Strategic Party Heterogeneity

Georgia Kernell
Northwestern University
(forthcoming at *Journal of Theoretical Politics*)

Abstract

Political parties field heterogeneous candidates and send a variety of messages about their policy positions. Yet most voting models maintain that office-seeking parties should enforce intraparty homogeneity and cultivate clear party reputations. This article reconciles theory with reality by identifying a strategic rationale for parties to pursue heterogeneity. I develop a model in which two parties each select a distribution of potential candidates to compete in an upcoming election. The model demonstrates that well-positioned parties should indeed offer homogeneous candidate teams, but that parties with platforms distant from the median voter should cast a wide net. Extensions allow for multiple candidate signals, voters who care about party platforms and candidates’ positions, and voter uncertainty.

I thank John Bullock, Daniel Galvin, Samuel Kernell, PJ Lamberson, Jennifer Larson, Sean Theriault, Alvin Tillery, Michael Ting, the anonymous reviewers, and seminar participants at Northwestern, Penn State, and Yale for helpful suggestions.
“Our standard should not be universal purity; it should be a more welcoming conservatism.”
–Republican National Committee “Growth and Opportunity” report to members, March 2013.

On the heels of defeat, the Republican National Committee formed a task force in December of 2012 assigned to “provide an honest review of the 2012 election cycle and a path forward for the Republican Party to ensure success in winning more elections.” The ensuing report highlighted the GOP’s inability to recruit diverse candidates and appeal to a wide range of voters. Its prescription? Drop the call for “universal purity,” and create an “Inclusion Council” designed to “identify, prepare and promote a diversified and talented pool of future candidates and leaders within the Party.”

Democrats adopted a similar strategy after suffering devastating losses in 2004. The following year, DNC Chair Howard Dean launched a “Fifty State Strategy” geared toward building a ground game and winning elections throughout the country, including in districts where the Democratic candidate appeared too progressive to win or where Democrats had achieved few victories in the past. The strategy was a success; in 2006, Democrats took back the House and Senate.

Political losers certainly must change course, and it is not surprising that their post-election reflections and strategies call for change. But existing models of electoral politics cannot explain when or why parties seek to diversify their message or recruit a more heterogeneous set of candidates, as in the examples above. Previous research argues that office-seeking parties should invariably cultivate reputations through consistent behaviors and disciplined candidates. As Aldrich and Griffin state, “both citizens and politicians benefit from the construction of and adherence to clear party labels in a self-enforcing equilibrium” (2010, 598). From this perspective, parties are only heterogeneous because they lack control over candidate selection or find it too costly to screen potential nominees; all else equal, they should act as a unified team to attract the support of voters averse to risk (see, e.g., Ashworth and Bueno de Mesquita 2008, Groeling 2010, Grynaviski 2010, Snyder and Ting 2002, Woon and Pope 2008).

Yet, this prescription too often differs from reality, and below I find that it relies heavily on the assumption that two parties’ positions are equal – or at least equally appealing to voters. By ignoring the possibility that parties are unevenly matched, we are left with unsatisfying explanations for variation in heterogeneity, both between the two parties and within either party over time. Moreover, it is not clear that voters really do prefer candidates or parties with clear signals (Tomz and Van Houweling 2009).

This article departs from previous research by offering a strategic, electoral rationale for party heterogeneity. I present a model in which two parties choose a distribution of potential candidates who may compete in an upcoming election. A party can guard against extremists by selecting a more homogeneous pool, or it can act as a broad coalition by allowing candidates with varying viewpoints to seek the nomination. I find that for any pair of party platforms, there exists a unique Nash equilibrium pair of heterogeneity levels. When platforms perfectly match the position of the median voter, parties should adopt perfect homogeneity. In contrast, when parties are constrained from converging to the median voter (as they most often are), they should field a more heterogeneous set of candidates.1 The intuition
is straightforward: by recruiting a diverse team, out-of-step parties can increase their appeal to moderate voters.

The model gives rise to a number of additional expectations. First, there exists a unique level of heterogeneity that maximizes a party’s electoral prospects. Second, in a Nash equilibrium, the party closest to the median voter should adopt a narrower distribution than its more distant opponent. Third, parties will become increasingly heterogeneous with polarization, but more homogeneous with heightened competition. Extensions to the model validate its applicability to various electoral contexts, while providing a more nuanced understanding of strategic heterogeneity. As voters care more about a party’s platform or reputation relative to the position of the candidate in their district, parties will increase optimal heterogeneity. Adding voter uncertainty to the model also has an impact. If uncertainty is unrelated to heterogeneity, those parties that are farther from the median (and adopt higher heterogeneity in equilibrium) benefit most, and equilibrium levels of heterogeneity remain unchanged. If, however, uncertainty is a function of heterogeneity, the farther, more heterogeneous party’s chance of victory diminishes, and optimal heterogeneity should decline as well. In all of these settings, the model’s primary implication – that parties closer to the median voter should be more homogeneous than their more distant competitors – is upheld.

The following section develops the model. I then derive the primary results of the paper: optimal platform location (given heterogeneity), optimal heterogeneity (given platform location), and Nash equilibria. I conclude by considering the model’s implications for legislative behavior and party governance, as well as discussing avenues for future research.

A Spatial Model of Party Heterogeneity

Previous studies have investigated the influence of primary elections, party activism, campaign donors, interest groups, strong reputations, and institutional stability on party positioning and candidate responsiveness (see, e.g., Adams and Merrill 2008, Aldrich 1983, Aldrich and McGinnis 1989, Aranson and Ordeshook 1972, Bawn et al. 2012, Burden 2004, Grofman 2004). While many of these studies acknowledge intraparty heterogeneity by incorporating a pre-electoral bargaining stage over a party’s platform or public position, once a bargain is reached, parties are assumed to put aside their differences, pursue candidates in line with their position, and present a unified front. In contrast, the contribution of the current paper is to ask how party strategies differ when policy positions are described by a distribution rather than a single point. Said differently, this paper asks how leaders may use tools beyond platform positioning, such as widening or narrowing the field of potential candidates, to increase their party’s chance of victory.

The model begins with two office-seeking parties, $i$ and $j$. Each party $p \in \{i, j\}$ has a policy platform $\mu_p$ that lies in a single dimension. Platform selection is not modeled explicitly. Instead, the model is solved generally to serve as a secondary stage to most existing models of party positioning, including those where the two parties converge, diverge, or are asymmetrically positioned around the median voter. Each party has a pool of potential candidates whose positions are represented by a random variable $c_p$ with distribution $f_p$, such that $E[c_p] = \mu_p$. The heterogeneity
of the candidate pool is captured by the standard deviation of $f_p$ and is denoted $\sigma_p$. Parties that exclude candidates who are positioned far from the platform impose greater homogeneity, or “screening,” by setting $\sigma_p$ close to zero. Parties that allow for potential candidates with diverse positions create a “big tent” by choosing a high level of heterogeneity.

Once the party distributions are specified, a single candidate (whose position is also denoted $c_p$) is nominated by each party to run in the election. This selection process is modeled as a random draw from a party’s distribution of potential candidates. The candidates truthfully announce the policy positions they will pursue if elected, and voters observe these positions with certainty. (Later, I examine the case where heterogeneity increases uncertainty.) Candidate behavior is not modeled directly, and it is assumed that a sufficient supply of candidates allows parties to recruit individuals from any position along the policy continuum. Voters’ utility functions are given by $u(v, c_p)$, where $v$ is the voter’s ideal point, and $u$ is any function decreasing in $|v - c_p|$, such as a quadratic or linear loss function. (No assumption is made about voters’ attitudes toward risk.) An election is held, and the candidate receiving a majority of votes wins. In the event of a tie, each party wins with probability 0.5.

As described above, this model may serve as a secondary stage to many previous models of party positioning. To take one example, consider the case where party supporters – partisans, activists, or registered voters – exert strong control over party locations and candidate positions through primary elections. Let party positions (here, platforms) equal the ideal point of the median party supporter (as in, for e.g., McGann 2002). Hindered by election rules that put selection in the power of voters, leaders are left to rely on their ability to widen or narrow recruitment. Leaders choose a level of heterogeneity that describes the distribution from which two potential candidates emerge. A primary election occurs, and the candidate closest to the median supporter wins and advances to the general election. This scenario produces identical comparative statics to the one described above, where a single candidate is chosen from each party’s distribution.

