



The Imperfect Beliefs Voting Model

Benjamin Ogden
SBS-EM, ECARES, Université libre de Bruxelles
and Texas A&M University

April 2017

ECARES working paper 2017-20

The Imperfect Beliefs Voting Model*

Benjamin G. Ogden[†]

4-8-2017

Abstract

In real-life elections, voters do not have full information over the policy platforms proposed by political parties. Instead, they form (imprecise) beliefs. I propose a new model of partisan competition to represent the interaction of these beliefs with platform selection. Both parties gain more from appealing to the voters with more precise beliefs over their platform. Minority candidates viewed with less precision overall gain relatively more from outliers. Therefore, the Median Voter Theorem is recovered if and only if voters' policy preferences are uncorrelated with the precision of their beliefs about each candidate, and preferences are distributed symmetrically. Otherwise, even election-motivated parties diverge away from each other. As the population becomes polarized in how they form beliefs about politics, they will become polarized on political grounds as well, providing a new explanation for recent political polarization in the United States which, under reasonable assumptions, is more in line with the stylized facts than models with perfect observability.

Keywords: Cultural Distance; Imperfect Communication; Inequality; Polarization; Policy Divergence; Redistribution; Social Groups; Voter Beliefs

*I thank, among others, Laurent Bouton, Horacio Larreguy, Bart Lipman, Michael Munger, Jawwad Noor, and Leeat Yariv for insightful comments and advice. In addition, I am grateful for the useful discussion of audiences at the meetings of the Association for Public Economic Theory, the Southern Economic Association, the Eastern Economic Association, and the Public Choice Society, the Texas A&M University Political Science Department, and participants in graduate workshops at Boston University, Concordia University, New York University, and Princeton University. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 637662). In addition, I gratefully acknowledge the Boston University Summer Research Grant, the Charles Huse Fellowship, and the Humane Studies Fellowship for providing funding for my graduate studies, which enabled this project.

[†]Texas A&M University, Department of Political Science & Université libre de Bruxelles, European Centre for Advanced Research in Economics and Statistics. Contact: bgogden@tamu.edu

1 Introduction

It has long been understood by social scientists that voters do not have full information concerning the policy platforms of parties. Voters are unable to recall the most basic of political facts (Cambell et al. (1960); Kinder and Sears (1985); Carpini and Keeter (1996)), misunderstand statements made by politicians (Mason (2015)), make voting decisions based upon irrelevant information (Leigh (2009); Wolfers (2009); Healy, Malhotra and Mo (2010); Huber, Hill and Lenz (2012); Achen and Bartels (2013)), and base their beliefs on a slew of private conversations and biased opinions (Levy and Razin (2015); Ortoleva and Snowberg (2015)). What’s worse, they often have little incentive to improve the knowledge that they do possess, i.e. they are “rationally ignorant”. As noted by Lupia and McCubbins (1998), “Ironically, for many political issues, information is not scarce; rather it is the cognitive resources that a person can use to process information that are scarce”. Indeed, recent experimental evidence (Woon (2015); Hill (2016)) shows that even in relatively simple settings voters display out-of-equilibrium beliefs, while Al-Najjar and Shmaya (2015) show that inconsistent beliefs are necessary to model disagreement about the fundamentals of the world. Despite this, our major models of partisan competition assume perfect observability.

Similarly, parties are simultaneously becoming more polarized both along cultural and socioeconomic grounds, as well as in policy (e.g., Brewer and Stonecash (2007); Mason (2015); Gentzkow, Shapiro and Taddy (2016)). For example, the Pew Research Center has shown that white, working class voters, a historically split demographic, have strongly and rapidly sorted into the Republican Party over the past decade, with nearly 60% now identifying as strong Republicans (Pew (2012)). As there is evidence that different socioeconomic and cultural groups have difficulty communicating with each other (e.g., Hymes (1964); Lang (1986); Lazear (1999); Wedeen (2002)), this problem will be amplified by the extension to politics, where voters have little incentive to bridge the gap in understanding between themselves and the candidates or the candidates’ supporters. Therefore, voters will have more precise beliefs over the platforms presented by parties that “speak their language”. If parties know this, they will set their policies accordingly, effectively committing to their own cultural base rather than the general population.

I propose a new voting model, the *imperfect beliefs voting model*, to represent this interaction of beliefs, culture, and communication. In a fashion similar to the prob-

abilistic voting framework¹, my model creates continuity by adding a stochastic element. Unlike probabilistic voting, however, the *beliefs*, rather than preferences, of voters are stochastic; the more culturally distant a voter is from a party, the greater the potential ex-post error in beliefs. Therefore, the results depend on how cultural distance interacts with personal policy preferences.

Simply allowing voters to have imperfect (non-equilibrium) beliefs over exactly what platforms candidates will implement leads to substantively different policy outcomes than the previous models, even without exogenous candidate preferences or lobbying. In particular, this leads to a model with purely election-motivated candidates better matching the empirical evidence with regards to political polarization within the United States without relying upon changes in underlying politician or voter preferences. It does this while maintaining the tractability of the Downsian framework, allowing for positive policy analysis. A preview of the results follow:

When policy preferences are correlated with belief precision, there will be divergence away from the median voter as both parties gain marginally more from appealing to the voters with whom they can communicate more easily. Therefore, as the population becomes polarized on cultural grounds, they will become polarized on political grounds as well.

If policy preferences are uncorrelated with cultural distance, and hence precision of beliefs, we *may* retrieve the classic Median Voter Theorem. Even here, however, the Median Voter Theorem may fall apart if one politician is viewed with more precision by a larger portion of the population and there is some skew in the distribution of voters' preferences (for example, due to the log-normal distribution of income). In this case, the party which is viewed with less precision will have an incentive to appeal to outliers in the distribution who are less likely to flip their preferences due to small amounts of noise. In this way, non-incumbent, "outsider" parties will appeal to more extreme elements of the population, as seen in, e.g., Fiorina (1973) and Ansolabehere, Snyder and Stewart (2001). This result shows the power of imperfect beliefs beyond the results of other models of party bases, such as, e.g., Glaeser et al. (2005) and Virag (2008).

Taken together, this provides an explanation for the party polarization in the United States that has occurred over the last half-century. While there is little evidence that the nation at large is becoming more polarized on political values, politicians currently are at their most polarized when it comes to voting behavior (e.g. Ozdemir and Ozkes (2014)). Figure 1, recreated from McCarty, Poole and Rosenthal

¹e.g., Enelow and Hinich (1982), Ledyard (1984), and Lindbeck and Weibull (1987)

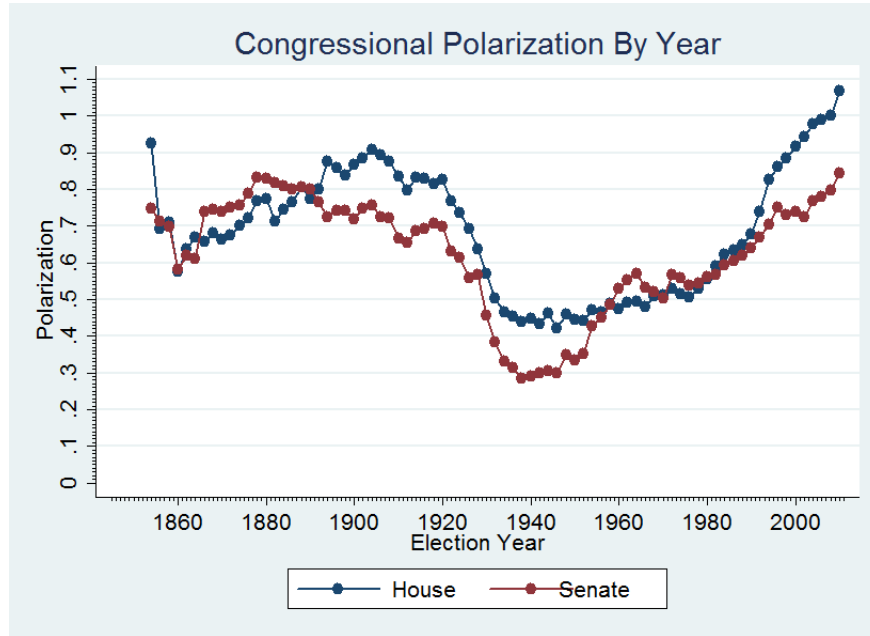


Figure 1:

(2006), maps the difference between the average DW Nominate scores, a measure of partisan voting behavior, for party politicians in the U.S. House and Senate (Carroll et al. (2013)). A sizable literature has emerged, reviewed by Fiorina, Abrams and Pope (2005) and McCarty et al. (2006), questioning several potential explanations, such as a polarizing electorate, changes in electoral institutions, gerrymandering, and the rise of primary systems within the parties. At the least, it appears that there is still a substantial residual even controlling for these institutional changes: Fowler and Hall (2016) find that parties have diverged as much on salient issues as non-salient issues, while Brunell, Grofman and Merrill (2015) find that 80% of the increase in polarization must be explained by within-district polarization between candidates rather than increased partisan bias or sorting of voters. The model provides endogenous results consistent with the time-series evidence, showing that politicians have polarized as voters have begun sorting based on socioeconomic characteristics and as inequality has grown.² It implies that cultural segregation and sorting, such as the selective move into the suburbs and assortative mating, have impacted positive policy outcomes.

In particular, one difference between the imperfect beliefs voting model and Downsian/probabilistic voting models comes via the role of identity. Probabilistic voting

²While the data on polarization and housing segregation referenced within this paper rely on time series at a national level, the results of, e.g., Brunell, Grofman and Merrill (2015), show that the increase in congressional polarization must be related in a strong way to within-district polarization, and hence a within-election explanation is necessary.

models predict that voters whose sociocultural identity is tied to a party or candidate will actually be relatively ignored in terms of policy outcomes. For example, this leads to the oft-repeated prediction that African American priorities are ignored by both parties as they vote overwhelmingly for the Democrats. However, this prediction relies on the assumption that such a preference is based on valence and not tied to the policies supported by the candidates. If, instead, the role of identity is to increase the perceptiveness of the voter with respect to a particular candidate, we should expect that they will be better represented by that candidate and less represented by the opposition. So, instead of being equally ignored by both parties, the model predicts that African Americans should be best represented by those candidates (usually Democrats) who are culturally tied into their community, while being less represented by Republicans. In equilibrium, this will still lead to the observed vote shares, but does not require an exogenous assumption of valence.

Finally, I apply the model to the classic Meltzer and Richard (1981) model of political redistribution. The implications of the model in this context are two-fold. First, we should be focusing on the interaction between the rich and the poor and the ability of the two groups to communicate ideas of redistribution to each other; it is this sociological relationship which will have a greater impact than inequality itself. Second, under the realistic assumption that incomes are distributed log-normally, the two parties will diverge, leading one to offer inefficiently high levels of government spending, and the other to potentially offer inefficiently low spending. The model predicts that divergence will increase between a small government and a big government party, with an ex-ante ambiguous impact on the ex-post fiscal policy implemented. Moreover, if the two parties tend to be divided by income, as is seen in most nations, this divergence will be even greater.

