_n^*$ located strictly between $m_D$ and $m_G$. Then a unilateral increase (decrease) in one or more of the $V_{Rk}$’s gives all Democratic candidates incentives to shift unilaterally toward (away from) $m_G$; a parallel statement holds for the Republicans.

**Proof.** Given that $V_{D1} = V_{D2} = ... = V_{Dn}$ and $V_{R1} = V_{R2} = ... = V_{Rm}$, the numerator of the derivative of the function $\frac{\partial P(D)}{\partial d_1}$ in the $n$-Democrat, $m$-Republican case (given by equation WS7) differs from that for the 2-Democrat, 1-Republican case (see the proof of Theorem 2) only by a factor of $m$ in one term. Accordingly, it is easy to see that the same argument used in the proof of Theorem 2 completes the proof of Theorem 4.
A Comparison of Candidate Strategies for
Expressive versus Strategic Primary Voting

James Adams and Samuel Merrill, III

(Working paper 10-07 – Please do not quote without authors’ permission. Note that the numbering of equations below is specific to this working paper.)

For these analyses of strategic primary voting we assume the same model of voters’ candidate evaluations that we use in our paper, “Candidate Strategies in Two-Stage Elections Beginning with a Primary,” and we add an additional assumption: namely that all primary voters know the location of the median general election voter $m_G$, and that the primary election voters know that all voters in the electorate have identical policy loss functions $g(j, x_i)$. For our model of strategic voting, we assume that primary voters weigh the candidates’ projected appeal in the general election. Specifically, we assume that strategic primary voters assume that the rival party will nominate its strongest general election candidate – i.e. the candidate whose combination of policy positions and campaigning abilities is most attractive to the median general election voter – and that primary voters project whether the focal candidate from their party is electable, in the sense that this focal candidate can defeat the rival party’s strongest general election candidate (we discuss the details of such projections momentarily). We assume that strategic primary voters employ the following decision rule, when voting in a competitive primary: If they project that both of their party’s primary candidates are electable, then they vote for the candidate they sincerely prefer; however if they project that only one or neither of the primary candidates is electable then they support the stronger general election candidate, regardless of their sincere preference. This contrasts with the expressive voting model we analyze in our paper, for which voters invariably support their sincerely-preferred candidate.

With respect to how strategic primary voters project whether a focal candidate is electable, note that at the time the primary electorate goes to the polls these voters have observed the candidates’ policy positions and their campaign-based valence images – and, furthermore, they have observed these attributes for both parties’ candidates. Therefore, given our specification that the candidates’ policy positions and campaigning abilities do not change between the primary and the general election, and the assumptions we employ here that primary election voters know the median general election voter’s position $m_G$ as well as voters’ policy loss function $g(j, x_i)$, it follows from Remark 2 in our manuscript (i.e. that the winning candidate in the general election is the one preferred by the median general election voter) that primary voters can assess which of the rival party’s primary candidates is the stronger general election opponent, and they can also forecast how each of

---

4 Alternatively, we may consider a “minimax” decision rule, whereby they assume that the rival party will make the best possible response to whichever candidate the focal party’s voters nominate; this would lead to the same outcome.
their own party’s primary candidates will fare in a general election contest against this opponent.\(^5\)

We emphasize that our specification that strategic primary voters can accurately forecast the outcome of any possible general election match-up is a very strong assumption. However if expressive and strategic primary voting motivate identical candidate policy strategies even in the extreme case where strategic voters can perfectly assess each candidate’s electability, then this conclusion plausibly extends to situations where primary voters are uncertain about the outcome of the general election.

**The probability that a candidate wins office when primary voters are strategic**

We now derive formulas for the probability that a candidate wins office for strategic primary voting. For these derivations we assume that the Democrats hold a competitive primary involving the candidates \(D_1\) and \(D_2\).

**Remark 1.** Under the strategic primary voting model described above, the probability that a party \(j\) wins office equals the probability \(P_G(j)\) that the median general election voter prefers one of party \(j\)'s candidates, i.e. the probability that this voter ranks one of \(j\)'s candidate’s first over all the primary candidates from both parties.\(^6\)

From Remark 1, it follows that the probability \(P_G(D)\) that the Democrats win office, i.e. the probability that either \(D_1\) or \(D_2\) is elected, is

\[
P_G(D) = \frac{e^{M_G(D_1)} + e^{M_G(D_2)}}{e^{M_G(D_1)} + e^{M_G(D_2)} + e^{M_G(R_1)}},
\]

for scenarios where the candidate \(R_1\) runs unopposed in the Republican primary, and that this probability is

\(^5\) Note that strategic voters’ abilities, at the time of the primary election, to forecast the general election outcome, contrasts with the electoral uncertainty candidates confront at the time they must select their policy strategies, which is prior to the primary campaign. This difference arises because the candidates have not observed their own (and the rival candidates’) campaigning skills at the time they select their positions, while strategic voters have observed these skills by the date of the primary election.