While the model describes a context in which party leaders strategically select the heterogeneity of their candidates, the set-up applies equally well to a case in which parties choose the degree of heterogeneity in the signals or messages they present to the public about their policy positions and platform location. Parties wishing to cultivate a clear brand name will limit communication so that voters only observe statements in line with their platform. In contrast, parties that prefer to appear vague or variable will choose a higher level of heterogeneity and issue a variety of, potentially conflicting, statements about their preferred policy outcomes. The following discussion focuses on the case where parties choose heterogeneous candidate pools, rather than signals, but the latter may be more relevant in situations where voters depend on party reputations more than candidate positions when deciding how to vote (for example, in list systems or when candidates’ positions are less credible). This is explored in the first extension to the model, where voters observe multiple candidates (signals) from each party.
Election Outcomes

In two-party competition with full information and sincere voting, the candidate who captures the median voter’s support wins the election. Hence, the discussion can be simplified by stipulating that parties compete for the support of a single voter who represents the median in the electorate, and by setting this voter’s ideal point to zero. Given the distribution of potential candidates’ positions, \( f_i \) and \( f_j \), party \( i \)'s probability of victory is

\[
P_i = \int_{-\infty}^{\infty} f_j(c_j) \int_{-|c_j|}^{|c_j|} f_i(c_i) \, dc_i \, dc_j.
\]

Because parties are office-seeking, they set heterogeneity to whatever level maximizes their chance of winning the election. To solve for the optimal strategies, I assume that each party’s distribution is symmetric and single-peaked. I make the additional technical assumption that \( \frac{f_{p(a/\sigma)}}{f_{p(b/\sigma)}} \) is monotonic in \( \sigma > 0 \) for all \( a \neq b \) and \( p \in \{i, j\} \). These assumptions are fairly unrestrictive: candidate preferences may follow most common symmetric distributions, such as the normal or logistic distribution. I also assume that \( f_p \) can be fully described by its mean and variance, allowing each party’s distribution to be written in terms of a single standard distribution, \( \phi \), with mean zero and standard deviation one: \( f_p(c_p) = \frac{1}{\sigma_p} \phi \left( \frac{c_p - \mu_p}{\sigma_p} \right) \). When \( \sigma_p = 0 \), let \( f_p \) equal a delta distribution at \( \mu_p \): \( \delta_{\mu_p}(c_p) = \lim_{\sigma_p \to 0} \frac{1}{\sigma_p} \phi \left( \frac{c_p - \mu_p}{\sigma_p} \right) \).

Substituting this notation into equation (1) gives:

\[
P_i = \int_{-\infty}^{\infty} \frac{1}{\sigma_j} \phi \left( \frac{c_j - \mu_j}{\sigma_j} \right) \int_{-|c_j|}^{|c_j|} \frac{1}{\sigma_i} \phi \left( \frac{c_i - \mu_i}{\sigma_i} \right) \, dc_i \, dc_j.
\]

Comparative Statics and Equilibria

Before examining party heterogeneity, it is useful to confirm that the equilibrium prescription for unitary parties – converge to the median voter – also holds when parties are heterogeneous.

Shifts in a Party’s Platform

Proposition 1. A party’s probability of winning increases as its platform moves closer to the median voter.

(Formal Statement: For all \( i \neq j \), decreasing \( |\mu_i| \) increases \( P_i \) and decreases \( P_j \).)

In the standard spatial model, where parties are represented by points instead of distributions, a change in position only affects election outcomes if it alters which party is closest to the median voter. But when candidate distributions are probabilistic, any platform shift toward the median increases a party’s probability of winning. Proposition 1 confirms our intuition that a party described by a symmetric, single-peaked distribution maximizes its probability of winning the election by setting its platform at the position of the median voter, regardless of its level of heterogeneity. (The proofs for all propositions are in Appendix A.)
Shifts in platform location will help or hurt a party most when that party is neither very close to the median voter (and therefore likely to win the election) nor very far from the median voter (and highly unlikely to win). Parties located somewhere in between face the stiffest competition and have the most to gain or lose by adjusting their stance. Similarly, a change in platform matters less as heterogeneity increases. Thus, if platform movement is costly, homogeneous parties that are neither very far nor very close to the median voter should be most likely to adjust their position. (For a detailed examination of the marginal effects of a change in platform location on the probability of winning see Appendix B.)

**Party Heterogeneity**

Proposition 1 confirms our intuition that vote-maximizing parties should still converge to the position of the median voter. But what happens when platforms are off median? In reality, party platforms and candidate positions rarely converge to the position of the median voter (Ansolabehere, Snyder and Stewart 2001, Burden 2004, Frendreis et al. 2003, Jessee 2010). Parties adopt more extreme positions to woo party activists (Aldrich 1983, Aldrich and McGinnis 1989, Frendreis et al. 2003) and primary voters (Burden 2004, Owen and Grofman 2006), or to deter abstention (Downs 1957) or the entry of an independent or third-party candidate (Lee 2012). Candidates may also have policy preferences of their own that are more extreme than those of the electorate (Aldrich 2011, Calvert 1985, Wittman 1983). And even those parties that successfully locate at the position of the median voter may find themselves unable to quickly adjust their platform when exogenous shocks (such as terrorist attacks or redistricting) dramatically change the partisan landscape and shift the position of the median voter (Burden 2004). Thus, it is important to ask if increasing heterogeneity can offer parties an attractive alternative to adjusting their policy platform. Should parties that are disadvantaged by unpopular platforms choose to be strategically heterogeneous? And if so, what distribution will maximize their success?

**Proposition 2.** For any distribution of the opposing party’s candidates, there exists a unique level of heterogeneity that maximizes a party’s probability of winning, given its platform location. This optimal level of heterogeneity is increasing in distance between the party’s platform and the median voter.

(Formal Statement: For a fixed \( \mu_i \) and \( f_j \), where \( f_j \neq \delta_0 \), there exists a unique value of \( \sigma_i^* \), that maximizes \( P_i \). For a fixed \( f_j \), \( \sigma_i^* \) is increasing in \( |\mu_i| \).)\(^8\)

As this theorem demonstrates, heterogeneity is indeed an important determinant of party success: a unique level of heterogeneity maximizes a party’s chance of winning, and optimal heterogeneity varies systematically with a party’s distance to the median voter. The finding is particularly important because heterogeneity is largely absent from models of spatial competition. Casting a wide net or broadening the party’s campaign message may be a critical, yet unexplored, component of many party leaders’ electoral strategies.
Figure 1: The distribution of candidates for three possible values of \( \sigma_i \) and a fixed \( c_j \). Party \( i \) maximizes the probability that it selects a candidate in \([-|c_j|, |c_j|]\) when its heterogeneity equals \( \sigma_{i2} \).

The proof’s underlying logic is as follows. A homogeneous party positioned far from the median voter has no chance of winning. By diversifying its candidate pool, the party increases its likelihood of nominating a candidate closer to the voter than the opposing party’s nominee. Yet, too much heterogeneity quickly becomes a liability. Beyond a certain point, increasing heterogeneity decreases the probability that a party’s candidate is closer to the voter than their opponent.

For intuition, Figure 1 depicts a scenario where party \( i \)'s platform (\( \mu_i \)) is farther from the median voter (0) than the opposing party’s nominee. (In the example, \( c_j \) is fixed.) The figure shows three candidate distributions for party \( i \). Its platform is identical in each; the only variation is in its heterogeneity, where \( \sigma_{i1} < \sigma_{i2} < \sigma_{i3} \). As we can see, the density between \(-|c_j|\) and \(|c_j|\) is greatest for the distribution with medium variance (\( \sigma_{i2} \)). A fairly homogeneous party (e.g. \( \sigma_{i1} \)) can improve its chances by widening the candidate pool. But a party that is very diverse (e.g. \( \sigma_{i3} \)) too often selects candidates that are out of touch with the interests of the median voter.

In fact, perfect homogeneity is optimal only when a party’s platform is at the position of the median voter, or when the opposing party is also perfectly homogeneous and at least as far from the median.\(^9\) Taken with Proposition 1, this implies that if parties can freely move their platforms and adjust heterogeneity, two homogeneous parties will converge to the position of the median voter. In equilibrium, the parties perfectly resemble those in the Downsian model.