1.1 Literature

The increase in political polarization within the United States over the past half-century remains difficult to explain within a general, tractable formal framework. Voting models of non-ideological, strategic candidates, following Downs (1957), tend to predict convergence, while divergence required ideological candidates with some form of commitment problem.³ For example, Glaeser, Ponzetto and Shapiro (2005) showed that policy divergence could be acquired if candidates targeted their messages to different social groups, encouraging them to turnout. Similarly, Feddersen and

³For an overview of the results with perfect observability, see Persson and Tabellini (2002).

Gul (2014) and Krasa and Polborn (2014) recently have shown that polarization can be driven by rising inequality, but they relied on both ideological candidates and a need for donations to garner this result. However, the imperfect beliefs voting model shows that it is also possible to derive divergence and potentially explain polarization by focusing on voter uncertainty over policy, a phenomenon that has been observed throughout the political sphere.

Previous literature on uncertainty has focused on imperfect observation of candidate quality, defined as ability to deliver public services efficiently, rather than uncertainty over the state of the platform itself (e.g. Coate (2004); Carrillo and Castanheira (2008); Frenkel (2014)). For example, Carrillo and Castanheira (2008) find that, with asymmetric information over candidate investment in quality, candidates may have an incentive to diverge from the median voter in order to signal their quality. In contrast, the model shows that divergence can be driven without a separate voter preference for candidate quality, i.e. only ideological voters.

There are a number of major papers that deserve attention for doing work on *platform* uncertainty (e.g., Berger, Munger and Potthoff (2000); Alesina and Holden (2008); Virag (2008); Gul and Pesendorfer (2009); Acemoglu, Egorov and Sonin (2012); Agranov (2012); Ashworth and de Mesquita (2014); Matejka and Tabellini (2016)). Most of this work has focused on cases with candidate bias or restricted choices, while Virag (2008), Matejka and Tabellini (2016), and Agranov (2012) are similar to the model in that they feature purely Downsian candidates, though the former two can only achieve a special case of culturally differentiated preferences and the latter requires the context of a primary campaign. The model generalizes these results by relaxing the assumption of perfect Bayesian beliefs, obtaining divergence and realistic dynamics even with purely Downsian candidates and no correlation between observability and preferences within a general election. This allows the model to feature testable implications concerning the dynamics of polarization and divergence discussed above without making assumptions about exogenous changes to candidate preferences or electoral law.

The paper also adds to a growing literature on *behavioral political economy* (e.g., Ashworth and de Mesquita (2014); Matejka and Tabellini (2016); Klingelhöfer (2017)) which shows that strategic candidates behave in starkly different ways in light of boundedly rational voters than they would in the face of perfectly rational voters. This is an essential inquiry due to the long literature in political behavior showing just such imperfections by voters. The model adds to this literature by providing a general framework for analyzing policy formation in light of a very realistic and

commonly observed bound upon standard behavior.

2 The Model

2.1 Parties

Following standard convention, I model an election consisting of two parties, 1 and 2, competing for the votes of a set of citizens of measure one. The parties only care about the rents that can be obtained from holding office, i.e. they have no personal policy beliefs.⁴ Once elected, a victorious party implements the policy platform to which he committed himself during the campaign. For simplicity, let us consider policies $P \in [a, b] \subset \mathbb{R}$. This standard framework allows for increased tractability and comparison to other models of political competition.

I assume that the parties will pick their policy platform P_k to maximize their expected vote share, π for party 1; $1 - \pi$ for party 2.⁵

I allow for full commitment on the part of candidates, such that candidates will make their intentions known in the form of public platforms.⁶ These platforms will be chosen simultaneously by the two candidates.

2.2 Voters

Every voter possesses an ideal policy point, $x_i \in [a, b] \subset \mathbb{R}$, such that they prefer a policy closer to their ideal, with $x_i \sim \Sigma(\cdot)$, where Σ is a continuous, differentiable CDF. Therefore, the voter will vote for the party that he *believes* will maximize his utility:

⁴One can alternatively think of strategic parties choosing politicians (who do possess personal policy preferences) in order to maximize electoral success.

⁵While it may be of interest to consider maximizing the probability of winning instead a la Lindbeck and Weibull (1987), this exercise proves to be mathematically intractable due to the structure of the model, while not appearing to provide any additional illumination of candidate behavior since the asymmetry in candidate incentives holds in either case. For example, consider the probability of winning as an increasing function $\Psi(\pi)$, where π is the expected vote share. In this case, the relevant equilibrium condition for candidate 1 is $\frac{\partial \Psi}{\partial \pi} \frac{\partial \pi(P_1, P_2)}{\partial P_1} = 0 \propto \frac{\partial \pi(P_1, P_2)}{\partial P_1} = 0$. See Persson and Tabellini (2002) for further defenses of this assumption.

⁶While it is certainly true that commitment problems and candidate policy biases exist, their role in generating divergent behavior has been explored elsewhere in the literature; I seek to identify a new cause of divergence which has not otherwise been identified: voter beliefs and cultural distance. Indeed, adding limited commitment to the model will simply add to the incentives for divergence between the parties unless we assume that the parties share an ideal point over policy.

$$(1) \quad U_i(\hat{P}_k^i) = U(|\hat{P}_k^i - x_i|)$$

where \hat{P}_k^i is voter i 's subjective belief about the policy platform of party k and the utility function is single-peaked and symmetric about the ideal point. These voters will vote sincerely based upon their beliefs.

The driving assumption of the imperfect beliefs voting model is that these beliefs must only be consistent with a subjective belief \hat{P}_k^i , and that this belief depends in some way upon the true policy platform P_k . This provides the model with flexibility, allowing the social scientist to match voter behavior to the data.

Let $\delta_k^i \in \{0, \delta\}$ be a measure of sociocultural distance between the party and voter i , where $\delta > 0$, such that the variance of \hat{P}_k^i is increasing in δ_k^i .

For tractability, we separate the continuum of voters into groups based upon their sociocultural distance from the candidates:

Definition 1: Voter i is within the **cultural base** of party k if $\delta_i = 0$.

Definition 2: Voter i has **type** $\theta_i \in \{1, 2\}$ such that if $\theta_i = k$, i is within the cultural base of party k . Let μ be the probability $\theta_i = 1$.

Through various social networks and individualistic information sources, voters will always have a relatively better idea of the platform for one party compared to the other. This focuses emphasis on the role of intelligibility and social sorting in politics (e.g. Wedeen (2002); Mason (2015)). Therefore, a portion μ of the population is within the base of party 1, will have more precise beliefs over that candidate, and will correctly assess their utility from supporting that candidate, with the rest within the base of party 2, featuring the potential for imperfect signals. The starkness of this assumption allows for tractability in analysis of the model, but is not substantive, as adding voters within the base of both parties limits to the standard Downsian model. Therefore, as shown in Appendix B, the key assumption is simply the asymmetry such that some voters have a smaller propensity to have ex-post errors about a particular candidate.

Since the base represents those with whom a party can better relate their platform, the social network and communication asymmetries that divide the voting population into bases may also divide the population by policy preferences. Therefore, let $G(x_i) = \Sigma(x_i|\theta_i = 1)$ and $H(x_i) = \Sigma(x_i|\theta_i = 2)$, where G and H are single-peaked, continuous, and differentiable. Therefore, the policy preferences x_i for any voter i in the base of party 1 will be distributed according to G , and similarly for 2 and H . Larger

differences between elements of G and H represent greater correlation between the social-cultural and policy differences within the population.⁷

For notational convenience, let G_{med} be the median voter of G , H_{med} be the median voter of H , and x_{med} be the median voter of the general population, with similar definitions \overline{G} , \overline{H} , and \overline{x} for the means.

2.3 Equilibrium

I will analyze subgame perfect equilibria of the game between the parties described above. Specifically, I define the game and equilibrium of interest as follows:

Definition 3: An *Imperfect Beliefs Voting Game*, Π , consists of the tuple $\{G, H, \mu, \delta, \{\hat{P}_1^i, \hat{P}_2^i\}_i\}$.

Definition 4: A *political equilibrium* for Π consists of a pair of policy platforms (P_1, P_2) such that:

1. $P_1 \in \argmax \pi$
2. $P_2 \in \argmax (1-\pi)$
3. Each voter i votes for party 1 if $U_i(\hat{P}_1^i) > U_i(\hat{P}_2^i)$, party 2 if $U_i(\hat{P}_1^i) < U_i(\hat{P}_2^i)$, and votes for each with probability $\frac{1}{2}$ otherwise.
4. If voter i is within the base of party 1, $\hat{P}_1^i = P_1$.

Note that for the voters, we relax the standard Nash assumption that there must be consistency between a player's beliefs and the outcome; here, we simply require that this relationship exist for those politicians in whose base a given voter resides.⁸

⁷The "bases", however, may share common support. Therefore, they capture both party bases in the traditional sense, as well as opposition voters who may pay more attention to holding the other side accountable than to their own candidates. For example, in the 2016 election, some voters on the left may have been more aware of Donald Trump's policies vis a vis immigration, criminal policy, etc. than Hillary Clinton's particular history on these same issues. These voters would be part of Trump's base rather than Clinton's. The model simply assumes that they will be of smaller measure than the number of voters who have more information about the proposals on their own side.

⁸Note that if voters were fully rational in a Nash sense (e.g. with common knowledge of the game being played between candidates and voters), a voter in the base of party 1 would treat any communication from party 2 as cheap talk (as in Glaeser, Ponzetto and Shapiro (2005)). Therefore, the incentive discussed in Proposition 1 would be enhanced as parties would have no incentive to appeal to voters who did not have precise beliefs over their platform. The equilibrium of such a game would trivially be for each party to diverge to the median ideal point of their base. While this does not appear to describe actual voter behavior, it does help highlight the incentive structures that exist with Culturally Differentiated preferences.

2.4 Additively Symmetric Beliefs

For the remainder of the paper, however, I will limit the analysis to a special form of beliefs such that voter beliefs will be correct *in expectation*:

Definition 5: Beliefs are **additively symmetric** if and only if $\hat{P}_k^i = P_k + \delta_k^i \epsilon_i$.

where ϵ_i is an i.i.d. random variable representing errors in voter belief formation. ϵ_i will be drawn from a distribution with symmetric CDF F with mean 0, standard deviation σ^2 , and full support. In addition, let f be the well-behaved PDF of F , which is strictly decreasing in $|x|$, with $\lim_{|x| \rightarrow \infty} f(x) = 0$; hence, larger errors are less likely than smaller errors. Here, the key is that the precision of voter beliefs with respect to party platforms is diminishing in their cultural distance from a party and the bulk of the population has reasonable, albeit potentially wrong, beliefs about the platforms, though radically wrong beliefs are still permitted. This symmetry is consistent with empirical (e.g., Bartels (1996); Cabral and Hoxby (2012)) and experimental (e.g., Hill (2016)) evidence that individual voter “mistakes” average to zero in the aggregate.⁹

To see how such a reduced form may be generated, note that these beliefs are isomorphic to scenarios with overconfident voters (e.g. Ortoleva and Snowberg (2015)). Suppose that voters begin with some common prior distribution Z_k over party k ’s policy and receive noisy signals $\hat{P}_k^i = P_k + \delta_k^i \epsilon_i$. Now suppose that they are overconfident and believe that $\delta_k^i = 0 \forall k$; in other words, they are overconfident in their ability to interpret their platform. This would generate exactly the ex-post beliefs above, as voters believe they have received a precise signal when, in fact, they are receiving a noisier signal about one candidate. Such overconfidence is consistent with a model in the vein of Feddersen and Pesendorfer (1996) in which only voters who are confident in their political information turnout, and is also consistent with the *American National Election Studies* (2014) data which show that conditional on turning out, voters are very confident in their political beliefs.