\(^6\) To see this, note that if the median general election voter ranks a candidate first over all the primary election candidates from both parties, then this candidate is electable in the sense that she can defeat the rival party’s strongest general election candidate. Therefore the probability that party \(j\) has an electable primary candidate equals \(P_G(j)\). And, given our definition of strategic primary voting, it follows that if a party has an electable primary candidate then the primary voters will select an electable candidate; furthermore, if the party does not have an electable primary candidate, then the party will not win office since in this case the rival party must have one or more electoral candidates.
\[ P_G(D) = \frac{e^{M_D(D_1)} + e^{M_D(D_2)}}{e^{M_D(D_1)} + e^{M_D(D_2)} + e^{M_D(R_1)} + e^{M_D(R_2)}}, \]  

when candidates \( R_1 \) and \( R_2 \) compete in the Republican primary.\(^7\)

Our second remark applies to candidates running in a competitive primary election:

**Remark 2.** If a focal candidate’s policy strategy is at least as extreme as her primary opponent’s strategy relative to the median general election voter, then this focal candidate’s probability of winning office is identical under the strategic and expressive primary voting scenarios.\(^8\)

Remarks 1 and 2 have the following implication. Define \( P_e(D_i) \) and \( P_s(D_i) \) as the respective probabilities that candidates \( D_1 \) and \( D_2 \) win office when primary voters are expressive, and define \( P_e(D_1) \) and \( P_e(D_2) \) as the respective probabilities that candidates \( D_1 \) and \( D_2 \) win office when primary voters are strategic. Then

\[
\begin{align*}
\text{if } d_2 &\leq d_1, \quad P_e(D_1) = P_e(D_1) \quad \text{and} \quad P_s(D_2) = [P_G(D) - P_e(D_1)] , \\
\text{and} \\
\text{if } d_1 &\leq d_2, \quad P_e(D_1) = [P_G(D) - P_e(D_2)] \quad \text{and} \quad P_s(D_2) = P_e(D_2) ,
\end{align*}
\]

where the formulas for the expressive vote probabilities \( P_e(D_i) \) and \( P_s(D_i) \) are given in our paper, and the formula for \( P_G(D) \) is given by equation A1 above.

---

\(^7\) The probability formula given by equation A1 represents the sum of the probability that the median general election voter prefers \( D_1 \), which is \( P_G(D_1) = \frac{e^{M_D(D_1)}}{e^{M_D(D_1)} + e^{M_D(D_2)} + e^{M_D(R_1)} + e^{M_D(R_2)}} \), and the probability that the median general election voter prefers \( D_2 \), which is \( P_G(D_2) = \frac{e^{M_D(D_2)}}{e^{M_D(D_1)} + e^{M_D(D_2)} + e^{M_D(R_1)} + e^{M_D(R_2)}} \). (Formulas are given for the scenario where \( R_1 \) and \( R_2 \) contest the Republican primary.)

\(^8\) To see this, suppose that candidate \( D_1 \)’s policy strategy \( d_1 \) is at least as extreme as \( D_2 \)’s strategy \( d_2 \), relative to the median general election voter’s position \( m_G \), i.e. suppose that \( d_1 \leq d_2 \leq m_G \). Then, given that voters have concave policy loss functions, it is impossible to construct a scenario where the median Democratic primary voter (who is located to the left of \( m_G \)) prefers \( D_2 \) to \( D_1 \), but the median general election voter strictly prefers \( D_1 \) to \( D_2 \). And, given that this is the only scenario for which \( D_1 \) can be elected under strategic primary voting but not under sincere primary voting, the fact that this scenario is impossible when \( d_1 \leq d_2 \leq m_G \) implies that \( D_1 \)’s probability of winning office is identical for expressive and strategic primary voting.
Candidate Policy Strategies and Office-Seeking Equilibrium for Expressive and Strategic Voting: Illustrative Examples

We now present an illustrative example – for the two-primary case – that conveys intuitions about office-seeking candidates’ policy strategies for expressive and strategic primary voting scenarios. For this example we assume the same scenario that we analyze in our manuscript, namely a voter distribution where the median general election voter is located at $m_o = 4$ on the Left-Right scale, the median Democratic primary voter locates at $m_D = 2$, the median Republican primary voter locates at $m_R = 6$, and where voters have quadratic policy loss functions and the policy salience coefficient is $a = 1$, i.e. the loss function is $g(j, x_j) = -(j - x_j)^2$. We also assume that all candidates have equal pre-campaign valence images, i.e. $V_{D1} = V_{D2} = V_{R1} = V_{R2}$.

Candidate strategies for expressive voting. We begin with the case where primary voters are expressive. Figure 1 illustrates the resulting Nash equilibrium configuration in candidates’ office-seeking strategies: the Democratic candidates pair at the position $d_1^* = d_2^* = 3.33$ and the Republicans pair at $r_1^* = r_2^* = 4.67$, i.e. each candidate’s optimal position is located between his party’s median primary voter and the median general election voter. The solid line shows candidate $D_1$’s probability $P(D_1)$ of winning office under expressive voting motivations as a function of his policy position, with the candidates $D_2$, $R_1$, and $R_2$ fixed at their equilibrium positions. (Below we discuss the significance of the hatched line in the figure, which represents $D_1$’s election probability for strategic primary voting.) Note that $P(D_1)$ drops sharply if $D_1$ shifts substantially to the left or to the right of his optimal position $d_1^* = 3.33$. In the former scenario $D_1$’s leftist positioning puts him in a strong position to win the Democratic primary but at the cost of crippling his general election prospects; in the latter scenario $D_1$’s centrist positioning makes him a strong general election candidate, but severely depresses the probability that his expressive primary constituency will nominate him.