For parties off median, optimal heterogeneity increases in platform distance. Figure 2 plots the probability that party \( i \) wins the election as a function of its standard deviation, \( \sigma_i \), for varying values of its platform location, \( \mu_i \), and a fixed distribution for party \( j \). This function is single peaked in \( \sigma_i \) for all values of \( \mu_i \). (Note that every level
Figure 2: The probability that party $i$ wins the election, for varying values of $\mu_i$ and $\sigma_i$. Both parties have normal distributions, and $\mu_j = 4$ and $\sigma_j = 2$. For each $\mu_i$, there is a unique value of $\sigma_i$ that maximizes party $i$’s probability of winning.

Clearly, a party’s recruitment strategy will depend on its popularity, and the level of heterogeneity has a significant impact on election outcomes. In Figure 2, a party located twice as far from the median voter as its opponent ($\mu_i = 8$), can boost its chance of winning from practically zero (with perfect homogeneity) to almost a quarter of the time (with $\sigma_i = 7$). The same change in heterogeneity for a party located at the position of the median voter will cut its chances of winning from near certainty to less than half the time. More generally, if a party successfully positions itself close to the center of the voter distribution, it should hone its appeal and select candidates who adhere to the party line. In contrast, if their party’s platform or sitting president is unpopular, leaders may distance themselves by emphasizing the diversity of opinions within the party and recruiting a wide variety of candidates from different social organizations or by asking volunteers to suggest qualified individuals. Leaders may not wield perfect control over the set of candidates who vie for their party’s nomination, but because the model finds that a party’s probability of winning is single-peaked in heterogeneity, if they can exert any influence over heterogeneity the model predicts in which direction they will do so. This appears consistent with efforts by some party leaders to increase inclusiveness in the wake of electoral
Proposition 3. When two parties are equidistant from the median voter, the more homogeneous party is more likely to win. When they are not equidistant, the closer party’s optimal heterogeneity is strictly less than the farther party’s heterogeneity.

(Formal Statement: If \(|\mu_i| = |\mu_j|\) then \(P_i > P_j\) if \(\sigma_i < \sigma_j\). If \(|\mu_i| < |\mu_j|\) then \(\frac{\partial P_i}{\partial \sigma_i} < 0\) when \(\sigma_i \geq \sigma_j\), which implies that for any fixed \(\sigma_j\) the value of \(\sigma_i\) that maximizes \(P_i\) is strictly less than \(\sigma_j\).)

When two parties are equidistant from the median voter, they should both choose levels of heterogeneity that undercut their opponent. Anticipating their competitors’ best response distributions leads both parties (regardless of how far apart they are) to adopt zero heterogeneity from the start.

This equilibrium is very sensitive to small perturbations in platform locations. If two homogeneous parties are asymmetrically positioned around the median voter, the farther party loses with certainty. In fact, the closer party should always opt for a less diverse candidate pool than the farther party.

Proposition 4. For any pair of party platforms, there exists a pair of heterogeneity strategies that constitutes a Nash equilibrium. This equilibrium is unique except when exactly one party has a platform equal to the position of the median voter.

(Formal Statement: For any \(\mu_i\) and \(\mu_j\), there exists a \(\sigma_i^*\) and \(\sigma_j^*\) such that \(\sigma_i^*\) is the optimal standard deviation for party \(i\) when party \(j\) has standard deviation \(\sigma_j^*\) and vice versa; i.e. \((\sigma_i^*, \sigma_j^*)\) is a Nash equilibrium. This equilibrium is unique in all cases except when \(\mu_i = 0\) and \(\mu_j \neq 0\) or vice versa.)

Above we saw that when platforms are equidistant from the median voter, the equilibrium strategy for both parties is perfect homogeneity. Proposition 4 suggests we should also observe stability when platforms are unevenly matched. In equilibrium, the closer party will recruit candidates who toe the party line while the farther party expands its nomination pool, even at the expense of including potential candidates with positions more extreme than the party platform.\(^{12}\) Taken with insights derived in Proposition 2, the results indicate that equilibrium levels of heterogeneity are positive and increasing in platform distance for all parties with off-median platforms. In the context of party signals, Propositions 3 and 4 imply that popular parties should present less ambiguous messages and foster more consistent reputations than parties out of step with the electorate. As political events or “partisan tides” shift the electorate in favor of one party or another, parties should adjust their equilibrium levels of heterogeneity.

In other cases, party platforms change while the distribution of voters remains stable. Over the past 35 years, the Democratic and Republican Party platforms have increasingly moved toward opposite poles on the ideological spectrum (McCarty, Poole, and Rosenthal 2006, Theriault 2008). A number of factors may explain this recent divergence, including partisan redistricting (Carson et al. 2007), a rise in inequality (McCarty, Poole, and Rosenthal 2006), extreme party activists (Aldrich 2011, Theriault 2008), and increased voter sorting along ideological and partisan...
lines (Levendusky 2009). In this era of polarized partisanship, will parties adjust their heterogeneity? Proposition 5 suggests they should.

**Proposition 5.** As the party platforms become more polarized around the median voter, their equilibrium levels of heterogeneity increase.

(Formal Statement: Suppose \((\sigma^*_i, \sigma^*_j)\) is a Nash equilibrium for the pair of platforms \((\mu_i, \mu_j)\). Then \((k\sigma^*_i, k\sigma^*_j)\) is a Nash equilibrium for the platforms \((k\mu_i, k\mu_j)\) for any \(k \geq 0\).)

As the two parties’ platforms simultaneously move away from the median voter, parties should broaden their appeal and nominate candidates with a wider range of positions. Extreme conservatives and liberals will be as likely to run as moderate candidates, and both primary and general elections may prove more volatile, with greater swings in candidate positions from one election to another.

Proposition 5 appears to contradict evidence that parties have become more coherent as they have moved apart (McCarty, Poole, and Rosenthal 2006). But there are several important differences between the model’s prediction and the empirical record. First, by polarization Proposition 5 refers specifically to a simultaneous shift in party platforms away from the median voter. As the proof demonstrates, increasing polarization (as it is defined here) is formally equivalent to changing the scale on which parties and voters’ policy positions are measured. The level of overlap (that is, the probability that a Democrat is to the right of a Republican) remains constant. Second, the heterogeneity model describes the preferences of all candidates – not just those who win the election. Winners may disproportionately represent one branch of the party, and thus a more homogeneous pool.

Moreover, increased homogeneity may be explained by a growing symmetry in party platforms around the median voter. Recall from Proposition 3 that two equidistant parties will adopt perfect homogeneity in equilibrium (regardless of their absolute distance). As the two parties’ platforms are increasingly spaced evenly around the median voter (by either the farther party moving closer to the median voter, or the closer party moving farther away), their average equilibrium heterogeneity levels will decline and approach zero. If Democrats’ and Republicans’ positions have become more symmetric around the median voter as their platforms move apart, then their average level of heterogeneity should decline.

Testing this hypothesis requires a measure of the distance between each party’s platform and the median voter over time. Most studies on polarization focus on the distance between the two parties rather than their relative distance to the median voter. But, as the two parties approach equidistance, we would expect elections to become more competitive. Volatility should be high, and majorities in the legislature should be small. This is consistent with the House election record over the past 18 years. After decades of Democratic dominance, control of the House is now anyone’s game. Since 1995, the majority party has held an average of less than 55 percent of the seats, and small shifts in voter preferences have produced significant changes in party control. Consequently, it may prove valuable in future research to further investigate the relationship between party symmetry and heterogeneity.
More generally, the extent to which American parties are involved in identifying potential candidates is itself an empirical question for which current research does not provide a definitive answer. While conventional wisdom holds that the U.S. primary system minimizes the role of parties in candidate selection, recent research suggests that party leaders work behind the scenes to shape the candidate pool (Cohen et al. 2008, Frendreis, Gibson, and Vertz 1990, Sanbonmatsu 2006). Leaders may recruit candidates by attending local association meetings, running media campaigns, holding press conferences, or asking party volunteers to recommend qualified people. In a nationwide study, Sanbonmatsu (2006) finds that over 75 percent of state party and legislative caucus leaders are actively involved in recruiting candidates, providing endorsements in primaries, campaigning for incumbents’ renomination, or even dissuading candidates who are out-of-step with the party platform. As Sanbonmatsu argues, “the party may actively recruit precisely because its hands are tied at the primary stage” (2006, 45). Future research would benefit from better understanding the role of leaders in shaping the candidate pool.