Therefore, given platform positions P_1 and P_2 , the parties will know that the expected vote share of party 1, π , is

$$(2) \quad \pi = \mu \left[1 + \int_a^{P_1} \left[F\left(\frac{2x - P_1 - P_2}{\delta}\right) - F\left(\frac{P_1 - P_2}{\delta}\right) \right] dG(x) + \int_{P_1}^b \left[F\left(\frac{P_1 - P_2}{\delta}\right) - F\left(\frac{2x - P_1 - P_2}{\delta}\right) \right] dG(x) \right] \\ + (1 - \mu) \left[\int_c^{P_2} \left[F\left(\frac{P_2 - P_1}{\delta}\right) - F\left(\frac{2x - P_1 - P_2}{\delta}\right) \right] dH(x) + \int_{P_2}^d \left[F\left(\frac{2x - P_1 - P_2}{\delta}\right) - F\left(\frac{P_2 - P_1}{\delta}\right) \right] dH(x) \right]$$

⁹Systemic biases would simply increase the incentives for politicians to diverge away from the median voter’s preferences, as seen in previous work.

3 Political Equilibria Analysis

Within the classical probabilistic voting framework, both parties feature identical incentives. This symmetry in marginal voters is what leads to the convergence results which underlie positive policy analysis within such models. The probabilistic element enters as a linear shift in voter preferences, maintaining marginal utilities over policy platforms. Therefore, swing voters are equally appealing to both parties.

In contrast, with imperfect beliefs the voters do not change the marginal likelihood of their vote by the same amount. In this case, the probabilistic element is *within* the utility function rather than an external shift to preferences. This is what differentiates the imperfect beliefs voting model from a probabilistic voting model in which the parties simply have asymmetric beliefs over each group's swingability.

Equation 1 for any voter i can be rewritten in terms of the error ϵ :

Property 1:

1. A voter i of type 1 will vote for 1 iff
$$\begin{cases} \epsilon_i > \frac{P_1 - P_2}{\delta} \text{ or } \epsilon_i < \frac{2x_i - P_1 - P_2}{\delta} & \text{if } x_i < P_1 \\ \epsilon_i > \frac{2x_i - P_1 - P_2}{\delta} \text{ or } \epsilon_i < \frac{P_1 - P_2}{\delta} & \text{if } x_i > P_1 \end{cases}$$
2. Voter i of type 2 will vote for 1 iff
$$\begin{cases} \epsilon_i \in [\frac{2x_i - P_1 - P_2}{\delta}, \frac{P_2 - P_1}{\delta}] & \text{if } x_i < P_2 \\ \epsilon_i \in [\frac{P_2 - P_1}{\delta}, \frac{2x_i - P_1 - P_2}{\delta}] & \text{if } x_i > P_2 \end{cases}$$

Property 1 highlights the crucial incentives created by imperfect voter beliefs. In order to consider the swingability of each voter, it is useful to think in terms of their voting responsiveness with respect to voting for each candidate:

Definition 6: Voter i 's **responsiveness** with respect to P_k is the marginal change in the probability of supporting candidate k given a change in P_k (i.e. $\frac{\partial \text{Prob}[i \text{ votes } k]}{\partial P_k}$).

Proposition 1: Consider two voters, i and j , who differ only in that voter i is within the base of party 1 and j is within the base of party 2. The following is true:

1. The responsiveness of voter i is greater than the responsiveness of voter j with respect to P_1 , and vice-versa for P_2 .
2. If $P_1 > (<)x_i$ and $P_2 > (<)x_i$, then the differences in responsiveness are diminishing in $|P_2 - x_i|$.

As in standard Downsian models, it remains true that a party can increase its probability of winning a particular voter by moving in her general direction. If the

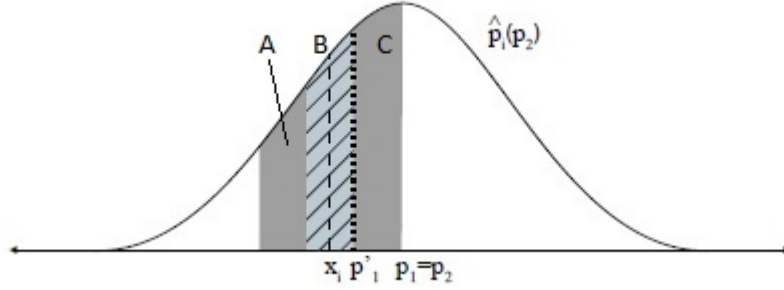


Figure 2:

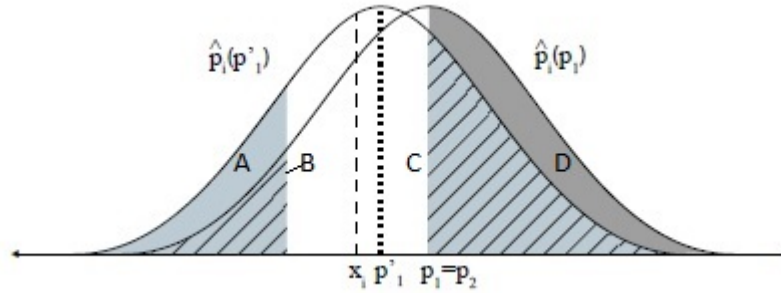


Figure 3:

voters possess imperfect beliefs, however, Proposition 1 tells us that there is now an asymmetry in the *relative* gains to be had to parties from given voters. To see the intuition, consider the extreme case in which a voter in the base of party 1 will have a fixed distribution of beliefs over party 2 independent of P_2 , and vice-versa for those within the base of party 2. In this case, voters in one party's base will only be affected by the policy of that party, and will be perfectly inelastic with respect to the other party's choice of platform.

For an example, consider figures 2 and 3. In Figure 2, the voter represented by ideal point x_i is within the base of party 1. Suppose we start at a position where both parties represent the same platform, shown here at $P_1 = P_2$. In this case, the ex-ante distribution of voter i 's beliefs with respect to candidate 2's platform are represented by $\hat{P}_2(P_2)$. Therefore, they will vote for party 2 if and only if their beliefs fall in the shaded area (i.e., $A+B+C$). Now consider party 1 moving in the direction of i 's ideal point to P'_1 . A move in her direction will decrease the chance she votes for party 2 (shrinking the area to B) by increasing the possibility that she thinks party 2 is "too extreme" relative to party 1. This is done along two dimensions: it both decreases the probability that she thinks that party 2 is to the left of party 1 (removing area C), while also increasing the probability, conditional on thinking that party 2 is to the left, that she thinks that party 2 is too far to her left (area A). As evidenced in figure

2, the latter possibility is less likely, but not impossible ($A < C$).

By contrast, if the voter is in the base of party 2, the distribution of errors for i at the point of convergence such that he would vote for 1 is the complement of that when he's in 1's base ($B+C+D$ in figure 3). Therefore, the effect of party 1 moving in her direction is not so simple: changes in the platform shift the distribution, not the boundaries. It is still true that the platform change will decrease the probability that party 1 will be viewed as to the right of party 2 (removing area D); however, it will also increase the probability that party 1 will be viewed as too far to the left (area A will now result in a vote for 2). Therefore, the voter is less responsive to party 1's platform change than she would have been if she had perfect beliefs as she may "overreact" to a shift to the left by thinking that party 1 is becoming too radical relative to her (perfect) belief of party 2.

From here forward, I will take one functional form assumption in order to allow us to talk about a "left" party and a "right" party:

Assumption 2: G and H have quasiconcave PDFs, with $G_{med} \leq H_{med}$.

Definition 7: Parties are said to be **Culturally Differentiated** if G is strictly first-order stochastically dominated by H .

Definition 8: Parties are said to be **Culturally Independent** if $G = H = E$ everywhere.

Another way to think of Culturally Differentiated preferences is that cultural distance δ_i is correlated with policy ideal points x_i . In these cases, we can differentiate ideal policies by group. Hence, Culturally Independent preferences are those where x_i is independent of δ_i . These two cases allow us to explore the most common ways in which communication and culture interact with policy preferences.

3.1 Culturally Differentiated Preferences

In much of the world, parties are increasingly looking very different from each other in terms of demographic attributes such as income, marriage rates, ethnic background, etc. Evidence has increasingly shown that such differences have an impact on the policies implemented (e.g., Levitt (1996)). Indeed, the language of the two parties is increasingly becoming more differentiated (e.g., Gentzkow, Shapiro and Taddy (2016)). Therefore, in a world of increasing social-cultural polarization, it makes sense to begin by allowing for correlation between policy preferences and sociocultural dis-

tance.¹⁰

As discussed above, the two parties may have a comparative advantage in certain classes of voters depending on the distribution of beliefs and preferences. This leads to the following major result:

Theorem 1: *If there are culturally differentiated preferences, then there exists a pure-strategy equilibrium.*

In particular, with $\mu = \frac{1}{2}$ and $G(x_{med} - \epsilon) = 1 - H(x_{med} + \epsilon)$ for any $\epsilon \geq 0$, then the unique equilibrium involves both parties place the same distance between x_{med} and the median of their base, and both win the same share of the vote.

The level of divergence, $|P_1 - P_2|$, is greater given greater cultural distance between groups, δ .

By Theorem 1, the Median Voter Theorem is generally violated. Even with a single-peaked, symmetric distribution of voter preferences and two parties with identically-sized bases, the inability to perfectly communicate with members of the opposing party's base causes a form of policy "risk aversion" on the part of candidates as described in Proposition 1. Therefore, each party will tend more towards those who can observe them perfectly, knowing that in expectation they are more likely to "get it right" and recognize that the party matches their preferences. If either party was to try to move towards the median of the general population, they risk losing more voters from their own base who hold incorrect ex-post beliefs over the other party than they are likely to gain by trying to capture the other party's voters. In general, the parties will locate at just the right position such that they will, in expectation, capture the voters within their own base while still remaining in play for moderates of the other side.