An important feature of this equilibrium configuration is that the candidates’ optimal positions are located closer to the median general election voter position $m_o = 4$ than they are to the median positions of the candidates’ primary constituencies ($m_D = 2$ for the Democrats and $m_R = 6$ for the Republicans). Thus in this example the office-seeking candidates have centrifugal incentives to appeal to expressive primary voters, but this incentive is weaker than the centripetal pull of the general electorate. Why is this true? The central intuition revolves around the candidates’ expectations about the relative campaigning abilities of their primary opponent compared to their general election opponent. Specifically, each candidate expects to face a better campaigner in the general election than he will in the primary. From the perspective of candidate $D_1$, for instance, the knowledge that two candidates ($R_1$ and $R_2$) are contesting the Republican primary – and that the Republican who demonstrates superior campaign-based valence will be nominated – implies that the Republican nominee’s campaign-based valence has a higher expected
value than that of \( D_1 \)’s primary opponent \( D_2 \). Accordingly, \( D_1 \) shades his policy strategy towards the general electorate.

[FIGURE 1 ABOUT HERE]

**Candidate strategies for strategic primary voting.** We now consider the scenario where primary voters are strategic. Intuitively, we might expect that compared to the expressive voting scenario discussed above, strategic primary voting will motivate office-seeking candidates to moderate their policies, since the candidates realize that primary voters weigh the candidates’ electability. Indeed, in an important paper Serra (2006) demonstrates that in a primary election model without campaign effects – i.e. one where voters evaluate candidates entirely on the basis of policies – strategic primary voting motivates the candidates to converge to the exact center of the general electorate. However, when we incorporate campaign effects, we find that this intuition no longer applies: for the scenario we analyze here the Nash equilibrium for strategic primary voting is identical to that for expressive voting, i.e. for strategic voting the Democratic candidates again pair at \( d_1^* = d_2^* = 3.33 \) and the Republicans at \( r_1^* = r_2^* = 4.67 \). The hatched line in Figure 1 displays \( D_1 \)’s probability \( P_e(D_1) \) of winning office as a function of his position when primary voters are strategic, with the rival candidates again fixed at their equilibrium positions. We see that although \( D_1 \)’s election probability peaks at the identical location \( d_1^* = 3.33 \) for strategic and for expressive primary voting, the two functions are not identical: whereas \( D_1 \)’s probability \( P_e(D_1) \) of election under expressive primary voting drops sharply as he shifts rightward away from his optimal position \( d_1^* = 3.33 \), under strategic primary voting \( D_1 \)’s election probability \( P_s(D_1) \) drops only gradually as \( D_1 \) shifts to the right of \( d_1^* = 3.33 \), towards the center of the general electorate.

Why, in this example, do strategic and expressive primary voting support identical equilibria in candidates’ office-seeking strategies? A possible answer again revolves around the strategic importance of campaign effects. Given that \( D_1 \)’s probability \( P_e(D_1) \) of election under expressive primary voting drops sharply as \( D_1 \) moderates away from his optimal position \( d_1^* = 3.33 \), it is possible for the added centrifugal incentives of strategic voting to increase \( D_1 \)’s probability \( P_s(D_1) \) without increasing them enough to motivate office-seeking candidates to shift closer to the center, compared to their incentives for expressive primary voting. Simply put, in this example strategic primary voting renders a completely moderate candidate policy strategy less unattractive than such a strategy would be for expressive voting; nevertheless, in this scenario expressive and strategic voting support identical optimal strategies. We label this effect, which we support theoretically below, the *expressive and strategic voting equivalence result*.

---

9 Technically, the campaign-based valence image of the Republican who wins the primary is \( \max(\epsilon_{R1}, \epsilon_{R2}) \), the expected value of which is higher than that of \( \epsilon_{D1} \), the campaign valence of \( D_1 \), since \( \epsilon_{R1}, \epsilon_{R2}, \epsilon_{D1} \) all have the same distribution.
Office-Seeking Equilibrium for Expressive and Strategic Primary Voting: Theoretical Results

We now demonstrate that the strategic logic suggested by the above examples generalizes to a large class of two-stage elections. Theorem 1 shows that when voters have quadratic policy losses, then under specified conditions a unique candidate equilibrium is guaranteed to exist and will be identical for expressive and strategic voting. Note that the theorem applies only to the limiting case where 1) the value of the policy salience coefficient $a$ approaches zero, and, 2) where candidate movement in the policy space is not continuous. In addition, our theorem applies to the scenario where both parties hold competitive primaries (although it can easily be extended to the case where only one party holds a competitive primary). The proof of the theorem is presented at the end of this memo (see pages 8-13).