Extensions

In the model presented above, parties are depicted as collections of candidates with varying positions rather than single actors with unique policy stances. Parties should strategically diversify their candidate pool to compensate for off-median positions, and a Nash equilibrium exists for any two party platform pairs where the party closest to the median is more homogeneous than its competitor. These results may appear to depend on two assumptions that could limit the model’s real world applicability: first, voters care only about the positions of the two candidates running in their district; and second, voters observe candidates’ positions with certainty. To address the first assumption, I extend the model in two ways. First, voters are allowed to observe the positions of multiple candidates within each party and then vote for the party with the most appealing set of candidates. Second, I consider the case in which voters care about the positions of both candidates and party platforms. Voters weigh the benefit of selecting a member who well-represents their position with one who will contribute to the legislative success of their preferred party. To test whether or not the findings persist in the face of voter uncertainty, the model is extended to constrain voters’ observations to a signal about candidate preferences rather than a clear and fully informative description of future behavior. When uncertainty is a function of party heterogeneity, voters are reluctant to support a candidate, and thus a party, whose position is vague or variable.

Multiple Candidates and Multiple Signals

Up to now, voters have simply observed the positions of the two candidates running in their district. In reality, however, voters may learn and care about the positions of many different candidates in each party. For example, voters surely incorporated new information about the Tea Party’s positions and candidates into their evaluations of Republican
contenders in 2010 and 2012, even in House districts where the GOP’s candidate was unaffiliated with the movement. Knowing that legislators must work together once they are in Congress, a voter may decide to support the party whose set of candidates is most sympathetic to their preferred policy outcomes instead of simply the candidate whose position is closest to their own (Austen-Smith 1984).

To incorporate this dynamic into the model, suppose voters observe \( N_p \) candidates’ positions for party \( p \):
\[
\{ c_{p1}, c_{p2}, \ldots, c_{pN_p} \},
\]
where \( N_p > 1 \), and each candidate is independently drawn from the distribution \( f_p \). Voters select the candidate whose party has an average candidate position,
\[
\bar{c}_p = \frac{1}{N_p} \sum_{l=1}^{N_p} c_{pl},
\]
closest to their ideal policy.

Because the standard deviation of the distribution of \( \bar{c}_p \) is smaller than the standard deviation of the distribution of \( c_p \), parties must increase their optimal heterogeneity as voters observe multiple candidates’ positions. For example, if \((\sigma_i^*, \sigma_j^*) \) is a Nash equilibrium for the pair of platforms \((\mu_i, \mu_j) \) when voters observe a single candidate, and candidate preferences in both parties are normally distributed, then \((\sigma_i^* \sqrt{N_i}, \sigma_j^* \sqrt{N_j}) \) is a Nash equilibrium when voters observe \( N_i \) and \( N_j \) candidates from parties \( i \) and \( j \), respectively. Optimal heterogeneity increases (at a decreasing rate) with the number of candidates a voter observes.

Recall that an alternative way of conceptualizing the model starts with the premise that voters care only about party positions, and parties signal their platform location through campaign messages, television advertisements, or the behavior of elected officials. Voters may learn about the party’s position indirectly from activists or interest groups, or they may observe “accidental data” about the party’s position by overhearing a political conversation among strangers or driving along a highway dotted with political billboards (Levendusky 2009). Applying the model to this context implies that parties with platforms in line with the median voter should send consistent signals about their positions. But a clear message to voters will convey the party’s position “for good or ill” (Aldrich and Griffin 2010, 602), and those parties with distant platforms may find it advantageous to be ambiguous and emphasize the diversity of opinions within the party. As voters learn more about their positions – due to high-profile campaigns, greater attention in the media, or heightened voter interest in politics – parties should broadcast a wider variety of messages.\(^{16}\)

**Voters Observe Platforms and Candidates**

Of course, voters most likely care about both a party’s platform (or the average position of its candidates) and the position announced by each candidate in their district (Cox and McCubbins 1993). They will weigh the importance of being represented by a party that is likely to pursue legislation close to their ideal point and electing a candidate with whom they identify. Voters may also weigh party and candidate positions if they expect candidates to move toward their party’s median member once elected, or if the party label conveys additional, or more credible, information about candidates’ positions. To incorporate party platforms into the model, suppose that a voter’s utility is based on their estimate of the party’s platform and the candidate that party nominates in their district:

\[
u(v,z), \text{ where } z = (1-\alpha)\bar{c}_p + \alpha c_p, \text{ and } 1-\alpha \text{ and } \alpha \text{ are the weights a voter assigns to their estimate of the party’s platform and}
\]
candidate’s position, respectively. When \( \alpha = 1 \), the model reduces to its original state, and a party’s optimal level of heterogeneity remains equal to \( \sigma^*_p \). When \( \alpha \) is less than one, the voter’s utility depends in part on their estimate of the party’s platform, and parties should increase heterogeneity.\(^{17}\) To the extent they can influence how voters consider the relative contributions of each position, leaders may strategically draw attention to the advantages of their party’s platform, record, or overall set of members (if they are close to the median) or run more candidate-centered campaigns (if they are far).

The value of strategic heterogeneity diminishes as voters learn more about each party’s position. Thus, to retain heterogeneity as a tool, both parties should increase the scope of their candidate pool in order to misrepresent their platform’s true location. This result differs significantly from previous work, which argues that parties should enforce homogeneity because risk-averse voters rely heavily on party labels for information about candidates’ positions (Aldrich and Griffin 2010, Ashworth and Bueno de Mesquita 2008, Grynaviski 2010, Snyder and Ting 2002).\(^{18}\) In contrast, the strategic heterogeneity model allows voters to observe candidates’ positions with certainty but assumes that parties are often unable to perfectly select candidates with ideal points that match that of the median voter. As a result, it is precisely when reputations matter most that parties should increase heterogeneity.

**Voter Uncertainty**

Thus far, I have assumed that candidates’ positions (and in the previous sections, party signals) are accurate and credible. Yet in reality, candidates may obscure their position – strategically or not – by sending imprecise signals about their policy stances. Once in office, candidates may deviate from the positions they declared during their campaigns. Savvy voters will factor this uncertainty into their assessments of candidates and possibly alter their vote. Previous research has argued that party heterogeneity is an electoral liability precisely because it increases uncertainty about candidate locations (Ashworth and Bueno de Mesquita 2008, Grynaviski 2010, Snyder and Ting 2002, Woon and Pope 2008). Risk-averse voters may be more willing to vote for a distant candidate whose position is known than a candidate whose expected position is closer, but uncertain.\(^{19}\) Thus, this section asks: when candidate positions are uncertain and voters are risk averse, do parties revert to acting as homogeneous teams?

Let us suppose that a voter interprets the candidate’s position for party \( p \) as a probability distribution, \( g_p(x) \), where \( E[x] = c_p \). The voter’s uncertainty about the candidate’s position is captured by the standard deviation, \( s_p \), of \( g_p(x) \). For simplicity, voter utility is modeled as a standard quadratic loss function: \( u(v,x) = -(v - x)^2 \). Thus, the median voter’s \((v = 0)\) expected utility from candidate \( c_p \) is

\[
EU(c_p) = -\int_{-\infty}^{\infty} g_p(x) x^2 \, dx = -c_p^2 - s_p^2,
\]

and they will vote for party \( i \) when
\[-c_i^2 - s_i^2 > -c_j^2 - s_j^2.\]

If both parties’ candidates have equal uncertainty (i.e. \(s_i = s_j\)), the results are identical to those of the model without uncertainty: voters simply support the candidate whose position most closely matches their own, and party strategies remain the same as in the baseline model.