To elucidate why convergence cannot be an equilibrium with asymmetric bases, consider Proposition 1. Suppose the two parties were to converge to the platform of the median voter. With perfect beliefs, this would be optimal for both sides as they would simultaneously lose as many voters as they gain by changing their platform. However, with imperfect beliefs, there is an asymmetry in the way that such probabilistic gains and losses would occur. Suppose party 1 was to move to the left. They would lose support in expectation from the members of party 2's base. However, this

¹⁰Consider, for example, the case of those who support a stronger social safety net and consider it an issue of "liberty" and "rights", but have difficulty expressing that to those who view such concepts through a predominately negative lens. In such a world, politicians will have to take into consideration how they can best word their platform to appeal to everyone, while simultaneously being aware that their base are the ones that will be most responsive to the positions that they take. This is because their base "speaks their language".

loss is only of a second order, as the *range* of errors such that a base voter of party 2 will vote for candidate 1 remains the same; the loss in expectation only comes from the fact that large errors are less likely than small errors. In fact, for moderate members of 2's base, the party actually reduces the probability of said voters thinking they are too extreme.

By contrast, for much of party 1's base (those that will remain to the left of his new policy profile), the errors will not only now need to be smaller, and therefore more likely, but the range of sufficient errors such that 1 wins is now larger in absolute terms. Therefore, while these ideologically committed voters still know that the party is "too moderate", there will now be a larger range over which they believe that the other party is either even more moderate or too extreme. In equilibrium, by Proposition 1.2, the two parties will diverge enough that the probabilities of such ex-post errors are so low that the candidate will have nothing to gain from further pandering to his base.

Note that in equilibrium it is unlikely that the voter will ever believe that party 1 is to the right of party 2. Therefore, part of the equilibrium dynamics which occur will limit the real impact of the voters' imperfect beliefs. While, in reality, few believe that the Republican party is to the left of the Democratic party, this is an equilibrium outcome.¹¹ If the two parties were actually to converge to a symmetric policy, there would be nothing unreasonable about believing any directional relationship between the two.

By way of an example, imagine that $\Sigma \sim U[-1, 1]$, with $G \sim U[-1, 0]$ and $H \sim U[0, 1]$, $\delta = 1$, and F a standard normal distribution. In this example, with $\mu = \frac{1}{2}$, $P_1 = -.32$ and $P_2 = .32$. Therefore, the population is divided almost into thirds: one-third of the population is to the left of candidate 1, one-third is to the right of candidate 2, and one-third is in between both candidates. Note, in addition, that the probability of casting a mistaken vote is very small: a voter only with $x_i = \pm\epsilon$ will only cast a mistaken vote with a probability of about .25, while a voter located at the extreme of the distribution ($x_i = \pm 1$) will make a mistake with a probability of about .17. Despite these relatively small equilibrium rates of mistakes, it is the off-path possibility that leads to substantial divergence on the part of the parties.

¹¹Though, as ANES data reveal, there do still exist some individuals who do perceive such a relationship between the two parties.

3.2 Culturally Independent Preferences

Culturally Independent preferences are a useful limiting case which can be used to examine situations in which the correlation between voter ideal points and political belief formation are weak.

With the assumption of Culturally Independent preferences, we *may* reacquire the Median Voter Theorem:

Theorem 2: *Let $\mu \geq \frac{1}{2}$ and $\bar{x} \geq x_{med}$. If there are Culturally Independent preferences, there exists a unique equilibrium such that $P_1 \leq x_{med} \leq P_2$ and $|P_1 - x_{med}| \leq |P_2 - x_{med}|$, with 1 winning a greater expected vote share.*

All are equalities if and only if Σ is symmetric or $\mu = \frac{1}{2}$.

Note that the conditions of the theorem are without loss of generality. Even with Culturally Independent preferences, the Median Voter Theorem is only a special case with imperfect beliefs. Convergence is an equilibrium if and only if both parties are viewed with precision by equal portions of the population or if the distribution of voter preferences is symmetric. Otherwise, a candidate who is viewed with greater precision will have the advantage in winning more “mainstream” voters, while those with weaker voter perceptions will specialize in appealing to outliers. This can generate policy divergence even without differentiated party bases.

Why does the Median Voter Theorem not hold without the above conditions? Suppose the two parties place their platform at the median voter’s ideal point. By Proposition 1, voters will make their decision more on the basis of party 1’s policy choice than party 2’s. However, this asymmetry is minimal for those outlier voters to the extreme right, as they are unlikely to ever think any candidate is too extreme. In this case, the party who is viewed with precision by a smaller portion of the population could move away from the mass of the distribution (i.e. towards the skew) since they gain relatively more from the outliers whom are less likely to make ex-post errors. Inversely, the candidate who is viewed with greater precision will have an incentive to move towards the mass of the distribution as trying to appeal to voters in the skew would be too costly as they are viewed with such precision.

One implication of Theorem 2 is that it provides an electorally-oriented explanation for the move by incumbents towards the majoritarian position (i.e. the mass of the distribution), while challengers find themselves appealing to outliers. Presidents Reagan, Clinton, Bush, and Obama were all perceived as running more moderate campaigns on their re-election tries than when they initially entered the White House as representatives of the outsider party (*American National Election Studies* (2014)).

If we consider that an incumbent is more likely to actually implement the policies proposed *prior* to an election, and will receive more broad and free coverage from the media, it is reasonable to believe that such a candidate will be perceived with less error. Therefore, as a candidate, the incumbent will have more to lose from appealing to extremists relative to the masses. By contrast, any challenger to an incumbent will find it harder to challenge the incumbent directly with said masses, but will instead find the need to appeal directly to those who are highly unlikely to be pleased with the majoritarian policies of an incumbent, i.e. the outliers. This incentive structure is similar to models of valence which gives one candidate a discrete advantage (e.g., Groseclose (2001)) or exogenous reputation (e.g., Bernhardt and Ingberman (1985)), but does not rely on the presumption that voters have an inherent preference for the incumbent; instead, this equilibrium advantage is derived endogenously.¹²

While Theorem 1 is reminiscent of the major results in Glaeser, Ponzetto and Shapiro (2005), these results do not rely on endogenous turnout, nor on activists who dismiss all signals from their out party as cheap talk. In addition, Theorem 2 shows that imperfect beliefs can lead to divergence even when the policy preferences of the two base groups are identical. Therefore, the imperfect beliefs voting model shows that divergence and polarization are a more fundamental function of plurality elections whenever voters may have imprecise beliefs over policy. In addition, the model can internalize the mechanism of targeting base groups while also retaining the tractability of the probabilistic voting model for applications, as in section 5, and while generating polarization even without differentiated activist groups.

By way of an example, consider as above $\delta = 1$ and $\epsilon \sim N[0, 1]$. In this case, however, let Σ be distributed according to an asymmetric triangle distribution, peaked at $\frac{1}{3}$, and let $\mu = 1$, such that $\theta_i = 1$ for all voters. In this case, $P_1 = \frac{1}{3}$ and $P_2 = 1$. This means that the “minority” party may actually take a position to the right of the most extreme voter within the distribution.

¹²In addition, such models require additional assumptions to generate the existence of a pure strategy equilibrium. On (mixed-strategy) equilibrium path, Groseclose (2001) does not generate this incentive structure. By contrast, Theorem 2 shows us that the model maintains this structure with only limited assumptions.

4 Cultural Groups, Imperfect Beliefs, and Polarization

As shown above, political polarization does not necessarily matter by itself, since politicians will still converge to the median as long as the precision of beliefs is not tied into their preferences or distributed asymmetrically. Candidates only become more radical when sociocultural distance increases and voters become more likely to have imperfect beliefs concerning certain candidates. This provides a new avenue for empirical exploration, as social sorting may be driving recent ideological divergence within the parties. In particular, the stark results of Gentzkow, Shapiro and Taddy (2016), finding that parties increasingly use particularistic, polarized language which is only intended to reach base voters, should indicate a potential social mechanism for political polarization.¹³

The model rationalizes a set of stylized facts concerning the rise of polarization noted in McCarty, Poole and Rosenthal (2006): not only has polarization risen even when controlling for the demographic characteristics of the electorate and the candidate, but said characteristics remain statistically and economically significant. In fact, as polarization has risen, the importance of voter demographics for policy outcomes has risen as well! These facts are difficulties for models which rely on the increasing homogeneity of parties alone to explain polarization, as well as for those models which credit institutional changes. By contrast, the model is able to explain this problem: as groups self-segregate, the accuracy of their beliefs become more asymmetric with respect to candidates of the different parties, driving polarization; however, in any given district, the median of the voting population, around which both parties diverge, will become more dependent on the differentiated underlying characteristics of the district. Meanwhile, candidate demographics remain important, as their cultural base will depend on those with whom they can more easily communicate, which is likely dependent on their socioeconomic descriptors.

The imperfect beliefs voting model offers an alternative, consistent explanation

¹³Recall Figure 1. The relative shock to polarization over the last half-century has occurred simultaneously with a move to the suburbs which has segregated the relatively wealthy from the relatively poor. This would be expected to reduce the social interactions between the groups that are likely to have different policy preferences, increasing the difficulty of communication and the passive acquisition of information across groups. In addition, as discussed above, this is the same period as we have seen the two parties diverge on a number of socioeconomic indicators. Therefore, it seems much more likely that we are in a world of divergent bases than one in which communication is independent of preferences. Indeed, as the model predicts, we have seen massive and continuous policy divergence between the two parties as this social segregation occurs.

for outcomes that had previously been credited as evidence to the citizen-candidate model (e.g. Pande (2003)). For example, Chattopadhyay and Duflo (2004) observe that quotas requiring that some districts only offer female candidates in India have led to public goods being offered which are preferred by women. While this has often been cited in support of citizen-candidate, the imperfect beliefs voting model shows how such results could be driven by the fact that now only candidates who can communicate perfectly with female voters are being nominated, while before candidates tended to be able to communicate more easily with men. Such an alternative result is consistent with Stadelmann, Portmann and Eichenberger (2014), which found that there was generally no difference between male and female politicians except on social and redistributive issues, which are the most likely to be clouded by differences between gendered language. In this way, the imperfect beliefs voting model offers many opportunities to add to the exploration of particularistic public goods provision.

The imperfect beliefs voting model also addresses concerns with primary-based explanations for divergence, such as Alesina and Holden 2008 and Agranov 2012. McCarty, Poole and Rosenthal (2006) show that polarization *decreased* in the United States as political primaries became more widespread, while Woon (2015) found no experimental effect from the presence of primaries. The imperfect beliefs voting model does not rely on such institutional structures to drive its results. The model provides an explanation for the a priori strange result that the rise of primaries in the United States may have actually *reduced* polarization between the parties.

We also consider additional comparative statics under various assumptions on the error distribution in Appendix C.

5 Application: Imperfect Beliefs and Redistribution

One of the most important questions in positive political economy is the relationship between inequality and real economic outcomes. In particular, there has long existed an intuition that the distribution of material wealth within a society will have an impact on the size and role of government within that nation, either via public goods allocations or direct redistribution and/or social insurance. Typically, such analysis is pursued via a standard model of redistributive taxation as popularized by Meltzer and Richard (1981). The model contains a continuum of voters who only differ over their effective wages. They first vote between candidates who campaign over a redistributive program that taxes income and redistributes evenly to the entire population. They then choose how much to work based upon the policy selected by the

winning candidate. This can and has been easily modified to consider the cases of public goods, unemployment insurance, and other measures of general government involvement, but I will limit myself to the straight redistribution case.