**Theorem 1.** Assume that each party has a two-candidate primary and that voters have quadratic policy loss functions. When the candidates select their positions from a finite set of platforms $Z$ that include the platforms $d_1(0)$, $d_2(0)$, $r_1(0)$, and $r_2(0)$ given below, then, when the value of the policy salience coefficient $a$ approaches zero, the unique equilibrium in the candidates’ office-seeking strategies for both expressive and strategic voting is \{\(d_1=d_1(0), d_2=d_2(0), r_1=r_1(0), r_2=r_2(0)\)\}, where

\[
d_1(0) = \frac{m_G + m_D w_{D1}}{1 + w_{D1}}, \quad \text{where} \quad w_{D1} = \frac{e^{v_{D2}}}{e^{v_{R1}} + e^{v_{R2}}},
\]

\[
d_2(0) = \frac{m_G + m_D w_{D2}}{1 + w_{D2}}, \quad \text{where} \quad w_{D2} = \frac{e^{v_{D1}}}{e^{v_{R1}} + e^{v_{R2}}},
\]

\[
r_1(0) = \frac{m_G + m_R w_{R1}}{1 + w_{R1}}, \quad \text{where} \quad w_{R1} = \frac{e^{v_{R2}}}{e^{v_{D1}} + e^{v_{D2}}},
\]

\[
r_2(0) = \frac{m_G + m_R w_{R2}}{1 + w_{R2}}, \quad \text{where} \quad w_{R2} = \frac{e^{v_{R1}}}{e^{v_{D1}} + e^{v_{D2}}}.
\]

(2)

In words, the theorem states that when the electoral salience of the candidates’ policy positions is sufficiently low, then there exists a unique paired equilibrium configuration, which is identical for expressive and strategic primary voting.
Candidates’ equilibrium positions for two-stage elections: Numerical results

Because Theorem 1 applies only in special cases, we turn to numerical calculation to demonstrate that equilibria are indeed identical for both expressive and strategic voting over a plausible range of conditions. Our calculations assume that each party has a two-candidate primary. Tables 1-2 report Nash equilibrium configurations in candidates’ office-seeking policy strategies for two-stage election contests, for expressive and for strategic primary voting. For these computations the median Democratic primary voter was again located at \( m_D = 2 \) and the median Republican primary voter \( m_R = 6 \). However we varied both the value \( a \) of the policy salience coefficient (between \( a=0.5 \) and \( a=1.0 \)) and the location \( m_G \) of the median general election voter (between \( m_G = 2 \) and \( m_G = 4 \)). \(^{10}\) We also varied voters’ policy loss metrics between quadratic losses (Table 1) and linear policy losses (Table 2). The top row of results presented in Table 1, for instance, reports the computed equilibrium configuration for the illustrative example pictured earlier in Figure 1, for which \( a=1 \), \( m_G = 4 \), and voters have quadratic policy losses. For this scenario, column 3 reports the equilibrium configuration for expressive primary voting \( \{ d_1^* = d_2^* = 3.33, r_1^* = r_2^* = 4.67 \} \), while column 4 reports that the equilibrium configuration for strategic primary voting is identical to that for expressive voting.

The computations reveal several patterns that are relevant to our theoretical results and illustrative examples. First, note that in every scenario the equilibrium configuration is identical for expressive and for strategic primary voting. These computations thereby suggest that the conclusion reported in Theorem 1 – which states that an identical paired equilibrium configuration exists for expressive and strategic primary voting, provided the policy salience coefficient \( a \) is sufficiently small – typically extends to scenarios where \( a \) is quite large. Second, note that in the scenarios with \( m_G = 4 \) – i.e. when \( m_G \) is equidistant from the two primary medians \( m_D = 2 \) and \( m_R = 6 \) – the candidates’ computed equilibrium positions are located closer to the median general election voter than to their median primary voter. Furthermore, note that this pattern also extends to the scenarios for which the median general election voter is not equidistant from the primary electorates (i.e. \( m_G \neq 4 \)). This suggests that the stronger centripetal pull of the general electorate is a general result.

Finally, note that when voters have linear policy losses (see Table 2), then, regardless of the location \( m_G \), the computed equilibrium configuration finds every candidate converging to \( m_G \) – i.e. for linear policy losses the candidates converge to the median general election voter position, just as they do in two-candidate general elections without a primary (Black, 1948). This demonstrates that – for the widely-used linear policy loss function – the primary electorate exerts no centrifugal pull on office-seeking candidates’ optimal strategies, compared to the centripetal pull of the primary electorate.

\[ \text{[TABLES 1-2 ABOUT HERE]} \]

\(^{10}\) We do not report equilibrium configurations for \( 4 < m_G \leq 6 \), because these configurations are the mirror images of the configurations we report for \( 2 \leq m_G < 4 \).
Proof of Theorem 1

Theorem 1. Assume that each party has a two-candidate primary and that voters have quadratic policy loss functions. When the candidates select their positions from a finite set of platforms \( Z \) that include the platforms \( d_1(0), d_2(0), r_1(0), \) and \( r_2(0) \) given below, then when the value of the policy salience coefficient \( a \) approaches zero, the unique equilibrium in the candidates’ office-seeking strategies for both expressive and strategic voting is \( \{d_1=d_1(0), d_2=d_2(0), r_1=r_1(0), r_2=r_2(0)\} \), where

\[
\begin{align*}
d_1(0) &= \frac{m_G + m_D w_{D1}}{1 + w_{D1}}, & w_{D1} &= \frac{e^{V_{D2}}}{e^{V_{D1}} + e^{V_{D2}}} \\
 d_2(0) &= \frac{m_G + m_D w_{D2}}{1 + w_{D2}}, & w_{D2} &= \frac{e^{V_{D1}}}{e^{V_{D1}} + e^{V_{D2}}} \\
r_1(0) &= \frac{m_G + m_R w_{R1}}{1 + w_{R1}}, & w_{R1} &= \frac{e^{V_{R2}}}{e^{V_{D1}} + e^{V_{D2}}} \\
r_2(0) &= \frac{m_G + m_R w_{R2}}{1 + w_{R2}}, & w_{R2} &= \frac{e^{V_{R1}}}{e^{V_{D1}} + e^{V_{D2}}}
\end{align*}
\]

Proof.