If the uncertainty levels for the two parties are unequal, the party with greater uncertainty is hurt most. As uncertainty increases, the share of a voter’s utility that is based on distance decreases. Thus, if both parties face an equal probability of being more uncertain than one another, uncertainty will on average help the farther party more than the closer party.\(^{20}\)

Of course, a party’s uncertainty may be positively related to its heterogeneity, causing risk-averse voters to punish parties with wide-ranging candidate pools. To investigate this, suppose \(s(\sigma_p)\) is the uncertainty around \(c_p\), where \(s\) is an increasing function. In this case, the voter supports party \(i\) when

\[-c_i^2 - s(\sigma_i)^2 > -c_j^2 - s(\sigma_j)^2.\]

Intraparty heterogeneity is now a partial liability. Although heterogeneity can increase a party’s chance of fielding a candidate close to the median voter, it may also hurt that party’s chance of winning when voters are risk averse. As important, the party with greater heterogeneity – which is the farther party in equilibrium – will be hurt most. Both parties should respond to voter uncertainty by reducing candidate heterogeneity, but the comparative statics stay the same: the optimal level of heterogeneity is higher for the farther party than the closer party.\(^{21}\)

For example, let \(s\) be a simple linear function: \(s(\sigma_p) = k\sigma_p\), for any \(k > 0.\)\(^22\) Figure 3 plots the probability of winning for varying values of heterogeneity and \(k\) for two parties with normal distributions. The solid line represents the probability that party \(i\) wins when there is no uncertainty (i.e. \(k = 0\)), and the dashed or dotted lines depict the same probability when \(k\) takes on three positive values. As we can see, to the left of \(\sigma_i = 2\), party \(i\) is more homogeneous than party \(j\), and its probability of winning increases in \(k\). To the right of \(\sigma_i = 2\), party \(i\) is more heterogeneous than party \(j\), and its probability of winning decreases in \(k\). Overall, the optimal level of heterogeneity is decreasing in \(k.\)\(^{23}\)

Thus, while extending the model in varying ways – adding multiple candidates, voters who care about party platforms, and uncertainty – changes a party’s optimal heterogeneity in important respects, the model’s primary predictions persist: a unique level of heterogeneity maximizes a party’s success; this optimal level of heterogeneity is typically non-zero; and the closer party should be more homogeneous than the farther party.
Figure 3: The probability party $i$ wins the election for varying values of $k$ and $\sigma_i$, where $c_i \sim N(5, \sigma_i^2)$ and $c_j \sim N(3, 2^2)$. 
Conclusion

Existing research typically assumes or implies that party leaders prefer homogeneity. When their members are in agreement, majority party leaders can pass partisan legislation more efficiently (Rohde 1991). Minority party leaders are also better positioned to block opposing legislation when their members act as a coherent team. In the electoral arena, homogeneous parties offer more reliable signals about their candidates’ preferences and their platform’s location (Ashworth and Bueno de Mesquita 2008, Grynviski 2010, Snyder and Ting 2002, Woon and Pope 2008). For risk-averse voters, greater certainty about a party’s position can make all the difference in which party to support (Alvarez 1998, Bartels 1986, Enelow and Hinich 1981). While party leaders cannot nominate every candidate who runs under their label, they can encourage some candidates more than others by offering campaign assistance, labor, or endorsements, or even by dissuading would-be candidates from running in the party primary. To the extent party leaders can influence candidate selection, political science offers a clear prescription: enforce homogeneity.

The heterogeneity model presented in this article largely agrees that homogeneity is the optimal strategy – for parties close to the median voter. Parties that are well positioned to win a majority of votes should recruit candidates whose positions align with the party platform. But a homogeneous party committed to a platform distant from the median voter is doomed to defeat. For parties that are out of step, increasing candidate heterogeneity may be the only viable option. Recent history suggests that national parties are playing an increasingly important role in candidate recruitment and campaigning. As a result, losing parties, like the Republicans in 2013, may put even greater efforts into expanding their base and opening up their recruitment process. In contrast, parties that are more in line with the interests of the median voter should focus steadily on their platform and hone their electoral appeal.

Thus, while previous research argues that heterogeneity decreases a party’s election prospects, the model in this article finds that parties may choose to be strategically heterogeneous precisely because their platforms are unpopular. The model offers clear predictions as to how leaders will alter heterogeneity even when they cannot reach an equilibrium, and the comparative statics present an interesting asymmetry: while the party close to the median voter should be fairly homogeneous, the farther party should seek to diversify its candidate pool.

Important extensions to the model – such as adding informative party labels or voter uncertainty – do not change the comparative statics. They do, however, affect the magnitude of a party’s optimal heterogeneity. When voters are hyper-vigilant, the media is widespread, or campaigns are long and drawn out, parties should send more varied signals about their positions. Similarly, when voters care about the location of the party platform as well as of their own candidate, parties should become more heterogeneous. And heterogeneity may heighten voter uncertainty about candidates’ positions, dampening the farther party’s prospects even more, and reducing optimal heterogeneity levels.

Future research may extend and test the model in several ways. Are majority parties (presumably closer to the median voter) indeed more homogeneous than minority parties? Because a party’s platform location and heterogeneity are not entirely independent, it is also important to investigate the tradeoffs for strategic parties in altering each. The
political or institutional environment may make it easier for some parties to change their platform or heterogeneity more than others. Majority parties can credibly claim to maintain the position they held in the previous term, while minority parties may enjoy more flexibility (but less credibility) in changing platforms or setting heterogeneity. In addition, future work may add a candidate affiliation stage to the model, as in Ashworth and Bueno de Mesquita (2008) or Snyder and Ting (2002), or endogenize party platform locations. Heterogeneity as a strategy depends in party on party and legislative caucus leaders’ control over recruitment. Candidate distributions are most likely to follow those prescribed by the heterogeneity model when state laws and party organizational rules afford leaders greater control over nominations. Finally, it is important to recognize that homogeneity may help parties pass or block legislation or attract partisans who donate valuable resources to political campaigns (Rohde 1991). Depending on the goals of legislators – policy or office – and their expected electoral prospects down the road, heterogeneity may not always be optimal, even for parties that are out-of-step. All told, modeling party heterogeneity as a strategic choice has important implications for understanding legislative success as well as electoral competition.
Appendix

A Proofs of Propositions

Proof of Proposition 1

Equation (2) can be rewritten as \( \int_{-\infty}^{\infty} f_j(c_j) \left[ \Phi\left(\frac{|c_j| - \mu_i}{\sigma_i}\right) - \Phi\left(\frac{-|c_j| - \mu_i}{\sigma_i}\right) \right] \, dc_j \), where \( \Phi \) is the integral of \( \phi \). Taking the derivative with respect to \( \mu_i \) gives

\[
\frac{\partial P_i}{\partial \mu_i} = \int_{-\infty}^{\infty} f_j(c_j) \left( \frac{-1}{\sigma_i} \right) \left[ \phi\left(\frac{|c_j| - \mu_i}{\sigma_i}\right) - \phi\left(\frac{-|c_j| - \mu_i}{\sigma_i}\right) \right] \, dc_j.
\]

Because \( \phi(x) \) has mean zero, when \( \mu_i > 0 \), \( \phi\left(\frac{|c_j| - \mu_i}{\sigma_i}\right) > \phi\left(\frac{-|c_j| - \mu_i}{\sigma_i}\right) \), and equation (3) is negative. Similarly, when \( \mu_i < 0 \), \( \phi\left(\frac{|c_j| - \mu_i}{\sigma_i}\right) < \phi\left(\frac{-|c_j| - \mu_i}{\sigma_i}\right) \), and equation (3) is positive. Thus, the marginal effect of an increase in \(|\mu_i|\) is negative, and the marginal effect of a decrease in \(|\mu_i|\) is positive. Since \( P_j = 1 - P_i \), the marginal effect of a decrease in \(|\mu_i|\) is negative for party \( j \). \( \square \)

Proof of Proposition 2

Part I. Existence and Uniqueness of a Global Maximum

I begin by proving existence, and then I prove uniqueness of a global maximum. Since throughout the proof I am concerned with the effect of \( \sigma_i \) on \( P_i \), I treat \( P_i \) as a single variable function of \( \sigma_i \), \( P_i(\sigma_i) \), and let \( P'_i(\sigma_i) \) denote the derivative of \( P_i \) with respect to \( \sigma_i \).

If \( \sigma_i = \sigma_j \), then party \( i \) wins the election with some positive probability \( p \). Since the \( \lim_{\sigma_i \to \infty} P_i(\sigma_i) = 0 \), there exists an \( S \) such that if \( \sigma_i > S \) then \( P_i(\sigma_i) < p \). Because \( P_i(\sigma_i) \) is continuous on the compact interval \([0, S]\), it must have at least one global maximum on \([0, S]\). This maximum must be a global maximum since \( P_i(\sigma_i) < p \) for \( \sigma_i \notin [0, S] \), but \( P_i(\sigma_i) = p \) for some \( \sigma_i \in [0, S] \).