The key result in the standard framework is that voters who have effective wages less than the average will prefer some level of redistribution, decreasing in their wage, while voters who are capable of earning wages higher than the average will prefer income subsidies. This derives a set of predicted political outcomes based upon a specific measure of inequality: the spread between the median and mean wages. Since wages and incomes tend to follow a log-normal distribution, the median wage will be less than the mean. A model with perfect observability, therefore, provides us with clear comparative statics:

Downsian with Perfect Observability Results:

1. *An increase in the wage held by the rich, holding the median wage constant, will increase the level of redistribution;*
2. *A decrease in the wage of the poor, holding the median wage constant, will decrease the level of redistribution.*

These results are driven by the fact that the median voter will pick a preferred tax rate based upon how much he expects to gain on net. Since he expects to receive the tax rate multiplied by mean income, while losing the tax rate multiplied by his own income, a larger amount of inequality will increase the mean wage relative to his own and increase his net gains. In contrast, when poverty increases, there is less to gain from a high tax rate, and therefore he will tend to prefer relatively small government.

The empirical data on the role of inequality on the level of redistribution is, however, mixed (see, e.g., de Mello and Tiongson (2006) for an overview); therefore, there is a growing literature attempting to find other sources of interaction between rising inequality and positive political outcomes.

Within the framework of the model, preferences are distributed from $\Sigma \sim [\tau^*_{min}, \tau^*_{max}]$. In addition, citizens sort into neighborhoods, social groups, and parties based upon their income. Such sorting is consistent with evidence presented by McCarty, Poole and Rosenthal (2006) that district level income has only begun to matter for policy outcomes since the 1990s; prior to that, the rich and poor were relatively mixed within the same districts and had more similar social behaviors. Therefore, let us consider a version of the imperfect beliefs voting model applied to this setting:

Redistribution with Imperfect Beliefs: $G \sim [\tau^*_{min}, b \geq \tau^*_{med}]$ and $H \sim [c \leq \tau^*_{med}, \tau^*_{max}]$, with $\bar{x} > x_{med}$.

In a world in which the model holds, candidate 1 is observed better by poorer voters, while candidate 2 is observed more easily by the rich, with potential overlap. Note that this does not require culturally dependent preferences, as $b = \tau^*_{max}$ and $c = \tau^*_{min}$ are allowed. However, it is realistic to assume that there is correlation between income and ability to observe and interpret platforms of different candidates, and such an assumption enhances the following results.

In total, the model provides a set of results that may more realistically match the empirical evidence than the model with perfect observability:

Result: P_1 will offer a larger tax rate than τ^*_{med} , and $P_2 < \tau^*_{med}$.

Let $P_1 = \tau^*_{med} + \epsilon$. Then, $P_2 < \tau^*_{med} - \epsilon$.

Also, the party with the poorer base will win a larger vote share.

This is a direct result of Theorems 1 and 2 and Proposition 2, which can be found in Appendix C. The candidate of the rich will have an incentive to offer less redistribution in order to limit the chance that moderately wealthy voters will vote for the other candidate, and vice-versa for the candidate of the poor. Therefore, ex ante, we could see either a small government candidate or a big government candidate win the election, though the candidate in favor of more redistribution will be more likely to win.

Due to the skewed nature of the income distribution, the difference between the wage of the richest and the wage of the median will be larger than the difference between that of the median and the poorest. By extension, the rich voters will prefer even smaller government relative to the median voter than the poorer voter will prefer large government. Therefore, the candidate of the rich has a more extreme base, ranging from those who prefer a mildly large government to those who prefer a radically small government that incentivizes work. Therefore, in order to maximize his chances of holding onto the richest members of his base, the rich candidate must be more extreme than his poor counterpart.

As in the Downsian model, however, the cause of the expected increase in poverty will impact the expected outcome on policy. The difference in the model is that the change will depend on the relative probabilities of the two candidates winning rather than on equilibrium policy. As noted above, any increase in inequality will cause further divergence in policy platforms; however, property 2 tells us that the cause of the change in inequality will determine the changes in the vote share:

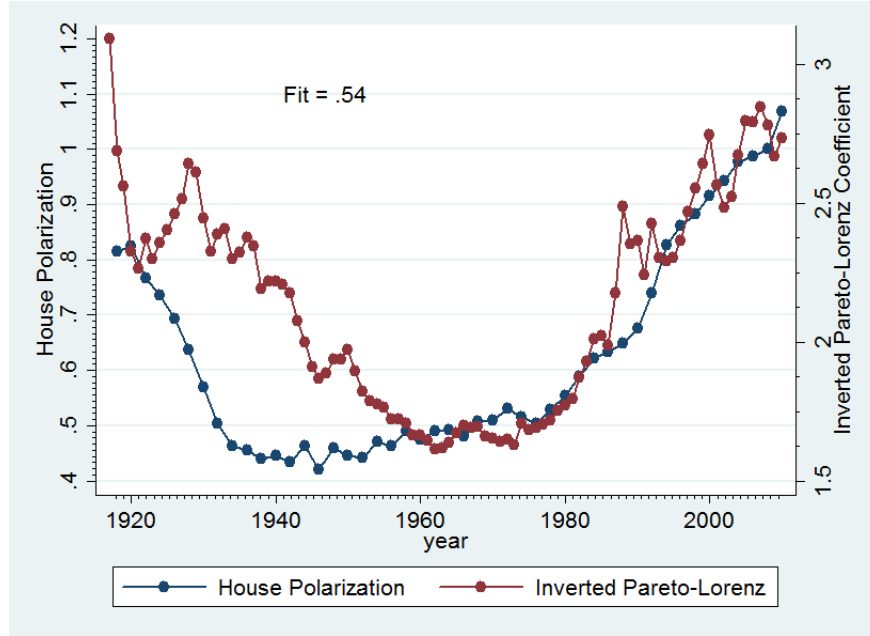


Figure 4:

Result: *An increase in the wage of the rich, holding median wages constant, will increase τ_1 and π , and will decrease τ_2 .*

A decrease in the wage of the poor, holding median wages constant, will increase τ_1 , and will decrease τ_2 and π .

Consider an increase in inequality driven by a rise in wealth at the top (an increase in τ^*_{max} in this framework). In this case, the spread of the income distribution within the base of the upper-class candidate becomes even wider relative to the base of the lower-class candidate. Therefore, π rises and the big government candidate becomes more likely to win. We have two reasons to expect to observe bigger government here: the big government candidate will want even more redistribution and he will win a higher vote share. While it is true that the small government candidate will want even less redistribution, he is similarly less likely to win.

In contrast, when inequality rises due to an increase in poverty (a decrease in τ^*_{min}), this increases the relative spread of the poor candidate's base with respect to the candidate of the rich. Therefore, while the platform dynamics are the same, in this case, π decreases. Thus, the small candidate will become relatively more likely to win.

Note, however, since incomes are still skewed, π is still greater than one-half. Thus, while the candidate of the rich will now offer an even more radical small-government program and will be more likely to win than he was before, the big government candidate is still more likely to win ex ante, and he is offering even more

taxes than before. This helps to explain why the evidence for the Downsian model’s claim that increases in poverty will decrease demand for redistribution tend to be weaker: while the median voter may now prefer relatively smaller government, the bigger government candidate will still be more likely to win, and now has a party base that wants even more government taxation and spending.

Indeed, taking data from the World Top Income Database and DW-Nominate Scores, the rise in inequality strongly tracks the rise in polarization (Alvaredo et al. (2014)). In Figure 7, I compare the DW-Nominate polarization used above with the Inverted Pareto-Lorenz Coefficient, which, as explained in Atkinson, Piketty, and Saez, is a strong determinant of the percentage of wealth increases that will accrue to the top of the income distribution (Atkinson, Piketty and Saez (2011)). For the sake of the analysis above, this is exactly what we need, as it holds the total wealth (i.e. that of the median voter) constant. Therefore, as the model predicts, the increase in divergence on economic issues is to a large extent driven by an increase in inequality.

6 Discussion of the Model

6.1 Beliefs

An alternative explanation for additively symmetric beliefs is that voters must undertake costly interpretation of platforms (e.g. Ergin and Sarver (2010)). We may consider voters who must expend some cost to interpret the platforms of each candidate, and hence their ex-post utility from voting for a particular candidate. In such an extended model, suppose that both platforms are initially observed as $P_j + \epsilon_i^j$ for both candidates, but that the voters may expend some cost ψ_j to reduce σ_j , therefore reducing their expected error. If a voter in the base of candidate 1 must expend a higher marginal cost to interpret the policy position of candidate 2, their ex-post error $|\epsilon_i^2|$ will be higher in expectation. We can therefore think of δ as being isomorphic to the cost of interpretation. The beliefs of the imperfect beliefs voting model are hence generated by the simplified scenario in which the marginal cost of interpreting a base candidate is zero, while being positive for the other candidate.

Search for information is also left unmodeled. Therefore, δ can also be interpreted as containing the difference in “free” (i.e. available without search) information available to members of a party’s cultural base. This is the information that is picked up tacitly via day to day interaction with different groups of people. Thus, even though the internet has lowered the cost of acquiring information, it has also allowed greater

freedom to choose particular sources, thereby increasing the asymmetry of information acquired for free by those who participate in certain social circles related to one candidate versus those related to the other. As voters are no longer exposed to the same baseline set of information and are given the ability to limit their exposure to “ideologically discordant” content (e.g. Bakshy, Messing and Adamic (2015)), we can consider this an increase in sociocultural distance in beliefs, and hence δ .

These results are also equivalent to a model with bounded rationality, in the sense that voters simply vote their beliefs without showing any concern for the possibility that they are wrong; they are unaware of the model. In general, this “irrationality” is both consistent with the literature on turnout (in which voters who show up for elections tend to believe that they have strong information) and actual voter behavior. In addition, due to the well-recognized incentive for voters to be ignorant and irrational due to their vanishingly small probability of affecting outcomes, it seems consistent to assume that they would not spend significant time modifying their voting behavior for the possibility of error. As noted by Schumpeter (1942): “The typical citizen drops down to a lower level of mental performance as soon as he enters the political field. He argues and analyzes in a way which he would readily recognize as infantile within the sphere of his real interests.”

It does seem reasonable that voters will use heuristics such as party and candidate history to aid in their voting decisions. For an overview of this literature, see Hinich and Munger (1996) and Lupia and McCubbins (1998). Indeed, Lau and Redlawsk (2001) showed that at least 75 percent of voters made the “correct” decision with respect to their policy preferences within an experimental setting. However, these results are *ex post*, ignoring the strategic role of the *possibility* for *ex-post* errors in candidate policy settings. Indeed, in equilibrium, the two parties will differentiate themselves enough such that voters are less likely to make *ex-post* errors. In addition, note that the use of heuristics actually makes those voters who do not receive precise signals less elastic with respect to “opposing” candidate policies, and therefore further reduces the incentive to appeal to these voters, increasing divergence in any equilibrium.