Part 1: Strategies when voters are expressive

Based on equation WS5 of the Web Supplement to the paper “Candidate Strategies in Two-Stage Elections Beginning with a Primary,” we know that under expressive voting:

\[
P_e(D_1) = P_e(D_1, R_1) + P_e(D_1, R_2)
\]

where

\[
P_e(D_1, R_1) = \frac{1}{1 + e^{[M_D(D_2) - M_D(D_1)]} + e^{[M_G(R_1) - M_G(D_1)]} + e^{[M_G(R_1) - M_G(D_2) + M_G(R_2) - M_G(R_1)]}} \times \frac{e^{M_G(R_1)}}{e^{M_G(R_1)} + e^{M_G(R_2)}}
\]

\[
P_e(D_1, R_2) = \frac{1}{1 + e^{[M_D(D_2) - M_D(D_1)]} + e^{[M_G(R_2) - M_G(D_1)]} + e^{[M_G(R_2) - M_G(D_2) + M_G(R_1) - M_G(R_2)]}} \times \frac{e^{M_G(R_1)}}{e^{M_G(R_1)} + e^{M_G(R_2)}}
\]

The derivative of \( P_e(D_1, R_i) \) with respect to \( d_1 \) is
\[
\frac{\partial P_e(D_i)}{\partial d_i} = \frac{\partial P_e(D_i, R_1)}{\partial d_i} + \frac{\partial P_e(D_i, R_2)}{\partial d_i}
\]

\[
= f'(d_1, m_D) e^{[M_D(D_1) - M_D(D_i)]} + f'(d_1, m_G) [e^{[M_G(R_1) - M_G(D_i)]} + e^{[M_G(R_2) - M_G(D_i) - M_R(R_1)]}] \\
\quad [1 + e^{[M_D(D_1) - M_D(D_i)]} + e^{[M_G(R_1) - M_G(D_i)]} + e^{[M_G(R_2) - M_G(D_i) - M_R(R_1)]}]^2 \times \frac{e^{M_R(R_1)}}{e^{M_R(R_1)} + e^{M_R(R_2)}}
\]

\[
+ f'(d_1, m_D) e^{[M_D(D_1) - M_D(D_i)]} + f'(d_1, m_G) [e^{[M_G(R_2) - M_G(D_i)]} + e^{[M_G(R_1) - M_G(D_i) - M_R(R_2)]}] \\
\quad [1 + e^{[M_D(D_1) - M_D(D_i)]} + e^{[M_G(R_2) - M_G(D_i)]} + e^{[M_G(R_1) - M_G(D_i) - M_R(R_2)]}]^2 \times \frac{e^{M_R(R_2)}}{e^{M_R(R_1)} + e^{M_R(R_2)}}
\]

where \( f \) is a single-peaked function that peaks at the voter’s ideal point and we write

\[
f'(d_1, m_D) = \frac{\partial f(d_1, m_D)}{\partial d_1}, \quad f'(d_1, m_G) = \frac{\partial f(d_1, m_G)}{\partial d_1}.
\]

Next we write \( g(d_1, m_D) = af(d_1, m_D), \quad g(d_1, m_G) = af(d_1, m_G) \), where \( a \) is a non-negative policy salience coefficient, that denotes the electoral salience of policies relative to the candidates’ campaign valence images.

When \( a=0 \), then \( M_D(D_1) = M_G(D_1) = V_{D_1}, \quad M_D(D_2) = M_G(D_2) = V_{D_2}, \quad M_R(R_1) = M_G(R_1) = V_{R_1}, \) and \( M_R(R_2) = M_G(R_2) = V_{R_2} \), and the equation for \( \frac{\partial P_e(D_i)}{\partial d_i} \) simplifies to

\[
\frac{\partial P_e(D_i)}{\partial d_1} = \frac{af'(d_1, m_D) e^{[v_{D_1} - v_{D_1}]} + af'(d_1, m_G) [e^{[v_{R_1} - v_{R_1}]} + e^{[v_{R_2} - v_{R_1}]}]}{[1 + e^{[v_{D_1} - v_{D_1}]} + e^{[v_{R_1} - v_{R_1}]} + e^{[v_{R_2} - v_{R_1}]}]^2} \times \frac{e^{v_{R_1}}}{e^{v_{R_1}} + e^{v_{R_2}}}
\]

\[
+ af'(d_1, m_D) e^{[v_{D_1} - v_{D_1}]} + af'(d_1, m_G) [e^{[v_{R_1} - v_{R_1}]} + e^{[v_{R_2} - v_{R_1}]}] \times \frac{e^{v_{R_2}}}{e^{v_{R_1}} + e^{v_{R_2}}}
\]