Now I turn to the proof of uniqueness. Let \( P_i(\sigma_i|c_j) \) denote the probability that party \( i \) wins the election as a function of \( \sigma_i \) given a fixed position, \( c_j \), for the opposing candidate,

\[
P_i(\sigma_i|c_j) = \int_{-|c_j|}^{\infty} \frac{1}{\sigma_i} \phi\left(\frac{c_i - \mu_i}{\sigma_i}\right) \, dc_i = \Phi\left(\frac{|c_j| - \mu_i}{\sigma_i}\right) - \Phi\left(\frac{-|c_j| - \mu_i}{\sigma_i}\right).
\]

For a given \( c_j \), the marginal effect of party \( i \)'s heterogeneity on the probability that party \( i \) wins is

\[
P'_i(\sigma_i|c_j) = \frac{-|c_j| + \mu_i}{\sigma_i^2} \phi\left(\frac{|c_j| - \mu_i}{\sigma_i}\right) - \frac{|c_j| + \mu_i}{\sigma_i^2} \phi\left(\frac{-|c_j| - \mu_i}{\sigma_i}\right).
\]
Because $\phi$ is symmetric around zero, this can be rewritten as

$$P'_i(c_j) = \frac{\mu_i - |c_j|}{\sigma_i^2} \phi \left( \frac{\mu_i - |c_j|}{\sigma_i} \right) - \frac{\mu_i + |c_j|}{\sigma_i^2} \phi \left( \frac{\mu_i + |c_j|}{\sigma_i} \right).$$ \hspace{1cm} (5)

Setting equation (5) equal to zero and multiplying both sides by $\sigma_i$, the first order condition is given by

$$\frac{\mu_i - |c_j|}{\sigma_i} \phi \left( \frac{\mu_i - |c_j|}{\sigma_i} \right) - \frac{\mu_i + |c_j|}{\sigma_i} \phi \left( \frac{\mu_i + |c_j|}{\sigma_i} \right) = 0. \hspace{1cm} (6)$$

Rearranging terms, equation (6) becomes

$$\frac{\mu_i + |c_j|}{\mu_i - |c_j|} = \phi \left( \frac{\mu_i - |c_j|}{\sigma_i} \right) / \phi \left( \frac{\mu_i + |c_j|}{\sigma_i} \right). \hspace{1cm} (7)$$

By assumption the right hand side of equation (7) is strictly monotonic as a function of $\sigma_i$ and therefore there is at most one $\sigma_i$ satisfying the equation. If there is no such solution, then equation (4) must be decreasing for all $\sigma_i$, so $P_i(c_j|c_j)$ is maximized at $\sigma_i = 0$. If there is a solution, it is unique, and thus there is a unique value of $\sigma_i$ that maximizes $P_i(c_j|c_j)$. Let $\sigma^*_i(c_j)$ denote this unique maximum. Then, because $f_j$ is symmetric and single-peaked, and because $\sigma^*_i(c_j)$ is continuous and monotonic in $c_j$, there exists a unique $\sigma^*_i$ that maximizes $P_i$. Numerical computations confirm this; for examples, see Figure 2.

\[\square\]

**Observation 1.** If $\mu_i = 0$, then $\frac{\partial P_i}{\partial \sigma_i} < 0$ for all $\sigma_i > 0$.

Observation 1 follows directly from evaluating equation (5) at $\mu_i = 0$.

**Part II. Optimal heterogeneity is increasing in $|\mu_i|$.**

Because $\sigma^*_i(c_j)$ is either zero or the unique solution to equation (7), if $|c_j| \geq |\mu_i|$ then $\sigma^*_i(c_j) = 0$ so the proposition holds trivially. Below I prove that $\sigma^*_i(c_j)$ is increasing in $|\mu_i|$ for the case where $|c_j| < |\mu_i|$ when $\mu_i > 0$. The case when $\mu_i < 0$ is similar.

Let $|\mu_1| < |\mu_2|$, and let $\sigma_i$ and $\sigma_j$ be the optimal heterogeneity levels given $\mu_i = \mu_1$ and $\mu_i = \mu_2$, respectively. Then $\sigma_1$ satisfies

$$\frac{(\mu_1 + |c_j|)}{(\mu_1 - |c_j|)} = \phi \left( \frac{\mu_1 - |c_j|}{\sigma_1} \right) / \phi \left( \frac{\mu_1 + |c_j|}{\sigma_1} \right),$$

and $\sigma_2$ satisfies

$$\frac{(\mu_2 + |c_j|)}{(\mu_2 - |c_j|)} = \phi \left( \frac{\mu_2 - |c_j|}{\sigma_2} \right) / \phi \left( \frac{\mu_2 + |c_j|}{\sigma_2} \right).$$
To show that $\sigma_1 < \sigma_2$, I introduce the intermediate point $s$, defined by

$$\frac{(\mu_2 + |c_j|)}{(\mu_2 - |c_j|)} = \frac{\phi \left( \frac{\mu_1 - |c_j|}{s} \right)}{\phi \left( \frac{\mu_1 + |c_j|}{s} \right)}.$$

Since $\frac{(\mu + |c_j|)}{(\mu - |c_j|)}$ is decreasing in $\mu$ and $\frac{\phi \left( \frac{\mu - |c_j|}{s} \right)}{\phi \left( \frac{\mu + |c_j|}{s} \right)}$ is decreasing in $\sigma$, $s > \sigma_1$. Then, since $\frac{\phi \left( \frac{s - |c_j|}{\sigma} \right)}{\phi \left( \frac{s + |c_j|}{\sigma} \right)}$ is increasing in $\mu$, $s < \sigma_2$. Therefore $\sigma_1 < \sigma_2$. Numerical computations confirm that this extends to any symmetric, single-peaked distribution of $c_j$.

**Proof of Proposition 3**

Suppose $\mu_i = \mu_j$ and $\sigma_j > 0$. Then, if $\sigma_i = \sigma_j$, $P_i = 0.5$. If instead $\sigma_i = 0$, then $P_i = \int_{-\infty}^{\mu_i} f_j(c_j) \, dc_j + \int_{\mu_i}^{\infty} f_j(c_j) \, dc_j = \int_{-\infty}^{\mu_i} f_j(c_j) \, dc_j + 0.5 > 0.5$. Because $P_i$ is a single-peaked function of $\sigma_i$, the value of $\sigma_i$ that maximizes $P_i$ must be strictly less than $\sigma_j$. Since Proposition 2 demonstrates that $\sigma_i^*$ is increasing in $|\mu_i|$ (holding $f_j$ constant), $\sigma_i^*$ must also be less than $\sigma_j$ when $|\mu_i| < |\mu_j|$.

**Proof of Proposition 4**

When $\mu_i = \mu_j$, Proposition 3 implies that $\sigma_i = \sigma_j = 0$ is a Nash equilibrium. If $\mu_i = 0 < |\mu_j|$ then $(\sigma_i^*, \sigma_j^*) = (0, \sigma_j^*)$ is a Nash equilibrium for any value of $\sigma_j^*$. Suppose $0 < |\mu_i| < |\mu_j|$. Let $BR_i : [0, \infty) \to [0, \infty)$ be defined as party $i$’s best response given $\sigma_j$: $BR_i(\sigma_j) = \sigma_i^*$. Similarly, define $BR_j : [0, \infty) \to [0, \infty)$ by $BR_j(\sigma_i) = \sigma_j^*$. Define the curves $\mathcal{I}$ and $\mathcal{J}$ in $[0, \infty) \times [0, \infty)$ by $\mathcal{I} = \{ (\sigma_i, \sigma_j) | \sigma_i = BR_i(\sigma_j) \}$ and $\mathcal{J} = \{ (\sigma_i, \sigma_j) | \sigma_j = BR_j(\sigma_i) \}$. Then any intersection of $\mathcal{I}$ and $\mathcal{J}$ is a Nash equilibrium.