6.2 Cultural Bases

It is the potential asymmetry in bases which will drive the important results, so it is useful to explore the rationale for this assumption. Parties draw their candidates from a pool of members (in our terminology, from their base) who will possess certain

traits, sociological characteristics, and verbal characteristics. This is explicit in nations with a number of particular dialects, such as many Asian and African polities. Even within the United States, however, terms such as “liberty”, “freedom”, and “education” take on very different meanings for different groups of people. Therefore, even if the candidate wishes to signal his platform perfectly, he will only be able to do so for those who are most similar to him and/or who run in social circles where his positions are more widely known (the other members of the base). While it may be possible, or even virtually costless, for voters within the other base to determine his platform, they are unlikely to be willing to invest much, if anything, to acquire this information.

Consider, for example, the use of the term “welfare” within the economics profession in contrast to the general population. When an economist discusses maximizing welfare, he would simply mean improving quality of life; many others would hear the same phrase and believe that the economist means to increase the size of what is commonly referred to as the *welfare state*. A cursory internet search would allow one to square the circle, so to speak, but does one expect the voter, knowing that she is unlikely to make a difference in end policy regardless of her vote, to take the time and effort to do so?

Being within a party’s base is not the same thing as being a member of that party; it simply means that that party can credibly signal his intentions to you, and therefore you can have more certainty over what they are offering should they be elected. The concept of cultural bases is analogous to the parties having a comparative advantage in communicating with and being observed by certain groups of voters. Within Appendix B, I relax this assumption, allowing some voters to be either within the base of both parties, or neither, and allowing the cultural base to have noisy, albeit less noisy, beliefs. Generically, this leads to relatively more convergence, as the incentives of the two parties are more in line, but does not distort the qualitative results.

Another justification for the assumption of separate bases comes from the realm of media choice. As media choices on television, and particularly the internet, have proliferated, so too have the focuses and political slants of these mediums. Voters, facing a limited amount of time and desire to consume news sources, tend to sort by demographics into viewership cohorts which consume news specifically targeted towards them. Political affiliations tend to be one of the most important determinants of media choice, and therefore these sources have an incentive to provide more accurate information on the parties and candidates represented by viewers (see e.g., Groseclose and Milyo (2005); Mullainathan and Shleifer (2005); Prior (2005); Gentzkow

and Shapiro (2006); Duggan and Martinelli (2010); Gentzkow and Shapiro (2010); Puglisi (2011)). Therefore, we can consider candidate cultural bases to be the viewers of news networks that favor (in terms of accuracy) a particular candidate (i.e. Fox News for Republicans, MSNBC for Democrats). Since they will be gaining accurate information about one candidate on average, they will have a more refined signal of that candidate's platform, while they will be left with whatever passive information they acquire concerning the other candidate.

Note that property 1 provides a connection between the model's definition of a "base" and the intuitive notion of a candidate's base being those voters whom the candidate must win in order to take an election. Base voters will be those whom are most responsive to the candidate's policy pronouncements, and, therefore, on equilibrium path they will vote for the candidate with higher probability. These are the voters to whom the candidate will primarily seek to appeal. Therefore, an evangelical Christian candidate such as Mike Huckabee will seek to win evangelical Christians first and foremost, while an older veteran such as John McCain will focus on older voters with connections to the military. These will not be the only voters they seek out, but they will be the voters who take priority. Note that, unlike models of valance, these voters are not defined as exogenous voters who will naturally vote for these candidates, but instead arise endogenously as they are easier to be reached by said candidates.

7 Conclusion

In the previous literature, Downsian convergence was the norm and divergent outcomes were the cause of some external force, such as lobbying or commitment problems combined with personal candidate preferences, and therefore policy solutions to divergence have focused on regulating these areas through campaign finance reform and gender/racial quotas within legislatures. The imperfect beliefs voting model, however, shows another explanation for divergence: voters are rationally ignorant over policy and have little incentive to improve upon such ignorance combined with the fact that it is more difficult to communicate information to groups with whom you have less in common. When the core supporters of the two parties communicate less with each other, then voters are marginally more valuable to those whom they observe better. Therefore, unless the candidates and bases are perfectly identical, we will not have convergence to the median voter, but instead the divergence which we witness in reality.

This changes the role that culture and communication play in policy formation. As communities and cultures become polarized, their politics will become more polarized as well. It does not necessarily matter if the politics are polarized in and of themselves: as long as communication is symmetric for the two parties, we may still get the median voter outcome. It is usually when right-wingers and left-wingers come from different socioeconomic groups that do not communicate well with each other that we get divergence. This has potential implication for public goods and inequality, particularly in ethnically diverse developing countries, which need to be explored.

In some ways, the imperfect beliefs voting model underestimates the divergent behavior in political parties; politicians likely have some personal policy preferences as subjects to the laws that they will pass, and will therefore take advantage of their imperfect observability to pull policies towards their own preferences. This is particularly true if those with more extreme preferences are those more likely to be drawn to politics. In addition, when considering incumbency and the potential of having held office prior to election, it seems likely that there will always be one candidate whom voters can observe more perfectly.

It would be feasible and interesting to extend this framework to include endogenous search for information on the part of voters. Depending on the information structure available, the results could either be attenuated or enhanced. Therefore, one could look for interactions between this model and the literature on political information and media (e.g., Mullainathan and Shleifer (2005); Gentzkow and Shapiro (2010)).

In addition, the model should be extended to a dynamic setting to find implications for party formation, since the ethnic and cultural components of a party are not in fact exogenous and static, as modeled here. In addition, voters will have idiosyncratic ways of thinking about campaigns and will respond differently to differing messages. This has clear strategic implications for parties over and across campaign cycles, and could help shed light on why campaigns matter at all. It would also be of use to consider how the model would change the results with targeted, rather than general, redistribution, a la Lizzeri and Persico (2001).

The same case can be made for using the imperfect beliefs voting model to examine candidate selection, which would provide additional testable implications between the model and citizen-candidate models. One could imagine that such repetition would allow the use of heuristics to form prior over party and/or candidate behavior, though, as discussed above, such priors will actually increase political polarization as voters with noisier beliefs become even less elastic to party policy platforms. Therefore, the

model implies that an increase in polarization is potentially a natural outcome of voter learning over time.

References

- Acemoglu, Daron, Georgy Egorov and Konstantin Sonin. 2012. "A Political Theory of Populism." *The Quarterly Journal of Economics* 128(2):771–805.
- Achen, Christopher H. and Larry M. Bartels. 2013. "Blind Retrospection: Electoral Responses to Drought, Flu, and Shark Attacks." Working Paper.
URL: http://www.vanderbilt.edu/csdi/research/CSDI_WP_05-2013.pdf
- Agranov, Marina. 2012. "Flip-Flopping, Intense Primaries, and the Selection of Candidates." Working Paper.
URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2030395
- Al-Najjar, Nabil I. and Eran Shmaya. 2015. "Uncertainty and Disagreement in Equilibrium Models." *Journal of Political Economy* 123(4):778–808.
- Alesina, Alberto and Richard Holden. 2008. "Ambiguity and Extremism in Elections." NBER Working Paper No. 14143.
URL: <http://www.nber.org/papers/w14143.pdf>
- Alvaredo, Facundo, Tony Atkinson, Thomas Piketty and Emmanuel Saez. 2014. "The World's Top Income Database."
URL: <http://topincomes.g-mond.parisschoolofeconomics.eu/>
- American National Election Studies. 2014.
URL: electionstudies.org/index.htm
- Ansolabehere, Stephen, James M. Snyder and Charles Stewart. 2001. "Candidate Positions in Congressional Elections." *American Journal of Political Science* 45.
- Ashworth, Scott and Ethan Bueno de Mesquita. 2014. "Is Voter Competence Good for Voters?: Information, Rationality, and Democratic Performance." *American Political Science Review* 108(3):565–87.
- Atkinson, Anthony B., Thomas Piketty and Emmanuel Saez. 2011. "Top Income in the Long Run of History." *Journal of Economic Literature* 49(1):3–71.

- Bakshy, Eytan, Solomon Messing and Lada A. Adamic. 2015. "Exposure to ideologically diverse news and opinion on Facebook." *Science* 348 (6239):1130–1132.
- Bartels, Larry M. 1996. "Uninformed Votes: Information Effects in Presidential Elections." *American Journal of Political Science* 40(1):194–230.
- Berger, Mark M., Michael C. Munger and Richard F. Potthoff. 2000. "The Downsian Model Predicts Divergence." *Journal of Theoretical Politics* 12(2):228–240.
- Bernhardt, M. Daniel and Daniel E. Ingberman. 1985. "Candidate Reputation and the "Incumbency Effect"." *Journal of Public Economics* 27(1):47–67.
- Brewer, Mark D. and Jeffrey M. Stonecash. 2007. *Split: Class and Cultural Divides in American Politics*. CQ Press.
- Brunell, Thomas L., Bernard Grofman and Samuel Merrill. 2015. "Components of party polarization in the US House of Representatives." *Journal of Theoretical Politics* 28:598–624.
- Cabral, Marika and Caroline Hoxby. 2012. "The Hated Property Tax: Salience, Tax Rates, and Tax Revolts."
- Cambell, Angus, Philip E. Converse, Warren E. Miller and Donald E. Stokes. 1960. *The American Voter*. John Wiley & Sons.
- Carpini, Michael X. Delli and Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters*. Yale University Press.
- Carrillo, Juan and Micael Castanheira. 2008. "Information and Strategic Political Polarization." *Economic Journal* 118:845–874.
- Carroll, Royce, Jeff Lewis, James Lo, Nolan McCarty, Keith Poole and Howard Rosenthal. 2013. "DW-NOMINATE Scores With Bootstrapped Standard Errors."
URL: voteview.com/dwnomin.htm
- Chattopadhyay, Raghabendra and Esther Duflo. 2004. "Women as Policy Makers: Evidence from an India-wide Randomized Policy Experiment." *Econometrica* 72(5):1409–1443.
- Coate, Stephen. 2004. "Political Competition with Campaign Contributions and Informative Advertising." *Journal for the European Economic Association* 2(5):772–804.