Since the denominators in both expressions on the RHS of the above equation are identical – i.e. the denominators are identical for \( \frac{\partial P_e(D_i, R_1)}{\partial d_i} \) and \( \frac{\partial P_e(D_i, R_2)}{\partial d_i} \) – it follows that the derivative of

\[
\frac{\partial P_e(D_i)}{\partial d_i}
\]

with respect to \( a \),

\[
\frac{\partial [\frac{\partial P_e(D_i)}{\partial d_i}]}{\partial a}
\]

, evaluated at \( a=0 \), increases most rapidly when the numerators for \( \frac{\partial P_e(D_i, R_1)}{\partial d_i} \) and \( \frac{\partial P_e(D_i, R_2)}{\partial d_i} \) in the above equation sum to zero, i.e. when
\[
\left[ f'(d_1, m_D) e^{[v_{d2} - v_{d1}]} + f'(d_1, m_G) [e^{[v_{g1} - v_{d1}]} + e^{[v_{g2} - v_{d1}]}] \right] e^{v_{g1}}
\]
\[
= -\left[ f'(d_1, m_D) e^{[v_{d2} - v_{d1}]} + f'(d_1, m_G) [e^{[v_{g1} - v_{d1}]} + e^{[v_{g2} - v_{d1}]}] \right] e^{v_{g2}}
\]

Algebraic calculation shows that for \(a=0\),
\[
\frac{\partial}{\partial d_1} \left[ \frac{\partial P_e(D_1)}{\partial a} \right] = 0 \iff -f'(d_1, m_D) e^{v_{d2}} + f'(d_1, m_G) [e^{v_{g1}} + e^{v_{g2}}] = 0
\]
which can be simplified to
\[
\frac{\partial}{\partial d_1} \left[ \frac{\partial P_e(D_1)}{\partial a} \right] = 0 \quad \text{for} \quad a = 0 \iff f'(d_1, m_G) [e^{v_{g1}} + e^{v_{g2}}] \Rightarrow \frac{f'(d_1, m_G)}{1} = \frac{e^{v_{d2}}}{e^{v_{g1}} + e^{v_{g2}}}
\]
i.e. that
\[
\frac{\partial}{\partial d_1} \left[ \frac{\partial P_e(D_1)}{\partial a} \right] = 0 \quad \text{for} \quad a = 0 \iff d_1(0) = \frac{m_G + m_D w_{D1}}{1 + w_{D1}}, \quad \text{where} \quad w_{D1} = \frac{e^{v_{d2}}}{e^{v_{g1}} + e^{v_{g2}}}
\]

When voters have quadratic policy losses, then when \(m_D \leq d_1 \leq m_G\) it follows that
\(f'(d_1, m_G) = 2(m_G - d_1)\) and \(f'(d_1, m_D) = 2(m_D - d_1)\), and in this case it is easily verified that
\[
\frac{\partial}{\partial d_1} \left[ \frac{\partial P_e(D_1)}{\partial a} \right] = 0 \quad \text{for} \quad a = 0 \iff d_1(0) = \frac{m_G + m_D w_{D1}}{1 + w_{D1}}, \quad \text{where} \quad w_{D1} = \frac{e^{v_{d2}}}{e^{v_{g1}} + e^{v_{g2}}}
\]

So far we have established that for \(a=0\), \(\frac{\partial P_e(D_1)}{\partial a}\) increases most rapidly with respect to increases in \(a\) when \(d_1 = d_1(0)\), where \(d_1(0)\) is given by equation 4. This in turn implies that for any two policy platforms \(d_1(0)\) and \(d_1'\) (where \(d_1(0) \neq d_1'\)), there is some positive value of the policy-salience coefficient \(a(d_1')\) – which depends on the platform \(d_1'\) – such that
\(P_e(D_1, d_1(0)) > P_e(D_1, d_1')\) for any \(a, \ 0 < a < a(d_1')\). This inequality is true by continuity of the
Partial derivative \( \frac{\partial}{\partial d_i} \). Now, with candidate \( D_1 \) fixed at \( d_1(0) \) let \( a_m(d_i') \) be the smallest value of \( a(d_i') \) for \( d_i' \in Z \) (with \( d_i(0) \neq d_i' \)). Thus, for \( 0 < a < a_m(d_i') \), \( P_s(D_1, d_1(0)) > P_s(D_1, d_1') \) for \( d_i' \in Z \) (with \( d_i(0) \neq d_i' \)). Note that for \( a=0 \), \( P_s(D_1) \) does not depend on the candidate strategy \( d_1 \). It follows that for \( 0 < a < a_m(d_i') \), candidate \( D_1 \)'s optimal position is \( d_1(0) \), no matter which platform the rival candidates select. We can use the identical approach to specify the value \( a_m(d_2') \), such that when \( 0 < a < a_m(d_2') \), candidate \( D_2 \)'s optimal position is \( d_2(0) \) no matter which platforms the rival candidates select, and similarly for candidates \( R_1 \) and \( R_2 \). It then follows that for \( 0 < a < \min\{a_m(d_1'), a_m(d_2'), a_m(r_1'), a_m(r_2')\} \), the equilibrium configuration in candidates’ optimal strategies is \( \{d_1=d_1(0), d_2=d_2(0), r_1=r_1(0), r_2=r_2(0)\} \). This completes the description of strategies for the expressive-vote case.