By Proposition 3, the curve $\mathcal{I}$ is contained in $\{ (\sigma_i, \sigma_j) | \sigma_i \leq \sigma_j \}$. If $\sigma_i = 0$ then the best response for party $j$ is to adopt some non-zero heterogeneity, $\sigma_j > 0$ (otherwise party $i$ wins the election with certainty). As $\sigma_i \to \infty$, $BR_j(\sigma_i) \to 0$. Since the curve $\mathcal{J}$ is continuous and goes from the point $(0, \sigma_j)$ on one end towards the limit $(\infty, 0)$ at the other, there must exist a point $(s_i, s_j)$ on $\mathcal{J}$ such that $s_i = s_j$. Let $X$ denote the closed subset of $(\sigma_i, \sigma_j) \in [0, \infty) \times [0, \infty)$ bounded by $\sigma_i = 0$, $\sigma_i = \sigma_j$, and the segment of $\mathcal{J}$ connecting $(0, \sigma_j)$ and $(s_i, s_j)$. Since $BR_i(0) = 0$ the curve $\mathcal{I}$ goes through the point $(0, 0) \in X$, and for $\sigma_j$ sufficiently large the point $(BR_i(\sigma_j), \sigma_j) \in \mathcal{I}$ is outside of $X$. Therefore, since $\mathcal{I}$ is continuous, this implies that $\mathcal{I}$ crosses the boundary of $X$. However, $\mathcal{I}$ cannot cross $\sigma_i = 0$ since every point of $\mathcal{I}$ has $\sigma_i \geq 0$, and $\mathcal{I}$ cannot cross $\sigma_i = \sigma_j$, since every point of $\mathcal{I}$ has $\sigma_i < \sigma_j$. Therefore, $\mathcal{I}$ must cross some part of the remaining boundary of $X$, which consists of the points of $\mathcal{J}$ with $\sigma_i \leq \sigma_j$. Thus, $\mathcal{I} \cap \mathcal{J} = \emptyset$, so a Nash equilibrium exists.

Next, I prove that this Nash equilibrium is unique except in the case where exactly one party has a platform equal to the position of the median voter. Because the case where $\mu_i = \mu_j = 0$ has already been shown to have a unique
equilibrium at \((0, 0)\), I will assume from here on that neither party’s platform is positioned at the ideal point of the median voter. Suppose that for the platform pair, \((\mu_i, \mu_j)\), there are two Nash equilibria: \((\sigma_{i1}, \sigma_{j1})\) and \((\sigma_{i2}, \sigma_{j2})\). Proposition 2 demonstrates that \(\sigma_i^*\) is unique given \(\mu_i\) and \(f_j\), except when \(f_j = \delta_0\). Then, if both \((\sigma_{i1}, \sigma_{j1})\) and \((\sigma_{i2}, \sigma_{j2})\) are Nash equilibria and \(\sigma_{i1} = \sigma_{i2}\), then \(f_j = \delta_0\). Similarly, if \(\sigma_{j1} = \sigma_{j2}\), then \(f_i = \delta_0\). Thus, if there are two Nash equilibria and neither party’s platform is positioned at the median voter, then it must be true that \(\sigma_{i1} \neq \sigma_{i2}\) and \(\sigma_{j1} \neq \sigma_{j2}\).

Given this, there are two possible cases: (1) \(P_i\) is greater at \((\sigma_{i1}, \sigma_{j1})\) than \((\sigma_{i2}, \sigma_{j2})\) or vice versa, or (2) \(P_i\) is equal at both \((\sigma_{i1}, \sigma_{j1})\) and \((\sigma_{i2}, \sigma_{j2})\). I will show that in both cases at least one party has an incentive to deviate.

For the first case, let us suppose without loss of generality that \(P_i(\sigma_{i1}, \sigma_{j1}) > P_i(\sigma_{i2}, \sigma_{j2})\). Because \((\sigma_{i1}, \sigma_{j1})\) is a Nash equilibrium, it must be the case that \(P_j(\sigma_{i1}, \sigma_{j1}) \geq P_j(\sigma_{i1}, \sigma_{j1}')\) for all \(\sigma_{j1}' \neq \sigma_{j1}\). This is equivalent to stating that \(P_i(\sigma_{i1}, \sigma_{j1}) \leq P_i(\sigma_{i1}, \sigma_{j1}')\). Because \(P_i(\sigma_{i1}, \sigma_{j1}) > P_i(\sigma_{i2}, \sigma_{j2})\) and \(P_i(\sigma_{i1}, \sigma_{j1}) \leq P_i(\sigma_{i1}, \sigma_{j1}')\), it must be the case that \(P_i(\sigma_{i1}, \sigma_{j1}) > P_i(\sigma_{i2}, \sigma_{j2})\), which implies that \(P_i(\sigma_{i1}, \sigma_{j1}) > P_i(\sigma_{i2}, \sigma_{j2})\). Thus, party \(i\) has an incentive to deviate from \((\sigma_{i2}, \sigma_{j2})\), and it is not a Nash equilibrium.

Now suppose that \(P_i(\sigma_{i1}, \sigma_{j1}) = P_i(\sigma_{i2}, \sigma_{j2})\). Then because \((\sigma_{i2}, \sigma_{j2})\) is an equilibrium, \(P_i(\sigma_{i1}, \sigma_{j1}) \geq P_i(\sigma_{i1}', \sigma_{j1})\) for all \(\sigma_{i1}' \neq \sigma_{i1}\). This implies that \(P_i(\sigma_{i2}, \sigma_{j2}) \geq P_i(\sigma_{i1}, \sigma_{j1})\), which is equivalent to saying that \(P_j(\sigma_{i1}, \sigma_{j1}) \leq P_j(\sigma_{i1}, \sigma_{j1}')\). Because \(P_i(\sigma_{i1}, \sigma_{j1}) = P_i(\sigma_{i2}, \sigma_{j2})\), it is also true that \(P_j(\sigma_{i1}, \sigma_{j1}) = P_j(\sigma_{i2}, \sigma_{j2})\). Thus,

\[
P_j(\sigma_{i1}, \sigma_{j1}) \leq P_j(\sigma_{i1}, \sigma_{j1}).
\]  

(8)

If the inequality in equation (8) is strict, party \(j\) has an incentive to deviate from \((\sigma_{i1}, \sigma_{j1})\), and it is not an equilibrium. If the two terms in equation (8) are equal, then \((\sigma_{i1}, \sigma_{j2})\) is also an equilibrium. But, as shown above, this is only possible if \(f_i = \delta_0\), which contradicts the assumption that neither party is positioned at the ideal point of the median voter.

**Proof of Proposition 5**

Let \(P_i^k(\sigma_i, \sigma_j)\) denote the probability that party \(i\) wins when parties \(i\) and \(j\) adopt platforms \(k\mu_i\) and \(k\mu_j\), and heterogeneity levels, \(\sigma_i\) and \(\sigma_j\), respectively. The proof proceeds by demonstrating that \(P_i(\sigma_i, \sigma_j) = P_i^k(\sigma_i, \sigma_j)\).

Substituting \(k\mu_i\) and \(k\mu_j\) into the equation for \(P_i\) gives

\[
P_i^k(\sigma_i, \sigma_j) = \int_{-\infty}^{\infty} \frac{1}{k\sigma_j} \phi\left( \frac{c_j - k\mu_j}{k\sigma_j} \right) \int_{-\infty}^{\infty} \frac{1}{k\sigma_i} \phi\left( \frac{c_i - k\mu_i}{k\sigma_i} \right) dc_i \; dc_j.
\]

In the second integral, changing variables using the substitution \(u = c_i/k\) gives

\[
P_i^k(\sigma_i, \sigma_j) = \int_{-\infty}^{\infty} \frac{1}{k\sigma_j} \phi\left( \frac{c_j - \mu_j}{\sigma_j} \right) \int_{-\infty}^{\infty} \frac{1}{\sigma_i} \phi\left( \frac{u - \mu_i}{\sigma_i} \right) du \; dc_j.
\]

21
Now, changing variables in the first integral using the substitution \( w = c_j/k \) gives

\[
P_k^i (k \sigma_i, k \sigma_j) = \int_{-\infty}^{\infty} \frac{1}{\sigma_j} \phi \left( \frac{w - \mu_j}{\sigma_j} \right) \int_{-|w|}^{|w|} \frac{1}{\sigma_i} \phi \left( \frac{u - \mu_i}{\sigma_i} \right) du \, dw,
\]

which is equal to \( P_i(\sigma_i, \sigma_j) \).

\[ \square \]

B Marginal Effects of Changes in Party Platform Location

Proposition 6. Shifting a party’s platform away from the median voter decreases its probability of winning:

a. less when its platform is very far from or very close to the median voter, and more when it is somewhere in between. (Formal Statement: The marginal effect, \( \frac{\partial P_i}{\partial |\mu_i|} \), approaches zero from below as \( |\mu_i| \) approaches infinity and as \( |\mu_i| \) approaches zero.)

b. less when it is very heterogeneous, and by twice the density of its opponent’s candidates at its platform when it is very homogeneous. (Formal Statement: The marginal effect, \( \frac{\partial P_i}{\partial |\mu_i|} \), approaches zero from below as \( \sigma_i \) approaches infinity and approaches \( -2 f_j(|\mu_i|) \) as \( \sigma_i \) approaches zero.)