- de Mello, Luiz and Erwin R. Tiongson. 2006. "Income Inequality and Redistributive Government Spending." *Public Finance Review* 34(3):282–305.
- Downs, Anthony. 1957. "An Economic Theory of Political Action in a Democracy." *Journal of Political Economy* 65:135–150.
- Duggan, John and Cesar Martinelli. 2010. "A Spatial Theory of Media Slant and Voter Choice." *Review of Economic Studies* 78:1–29.
- Enelow, James and Martin Hinich. 1982. "Ideology, Issues, and the Spatial Theory of Elections." *American Political Science Review* 76:493–501.
- Ergin, Haluk and Todd Sarver. 2010. "A Unique Costly Contemplation Representation." *Econometrica* 78(4):1285–1339.
- Feddersen, Timothy and Faruk Gul. 2014. "Polarization and Income Inequality: A Dynamic Model of Unequal Democracy." Working Paper.
URL: <http://econ.as.nyu.edu/docs/IO/33147/Polarizationtf.pdf>
- Feddersen, Timothy and Wolfgang Pesendorfer. 1996. "The Swing Voter's Curse." *American Economic Review* 86(3):408–424.
- Fiorina, Morris P. 1973. "Electoral Margins, Constituency Influence and Policy Moderation: A Critical Assessment." *American Politics Quarterly* 1:479–498.
- Fiorina, Morris P., Samuel J. Abrams and Jeremy C. Pope. 2005. *Culture Wars? The Myth of the Polarized Electorate*. Pearson Longman.
- Fowler, Anthony and Andrew B. Hall. 2016. "The Elusive Quest for Convergence." *Quarterly Journal of Political Science* 11:131–149.
- Frenkel, Sivan. 2014. "Competence and Ambiguity in Electoral Competition." *Public Choice* 159(1):219–234.
- Gentzkow, Matthew and Jesse M. Shapiro. 2006. "Media Bias and Reputation." *Journal of Political Economy* 114:280–316.
- Gentzkow, Matthew and Jesse M. Shapiro. 2010. "What Drives Media Slant? Evidence from U.S. Daily Newspapers." *Econometrica* 78:35–71.
- Gentzkow, Matthew, Jesse M. Shapiro and Matt Taddy. 2016. "Measuring Polarization in High-Dimensional Data: Method and Application to Congressional Speech.".
URL: <https://web.stanford.edu/~gentzkow/research/politext.pdf>

- Glaeser, Edward L., Giacomo A.M. Ponzetto and Jesse M. Shapiro. 2005. "Strategic Extremism: Why Republicans and Democrats Divide on Religious Values." *The Quarterly Journal of Economics* 120(4):1283–1330.
- Groseclose, Tim. 2001. "A Model of Candidate Location When One Candidate Has a Valence Advantage." *American Journal of Political Science* 45(4):862–886.
- Groseclose, Tim and Jeffrey Milyo. 2005. "A Measure of Media Bias." *The Quarterly Journal of Economics* 120:1191–1237.
- Gul, Faruk and Wolfgang Pesendorfer. 2009. "Partisan Politics and Election Failure with Ignorant Voters." *Journal of Economic Theory* 144:146–174.
- Healy, Andrew J., Neil Malhotra and Cecilia Hyunjung Mo. 2010. "Irrelevant Events Affect Voters' Evaluations of Government Performance." *Proceedings of the National Academy of Sciences* 107(29):12804–809.
- Hill, Seth J. 2016. "Learning Together Slowly: Bayesian Updating About Political Facts." *Journal of Politics* Forthcoming.
URL: http://www.sethjhill.com/SJH_InfoBayes.pdf
- Hinich, Melvin J. and Michael C. Munger. 1996. *Ideology and the Theory of Political Choice*. University of Michigan Press.
- Huber, Gregory A., Seth J. Hill and Gabriel S. Lenz. 2012. "Source of Bias in Retrospective Decision Making: Experimental Evidence on Voters' Limitations in Controlling Incumbents." *American Political Science Review* 106(4):720–41.
- Hymes, Dell. 1964. "Toward Ethnographies of Communication." *American Anthropologist* 66:1–34.
- Kinder, Donald R. and David O. Sears. 1985. *The Handbook of Social Psychology*. Vol. II. Random House chapter Public Opinion and Political Action, pp. 659–71.
- Klingelhöfer, Jan. 2017. "The Swing Voters' Blessing." *Journal of Economic Behavior and Organization* Forthcoming.
- Krasa, Stefan and Matthias Polborn. 2014. "Social Ideology and Taxes in a Differentiated Candidates Framework." *American Economic Review* 104(1):308–22.
- Lang, Kevin. 1986. "A Language Theory of Discrimination." *Quarterly Journal of Economics* 101(2):363–382.

- Lau, Ricahrd R. and David P. Redlawsk. 2001. "Advantages and Disadvantages of Cognitive Heuristics in Political Decision Making." *American Journal of Political Science* 45(4):951–71.
- Lazear, Edward P. 1999. "Culture and Language." *Journal of Political Economy* 107:95–126.
- Ledyard, John O. 1984. "The Pure Theory of Large Two-Candidate Elections." *Public Choice* 44:7–41.
- Leigh, Andrew. 2009. "Does the World Economy Swing National Elections?" *Oxford Bulletin of Economics and Statistics* 71(2):163–81.
- Levitt, Steve. 1996. "How do Senators Vote? Disentangling the Role of Voter Preferences, Party Affiliation, and Senator Ideology." *American Economic Review* 86(3):425–441.
- Levy, Gilat and Ronny Razin. 2015. "Correlation Neglect, Voting Behavior, and Information Aggregation." *American Economic Review* 105(4):1634–1645.
- Lindbeck, A. and J.W. Weibull. 1987. "Balanced-budget Redistribution as the Outcome of Political Competition." *Public Choice* 52(3):273–297.
- Lizzeri, Alessandro and Nicola Persico. 2001. "The Provision of Public Goods Under Alternative Electoral Incentives." *American Economic Review* 91:225–239.
- Lupia, Arthur and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge University Press.
- Mason, Lilliana. 2015. "'I Disrespectfully Agree': The Differential Effects of Partisan Sorting on Social and Issue Polarization." *American Journal of Political Science* 59(1):128–145.
- Matejka, Filip and Guido Tabellini. 2016. "Electoral Competition with Rationally Inattentive Voters." Working Paper.
URL: http://home.cerge-ei.cz/matejka/RI_voters.pdf
- McCarty, Nolan, Keith T. Poole and Howard Rosenthal. 2006. *Polarized America: The Dance of Ideology and Unequal Riches*. MIT Press.
- Meltzer, Allan and Scott Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy* 89:914–27.

- Mullainathan, Sendhil and Andrei Shleifer. 2005. "The Market for News." *American Economic Review* 95(4):1031–1053.
- Ortoleva, Pietro and Erik Snowberg. 2015. "Overconfidence in Political Behavior." *American Economic Review* 105(2):504–35.
- Ozdemir, Ugur and Ali Ihsan Ozkes. 2014. "Measuring Public Preferential Polarization." Ecole Polytechnique Centre National De La Recherche Scientifique Working Paper.
URL: <https://hal.archives-ouvertes.fr/hal-00954497/document>
- Pande, Rohini. 2003. "Can Mandated Political Representation Increase Policy Influence for Disadvantaged Minorities." *American Economic Review* 93(4):1132–1151.
- Persson, Torsten and Guido Tabellini. 2002. *Political Economics: Explaining Economic Policies*. Zeuthen Lectures.
- Pew. 2012. "A Closer Look at the Parties in 2012."
URL: <http://www.people-press.org/2012/08/23/a-closer-look-at-the-parties-in-2012/>
- Prior, Markus. 2005. "News vs. Entertainment: How Increasing Media Choice Widens Gaps in Political Knowledge and Turnout." *American Journal of Political Science* 49:577–592.
- Puglisi, Riccardo. 2011. "Being the New York Times: the Political Behaviour of a Newspaper." *The B.E. Journal of Economic Analysis & Policy* 11:1–34.
- Schumpeter, Joseph A. 1942. *Capitalism, Socialism, and Democracy*. Harper.
- Stadelmann, David, Marco Portmann and Reiner Eichenberger. 2014. "Politicians and Preferences of the Voter Majority: Does Gender Matter?" *Economics & Politics* Forthcoming.
- Virag, Gabor. 2008. "Playing for Your Own Audience: Extremism in Two-Party Elections." *Journal of Public Economic Theory* 10(5):891–922.
- Wedeen, Lisa. 2002. "Conceptualizing Culture: Possibilities for Political Science." *American Political Science Review* 96(4):713–728.
- Wolfers, Justin. 2009. "Are Voters Rational? Evidence from Gubernatorial Elections."
<http://users.nber.org/~jwolfers/Papers/Voterrationality>

Woon, Jonathan. 2015. "Primaries, Strategic Voting, and Candidate Polarization."
Working Paper.

URL: *<http://www.pitt.edu/~woon/papers/woon-primaries-aug15.pdf>*

A Proofs

A.1 Property 1

WLOG, let $P_1 < x$.

The probability of voter i voting for 1 is $1 + F(\frac{P_1 - P_2}{\delta}) - F(\frac{2x - P_1 - P_2}{\delta})$, and for voter j it is $F(\frac{2x - P_1 - P_2}{\delta}) - F(\frac{P_2 - P_1}{\delta})$.

Therefore, the responsiveness of voter i's propensity of voting w.r.t. P_1 is $\frac{1}{\delta}[f(\frac{P_1 - P_2}{\delta}) + f(\frac{2x - P_1 - P_2}{\delta})]$.

For voter j, it is $\frac{1}{\delta}[f(\frac{P_2 - P_1}{\delta}) - f(\frac{2x - P_1 - P_2}{\delta})]$.

Therefore, the difference is $\frac{2}{\delta}f(\frac{2x - P_1 - P_2}{\delta}) > 0$.

The derivative of the difference w.r.t. P_1 is $-\frac{2}{\delta}f'(\frac{2x - P_1 - P_2}{\delta})$, and the inverse for x.

A.2 Theorem 1

First, let us prove a lemma.

Lemma 1: Any pair of policy platforms (P_1, P_2) represent a political equilibria iff:

1. $\mu[\int_{P_1}^b dG(x) - \int_a^{P_1} dG(x)] = (1 - \mu)[\int_c^{P_2} dH(x) - \int_{P_2}^d dH(x)]$
2. $\mu[\int_a^{P_1} f(\frac{2x - P_1 - P_2}{\delta})dG(x) - \int_{P_1}^b f(\frac{2x - P_1 - P_2}{\delta})dG(x)] =$
 $(1 - \mu)[\int_c^{P_2} f(\frac{2x - P_1 - P_2}{\delta})dH(x) - \int_{P_2}^d f(\frac{2x - P_1 - P_2}{\delta})dH(x)]$

Proof) Since 4 and 5 must both be equal to zero for an equilibrium to exist, we can similarly set them equal to one another.

This leads directly to 1.2.

Plug 1.2 directly into either of the initial equations, and set equal to zero (an equilibrium condition):

$$\begin{aligned} & \mu[\int_{P_1}^b [f(\frac{P_1 - P_2}{\delta})]dG(x) - \int_a^{P_1} [f(\frac{P_1 - P_2}{\delta})]dG(x)] \\ & - (1 - \mu)[\int_c^{P_2} f(\frac{2x - P_1 - P_2}{\delta})dH(x) - \int_{P_2}^d f(\frac{2x - P_1 - P_2}{\delta})dH(x)] \\ & = (1 - \mu)[\int_c^{P_2} [f(\frac{P_2 - P_1}{\delta}) - f(\frac{2x - P_1 - P_2}{\delta})]dH(x) - \int_{P_2}^d [f(\frac{P_2 - P_1}{\delta}) - f(\frac{2x - P_1 - P_2}{\delta})]dH(x)] \\ & \Rightarrow \mu f(\frac{P_1 - P_2}{\delta})[\int_{P_1}^b dG(x) - \int_a^{P_1} dG(x)] = (1 - \mu)f(\frac{P_2 - P_1}{\delta})[\int_c^{P_2} dH(x) - \int_{P_2}^d dH(x)] \end{aligned}$$

By the symmetry of f, this simplifies to 1.1.