**Part 2: Strategies when voters are strategic**

Now we address the case where voters are strategic. Rather than go through all the steps of the proof we simply show that when the policy salience coefficient \( a=0 \), the position \( d_1 = d_1(0) \) for which \( \frac{\partial P_s(D_1)}{\partial d_1} = 0 \) is identical for strategic and expressive voting, i.e. that this derivative is once again given by equation 3. Recall that for strategic voting, the probability \( P_s(D_1) \) that candidate \( D_1 \) is elected takes a different functional form depending on whether \( D_1 \) is the more moderate or the more radical Democratic candidate, relative to the median general election voter position \( m_G \).

**Case A: The policy strategy for the more moderate Democratic candidate**

When voters are strategic and candidate \( D_1 \) is at least as moderate as \( D_2 \) relative to \( m_G \), i.e. when \( d_2 \leq d_1 \leq m_G \), then we know that

\[
P_s(D_1) = P_s(D) - P_s(D_2)
\]

where
\[ P_s(D) = \frac{e^{M_D(D_1)} + e^{M_D(D_2)}}{e^{M_D(D_1)} + e^{M_D(D_2)} + e^{M_G(R_1)} + e^{M_G(R_2)}} \]

\[ P_s(D_2) = \frac{1}{1 + e^{[M_D(D_1)-M_D(D_2)]} + e^{[M_G(R_1)-M_G(D_2)]} + e^{[M_G(R_2)-M_G(D_2)]}} \]

\[ = \frac{e^{M_D(D_2)}}{e^{M_D(D_1)} + e^{M_D(D_2)} + e^{[M_D(D_2)+M_G(R_1)-M_G(D_2)]} + e^{[M_G(D_2)+M_G(R_2)-M_G(D_2)]}} \]

where in the second equality for \( P_s(D_2) \) above, we have multiplied both the numerator and the denominator by \( e^{M_D(D_2)} \). This is done because it will help in the derivations below. Next take the derivative of \( P_s(D_1) \) with respect to \( d_1 \):

\[ \frac{\partial P_s(D_1)}{\partial d_1} = \frac{\partial [P_s(D_1) + P_s(D_2)]}{\partial d_1} - \frac{\partial P_s(D_2)}{\partial d_1} \]

\[ = \frac{af'(d_1, m_D)e^{M_D(D_1)}[e^{M_G(R_1)} + e^{M_G(R_2)}]}{[e^{M_D(D_1)} + e^{M_G(R_1)} + e^{M_G(R_2)}]^2} - \frac{-af'(d_1, m_D)e^{[M_D(D_1)+M_D(D_2)]}}{[e^{M_D(D_1)} + e^{M_G(D_2)} + e^{[M_D(D_2)+M_G(R_1)-M_G(D_2)]} + e^{[M_G(D_2)+M_G(R_2)-M_G(D_2)]}]^2} \]

When \( a=0 \), then \( M_D(D_1)=M_G(D_1)=V_{D1} \) and \( M_D(D_2)=M_G(D_2)=V_{D2} \), and mathematical computation shows that for \( a=0 \), the denominators of the above expressions for \( \frac{\partial [P_s(D_1)]}{\partial d_1} \) and \( \frac{\partial P_s(D_2)}{\partial d_1} \) on the RHS of the equation are equal. This in turn implies that for \( a=0 \):

\[ \frac{\partial P_s(D_1)}{\partial d_1} = 0 \Leftrightarrow f'(d_1, m_G)e^{V_{D1}}[e^{V_{R1}} + e^{V_{R2}}] = -f'(d_1, m_G)e^{[V_{D1}+V_{D2}]} \]

which in turn implies that

\[ \frac{\partial P_s(D_1)}{\partial a} = 0 \Leftrightarrow \frac{f'(d_1, m_G)}{-f'(d_1, m_D)} = e^{V_{D2}} e^{V_{R1}} + e^{V_{R2}} \].
The condition in equation 5 is identical to the condition given in equation 3, for expressive voting.

**Case B: The policy strategy for the more extreme Democratic candidate**

Finally, when voters are strategic and candidate $D_1$ is at least as extreme as $D_2$, i.e. when $d_1 \leq d_2 \leq m_G$, then

$$P_s(D_1) = \frac{1}{1 + e^{[M_G(D_2) - M_G(D_1)]} + e^{[M_G(R_1) - M_G(D_1)]} + e^{[M_G(R_2) - M_G(D_1)]}}.$$  

By straightforward calculation, we find that:

$$\frac{\partial P_s(D_1)}{\partial d_1} = 0 \text{ for } a = 0 \iff \frac{f'(d_1, m_G)}{-f'(d_1, m_D)} = \frac{e^{V_{D_2}}}{e^{V_{R_1}} + e^{V_{R_2}}},$$

which is again identical to the condition given in equation 3, for expressive voting.
Table 1. Equilibrium Configurations on Two-Stage Elections Beginning with a Primary: Results When Voters Have Quadratic Policy Losses