A change in party platform matters most when parties are relatively equidistant from the median voter. Moreover, if if a party’s heterogeneity is sufficiently high, its probability of winning approaches zero, and a change in platform has little effect. If a party is perfectly homogeneous, its chance of winning is simply the probability that its platform is closer to the median voter than the candidate from the other party.

Because the probability that party \( j \) wins the election is \( 1 - P_i \), these results apply to the opposing party as well. Shifts in either party’s platform affect election outcomes more when the party changing platform locations is neither very close nor very far from the median voter. Changes in the party platform have the smallest effect when heterogeneity is high.

Proof

Proposition 6a follows directly from substituting \( |\mu_i| = 0 \), or taking the limit as \( |\mu_i| \) approaches infinity, in equation (3). Likewise, the first part of Proposition 6b follows directly from taking the limit as \( \sigma_i \) approaches infinity in equation (3).

For the second part of Proposition 6b, suppose that \( \sigma_i = 0 \). Then \( c_i = \mu_i \), and

\[
P_i = 1 - \int_{-|\mu_i|}^{|\mu_i|} f_j(c_j) \, dc_j = 1 - \left[ \int_{-|\mu_i|}^{0} f_j(c_j) \, dc_j + \int_{0}^{|\mu_i|} f_j(c_j) \, dc_j \right] = 1 + \left[ \int_{0}^{-|\mu_i|} f_j(c_j) \, dc_j - \int_{0}^{|\mu_i|} f_j(c_j) \, dc_j \right].
\]

(9)

22
Taking the derivative of equation (9) with respect to $|\mu_i|$ gives us:

$$\frac{\partial P_i}{\partial |\mu_i|} = -f_j(-|\mu_i|) - f_j(|\mu_i|) = -2f_j(|\mu_i|).$$

(10)

Because the derivative of $P_i$ with respect to $\mu_i$ is continuous as a function of $\sigma_i$, $\frac{\partial P_i}{\partial |\mu_i|}$ approaches $-2f_j(|\mu_i|)$ as $\sigma_i$ approaches zero.

Note that the result in Proposition 6a when $|\mu_i|$ approaches zero only holds when $\sigma_i > 0$. When $\sigma_i = 0$, the marginal effect of an increase in $|\mu_i|$ at $\mu_i = 0$ is given by substituting $\mu_i = 0$ into equation (10). □

Notes

1Nearly every empirical study on the matter finds that parties diverge from each other and from the median voter (see, e.g., Ansolabehere, Snyder and Stewart 2001, Burden 2004, Frendreis et al. 2003, Jessee 2010).

2Candidates from the same party may hold heterogeneous positions that reflect the diversity of districts’ median voters, but this mechanism cannot account for when or why the two parties differ in their heterogeneity levels.

3Examples of first-stage models include: a) platforms are drawn randomly from a distribution with expectation equal to the median voter; b) leaders inherit a platform equal to the mean position of the winning candidates (or party supporters) from their party in the last election; and c) leaders and activists bargain over choosing a platform that either maximizes expected vote shares or is consistent with core ideology.

4Candidates may prefer to affiliate with a party that is ideologically close (Ashworth and Bueno de Mesquita 2008, Snyder and Ting 2002), and in line with this reasoning I later assume that $f_p$ is symmetric and single peaked around $\mu_p$. Voters’ ideal points lie in the same dimension as candidates’ positions. Instead of policy, this dimension could reflect partisanship, ideology, or a single issue.

5The choice of first stage model may affect long-run dynamics, but will not alter the comparative statics and equilibrium results derived below. In this example, platforms and optimal heterogeneity levels remain constant if the distribution of party supporters is unaffected by election results. If partisan support is fluid and only reflects the location of each party’s candidate from the last election, platforms and heterogeneity may change from one election to another. Alternatively, we could imagine a case where selection is centralized, and party leaders nominate candidates. Leaders sample from the distribution (at positive cost per draw) and
select the candidate most likely to win. I do not explore this option below.

6This assumption essentially implies that \( f_p \) changes concavity only one time on either side of its mean and guarantees that the optimal level of heterogeneity is unique.

7If party \( i \)'s distribution is asymmetric or has more than one mode, \( P_i \) is not always maximized when the party’s mean position matches that of the median voter. In contrast, Proposition 1 does not depend on the opposition party’s distribution of candidates.

8When \( f_j = \delta_0 \) and \( \mu_i = 0, \sigma^*_i = 0 \). When \( f_j = \delta_0 \) and \( \mu_i \neq 0 \), any value of \( \sigma_i \) is optimal.

9See Observation 1 in Appendix A.

10Because candidate distributions are symmetric, an increase in heterogeneity also raises the probability that the party’s candidate will be positioned even farther away from the voter than the party platform. For this reason, increasing heterogeneity can never propel a party with a probability of winning less than 0.5 to a status where it is more likely than not to win.

11When \( \mu_i = 0 \) and \( \mu_j \neq 0 \), \((0, \sigma^*_j)\) is a Nash equilibrium for any value of \( \sigma^*_j \).

12Requiring candidate distributions be symmetric presents a boundary case. Party leaders who increase heterogeneity accept the possibility of recruiting radical candidates along with those who take more moderate, competitive positions. Since parties benefit from heterogeneity when candidates are drawn from a symmetric distribution around the party platform, they should also do so when distributions are asymmetric and leaders can exert more control over the candidate distribution or “target” candidates to specific districts.

13See Theriault (2008) for a comprehensive discussion of the sources of polarization.

14Ashworth and Bueno de Mesquita (2008) also define polarization as a simultaneous shift in party platforms, and find that polarization increases heterogeneity as a result of a similar change in scale. In their model, voters care less about uncertainty relative to platform positions as parties move apart, making (costly) screening less consequential.

15Note that this does not contradict Proposition 2. A party’s optimal heterogeneity is still increasing as its platform moves away from the median voter, holding its competitor’s distribution constant.

16Previous research argues that candidates may be strategically ambiguous about their positions on unpopular issues by avoiding those issues altogether (Page 1976, Campbell 1983). Other studies have found that politicians may deliberately obfuscate their positions if voters are risk seeking or hold “intense” preferences (Aragonès and Postlewaite 2002, Shepsle 1972), or when politicians plan to enact their preferred policy position once in office (Alesina and Cukierman 1990, Aragonès and Neeman 2000). In contrast, the
model presented in this article finds that purely office-seeking parties may be strategically heterogeneous even when voters are averse to risk (for this extension, see below).

Specifically, if candidate positions in party \( p \) follow a normal distribution, party \( p \) will optimize its heterogeneity when \( \sigma_p = \frac{\sigma^*_p}{\sqrt{\frac{1-\alpha^2}{N_p} + (\alpha)^2}} \).

Although, see Carson et al. (2010), who find that party cohesion in the legislature is not always beneficial for party success.

In contrast with the experimental results mentioned previously, several observational studies have found that voters behave in accordance with risk-averse utility functions (Alvarez 1998, Bartels 1986).

Specifically, suppose that \( s_p \sim \Theta \) for \( p \in \{i,j\} \), where \( \Theta \) is a distribution with non-negative support. Then, as \( \Theta \) increases (in the sense of first-order stochastic dominance), the farther party’s probability of winning improves and approaches 0.5.

Alternatively, heterogeneity could be negatively associated with uncertainty if voters believe that a candidate’s position is more credible when they are nominated by a party with high variance (see, e.g., Levy 2004). When uncertainty and heterogeneity are inversely related, both parties will increase equilibrium levels of heterogeneity.

The constant \( k \) effectively captures a party’s cost of heterogeneity due to voter uncertainty. It also signifies the relative assurance voters and parties have over candidates’ positions. (When \( k < 1 \), parties are less certain about their nominee’s position before selection than are voters after selection. When \( k > 1 \), the opposite is true.) Note that Snyder and Ting (2002) essentially assume that \( k = 1 \).

The comparative statics also hold when \( s \) is a function of multiple signals drawn from the party’s distribution, as in the first extension.

If candidates are unwilling to affiliate with heterogeneous parties (because they send unclear signals to voters about their positions), parties may choose to decrease overall levels of heterogeneity. Similarly, the number of Independent candidates may increase, particularly near heterogeneous parties.
References