<table>
<thead>
<tr>
<th>Policy salience coefficient (1)</th>
<th>Median general election voter position (2)</th>
<th>Equilibrium configuration for expressive primary voting (3)</th>
<th>Equilibrium configuration for strategic primary voting (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a=1.0$</td>
<td>$m_G=4.0$</td>
<td>$d_1^<em>=d_2^</em>=3.33, r_1^<em>=r_2^</em>=4.67$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=1.0$</td>
<td>$m_G=3.5$</td>
<td>$d_1^<em>=d_2^</em>=2.94, r_1^<em>=r_2^</em>=4.22$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=1.0$</td>
<td>$m_G=3.0$</td>
<td>$d_1^<em>=d_2^</em>=2.57, r_1^<em>=r_2^</em>=3.77$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=1.0$</td>
<td>$m_G=2.5$</td>
<td>$d_1^<em>=d_2^</em>=2.27, r_1^<em>=r_2^</em>=3.29$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=1.0$</td>
<td>$m_G=2.0$</td>
<td>$d_1^<em>=d_2^</em>=2.00, r_1^<em>=r_2^</em>=2.84$</td>
<td>Same as expressive equilibrium</td>
</tr>
</tbody>
</table>

| $a=0.5$                         | $m_G=4.0$                                 | $d_1^*=d_2^*=3.33, r_1^*=r_2^*=4.67$                     | Same as expressive equilibrium                           |
| $a=0.5$                         | $m_G=3.5$                                 | $d_1^*=d_2^*=2.95, r_1^*=r_2^*=4.26$                     | Same as expressive equilibrium                           |
| $a=0.5$                         | $m_G=3.0$                                 | $d_1^*=d_2^*=2.60, r_1^*=r_2^*=3.83$                     | Same as expressive equilibrium                           |
| $a=0.5$                         | $m_G=2.5$                                 | $d_1^*=d_2^*=2.29, r_1^*=r_2^*=3.39$                     | Same as expressive equilibrium                           |
| $a=0.5$                         | $m_G=2.0$                                 | $d_1^*=d_2^*=2.00, r_1^*=r_2^*=2.95$                     | Same as expressive equilibrium                           |

Notes. For these computations the median position of the Democrats’ primary voter was set to $m_D=2$, and the median Republican primary voter was located at $m_R=6$. 


Table 2. Equilibrium Configurations on Two-Stage Elections Beginning with a Primary: Results When Voters Have Linear Policy Losses

<table>
<thead>
<tr>
<th>Policy salience coefficient (1)</th>
<th>Median general election voter position (2)</th>
<th>Equilibrium configuration for expressive primary voting (3)</th>
<th>Equilibrium configuration for strategic primary voting (4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a=1.0$</td>
<td>$m_G=4.0$</td>
<td>$d_1^<em>=d_2^</em>=4.00, r_1^<em>=r_2^</em>=4.00$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=1.0$</td>
<td>$m_G=3.5$</td>
<td>$d_1^<em>=d_2^</em>=3.50, r_1^<em>=r_2^</em>=3.50$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=1.0$</td>
<td>$m_G=3.0$</td>
<td>$d_1^<em>=d_2^</em>=3.00, r_1^<em>=r_2^</em>=3.00$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=1.0$</td>
<td>$m_G=2.5$</td>
<td>$d_1^<em>=d_2^</em>=2.50, r_1^<em>=r_2^</em>=2.50$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=1.0$</td>
<td>$m_G=2.0$</td>
<td>$d_1^<em>=d_2^</em>=2.00, r_1^<em>=r_2^</em>=2.00$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=0.5$</td>
<td>$m_G=4.0$</td>
<td>$d_1^<em>=d_2^</em>=4.00, r_1^<em>=r_2^</em>=4.00$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=0.5$</td>
<td>$m_G=3.5$</td>
<td>$d_1^<em>=d_2^</em>=3.50, r_1^<em>=r_2^</em>=3.50$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=0.5$</td>
<td>$m_G=3.0$</td>
<td>$d_1^<em>=d_2^</em>=3.00, r_1^<em>=r_2^</em>=3.00$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=0.5$</td>
<td>$m_G=2.5$</td>
<td>$d_1^<em>=d_2^</em>=2.50, r_1^<em>=r_2^</em>=2.50$</td>
<td>Same as expressive equilibrium</td>
</tr>
<tr>
<td>$a=0.5$</td>
<td>$m_G=2.0$</td>
<td>$d_1^<em>=d_2^</em>=2.00, r_1^<em>=r_2^</em>=2.00$</td>
<td>Same as expressive equilibrium</td>
</tr>
</tbody>
</table>

Notes. For these computations the median position of the Democrats’ primary voter was set to $m_D=2$, and the median Republican primary voter was located at $m_R=6$. 

27
Figure 1: Candidate $D_1$’s probability of being elected for expressive and for strategic voting, as a function of his policy position

Notes. For these computations the policy salience coefficient was set to $a=1$ for quadratic policy losses, and the median voter positions were set to $m_D=2$, $m_G=4$, $m_R=6$. This scenario supports identical equilibrium configurations for expressive and strategic voting, for which the Democratic candidates pair at $d_1^*=d_2^*=3.33$ and the Republicans pair at $r_1^*=r_2^*=4.67$. The figure displays candidate $D_1$’s probability of being elected as a function of his policy position, with the other candidates fixed at their equilibrium positions. $P_s(D_1)$ represents $D_1$’s election probability when primary voters are strategic, and $P_e(D_1)$ represents the election probability for expressive voting.