It is trivial to see that if 1.1 and 1.2 both hold, equations 4 and 5 is always equal to zero. \square

WLOG let $\mu < \frac{1}{2}$ and $\overline{G} > G_{med}$.

First, to see existence of equilibrium, note that 1.1 can be re-written as $P_2 = H^{-1}(\frac{1}{2(1-\mu)} - \frac{\mu}{1-\mu}G(P_1))$.

Therefore, 1.2 can be re-written as $\mu \left[\int_a^{P_1} f\left(\frac{2x-P_1-H^{-1}\left(\frac{1}{2(1-\mu)}-\frac{\mu}{1-\mu}G(P_1)\right)}{\delta}\right)dG(x) - \int_{P_1}^b f\left(\frac{2x-P_1-H^{-1}\left(\frac{1}{2(1-\mu)}-\frac{\mu}{1-\mu}G(P_1)\right)}{\delta}\right)dG(x) \right. \\ \left. (1-\mu) \left[\int_c^{H^{-1}\left(\frac{1}{2(1-\mu)}-\frac{\mu}{1-\mu}G(P_1)\right)} f\left(\frac{2x-P_1-H^{-1}\left(\frac{1}{2(1-\mu)}-\frac{\mu}{1-\mu}G(P_1)\right)}{\delta}\right)dH(x) - \int_{H^{-1}\left(\frac{1}{2(1-\mu)}-\frac{\mu}{1-\mu}G(P_1)\right)}^d f\left(\frac{2x-P_1-H^{-1}\left(\frac{1}{2(1-\mu)}-\frac{\mu}{1-\mu}G(P_1)\right)}{\delta}\right)dH(x) \right] \right]$

First consider $P_1 = a$.

The LHS will be $-\mu E_G[f(\frac{2x-a-H^{-1}(\frac{1}{2(1-\mu)})}{\delta})] < 0$.

The RHS will be $(1-\mu) \left[\int_c^{H^{-1}(\frac{1}{2(1-\mu)})} f(\frac{2x-P_1-H^{-1}(\frac{1}{2(1-\mu)})}{\delta})dH(x) - \int_{H^{-1}(\frac{1}{2(1-\mu)})}^d f(\frac{2x-P_1-H^{-1}(\frac{1}{2(1-\mu)})}{\delta})dH(x) \right]$,

which is > 0 by the symmetry of f and quasiconcavity of H .

Now consider $P_1 = b$:

The LHS will be $\mu E_G[f(\frac{2x-b-H^{-1}(\frac{1-2\mu}{2(1-\mu)})}{\delta})] > 0$

The RHS will be $\frac{1-2\mu}{2} E_H[f(\frac{2x-b-H^{-1}(\frac{1-2\mu}{2(1-\mu)})}{\delta})|x < H^{-1}(\frac{1-2\mu}{2(1-\mu)})] - \frac{1+2\mu}{2} E_H[f(\frac{2x-b-H^{-1}(\frac{1-2\mu}{2(1-\mu)})}{\delta})|x > H^{-1}(\frac{1-2\mu}{2(1-\mu)})] < 0$ by symmetry of f and quasiconcavity of H .

Therefore, by continuity, there must exist a $P_1 \in (a, b)$ that satisfies 1.2, and therefore an equilibrium exists.

Now take the special case.

Let us normalize $x_{med} = 0$. Therefore, $G(x) = 1 - H(-x)$

By lemma 1.1, we have that $P_1 = -P_2$.

Therefore, for 1.2 to hold, $\int_a^{P_1} f(\frac{2x}{\delta})dG(x) - \int_{P_1}^b f(\frac{2x}{\delta})dG(x) = \int_{-b}^{-P_1} f(\frac{2x}{\delta})dH(x) - \int_{-P_1}^{-a} f(\frac{2x}{\delta})dH(x)$.

Therefore, need $\int_a^{P_1} f(\frac{2x}{\delta})dG(x) = \int_{P_1}^b f(\frac{2x}{\delta})dG(x)$.

Note that the LHS is increasing in P_1 and the RHS is decreasing in P_1 . Therefore, there is a unique solution.

It is sufficient to show that the LHS $<$ RHS when $P_1 = G_{med}$ and that LHS $>$ RHS when $P_1 = x_{med}$.

Consider $P_1 = G_{med}$. We can re-write as $\frac{E[f(2x)|x < G_{med}]}{2}$ and $\frac{E[f(2x)|x > G_{med}]}{2}$.

By the fact that $G_{med} < 0$ and f is decreasing in $|x|$, the LHS $<$ RHS.

Consider $P_1 = x_{med}$. By assumption, $b \leq |a|$.

Therefore, we can re-write the LHS as $\int_a^{-b} f(\frac{2x}{\delta})\partial G(x) + \int_{-b}^0 f(\frac{2x}{\delta})\partial G(x)$.

By single-peakedness of G and symmetry of f , $\int_{-b}^0 f(\frac{2x}{\delta})\partial G(x) > \int_0^b f(\frac{2x}{\delta})\partial G(x)$.

Therefore, the LHS $>$ RHS.

Now, consider an equilibrium s.t. $\int_a^{P_1^*} f(\frac{2x}{\delta})dG(x) = \int_{P_1^*}^b f(\frac{2x}{\delta})dG(x)$ and consider an increase in δ .

At P_1^* , the LHS would be greater than the RHS. Therefore, new equilibrium $P_1 < P_1^*$. \square

A.3 Theorem 2

By lemma 1.1, we know that if $P_1 \neq P_2$, then either $P_1 < x_{med} < P_2$ or vice-versa. If $P_1 = P_2$, then both must equal x_{med} .

Note also that, by lemma 1.1, if G is symmetric about the median, $P_1 = -P_2$.

Next, looking at lemma 1.2: $\mu[\int_a^{P_1} f(\frac{2x-P_1-P_2}{\delta})dG(x) - \int_{P_1}^b f(\frac{2x-P_1-P_2}{\delta})dG(x)] = (1-\mu)[\int_a^{P_2} f(\frac{2x-P_1-P_2}{\delta})dG(x) - \int_{P_2}^b f(\frac{2x-P_1-P_2}{\delta})dG(x)]$.

If $\mu = 1 - \mu$, this is possible iff $P_1 = P_2$.

Note, also, that if G is symmetric, both the LHS and RHS are 0 iff $P_1 = P_2 = x_{med}$.

Otherwise, one side is negative, and the other is positive.

Suppose G is asymmetric and $\mu > 1 - \mu$.

By lemma 1.1, $\mu G(P_1) = \frac{1}{2} - (1 - \mu)G(P_2)$. Therefore, $P_1 \in (a, b)$.

Consider if $P_1 = P_2 = x_{med}$. Then $\int_a^{P_1} f(\frac{2x-P_1-P_2}{\delta})dG(x) - \int_{P_1}^b f(\frac{2x-P_1-P_2}{\delta})dG(x) = \int_a^{P_2} f(\frac{2x-P_1-P_2}{\delta})dG(x) - \int_{P_2}^b f(\frac{2x-P_1-P_2}{\delta})dG(x) \neq 0$ by asymmetry.

However, since $\mu \neq 1 - \mu$, the two sides of lemma 1.2 are not equal, and therefore convergence cannot be an equilibrium.

WLOG let $\mu > \frac{1}{2}$ and $\bar{G} > G_{med}$.

To construct the equilibrium, start from $P_1 = P_2 = x_{med}$. Note that $\int_a^{P_1} f(\frac{2x-P_1-P_2}{\delta})dG(x) - \int_{P_1}^b f(\frac{2x-P_1-P_2}{\delta})dG(x) = \int_a^{P_2} f(\frac{2x-P_1-P_2}{\delta})dG(x) - \int_{P_2}^b f(\frac{2x-P_1-P_2}{\delta})dG(x) > 0$, and therefore the LHS is greater than the RHS.

We can re-write the above as:

$$\int_a^{G^{-1}(\frac{1}{2\mu} - \frac{1-\mu}{\mu}G(P_2))} f(\frac{2x-G^{-1}(\frac{1}{2\mu} - \frac{1-\mu}{\mu}G(P_2))-P_2}{\sigma})dG(x) - \int_{G^{-1}(\frac{1}{2\mu} - \frac{1-\mu}{\mu}G(P_2))}^b f(\frac{2x-G^{-1}(\frac{1}{2\mu} - \frac{1-\mu}{\mu}G(P_2))-P_2}{\delta})dG(x) = \int_a^{P_2} f(\frac{2x-G^{-1}(\frac{1}{2\mu} - \frac{1-\mu}{\mu}G(P_2))-P_2}{\delta})dG(x) - \int_{P_2}^b f(\frac{2x-G^{-1}(\frac{1}{2\mu} - \frac{1-\mu}{\mu}G(P_2))-P_2}{\delta})dG(x) > 0$$

As P_2 increases, the LHS decreases and the RHS increases, until eventually the LHS becomes negative (at some point, P_2^*).

Note that the LHS stays negative for all $P_2 > P_2^*$, and note that the RHS always stays positive for all $P_2 > x_{med}$.

By continuity of F and G , the two will meet at one, and (due to the previous argument) only one, point. \square

B Relaxed Assumptions over Base Structure

It may be worthwhile to relax the assumption that each voter is within the cultural base of one, and only one, candidate. While that assumption is based on the idea that all voters can observe one candidate better than the other, whether through social network ties, communication, media choice, etc., it may be reasonable to wonder about

a situation in which a voter is either perfectly informed about, or equally unconnected to, the two candidates. For intuitive tractability, let $\delta = 1$, though one will note that none of these results depend on this assumption.

B.1 Voters in the base of both parties

First, let's consider voters who are in the base of both parties. These voters are perfectly informed about the policy platforms of both parties. Let CDF $I(\cdot) \sim [e, j]$, with associated PDF i , represent the distribution of preferences for those within the base of both parties. Therefore, voter i in the base of both parties will vote for candidate k iff:

$$(B.1) \quad \begin{aligned} \text{If } P_k < P_{-k}, \quad x_i &< \frac{P_k + P_{-k}}{2} \\ \text{If } P_k > P_{-k}, \quad x_i &> \frac{P_k + P_{-k}}{2} \end{aligned}$$

Thus, if we take μ_1 voters to be in the base of candidate 1 and μ_2 voters to be in the base of candidate 2, we can re-write the 1's vote share as:

$$(B.2) \quad \begin{aligned} \pi = \mu_1 & \left[\int_a^{P_1} [F(2x - P_1 - P_2) + 1 - F(P_1 - P_2)] dG(x) + \int_{P_1}^b [F(P_1 - P_2) + 1 - F(2x - P_1 - P_2)] dG(x) \right] \\ & + \mu_2 \left[\int_c^{P_2} [F(P_2 - P_1) - F(2x - P_1 - P_2)] dH(x) + \int_{P_2}^d [F(2x - P_1 - P_2) - F(P_2 - P_1)] dH(x) \right] \\ & + (1 - \mu_1 - \mu_2) \left[I\left(\frac{P_1 + P_2}{2}\right) \mathbf{1}(P_1 < P_2) + (1 - I\left(\frac{P_1 + P_2}{2}\right)) \mathbf{1}(P_1 > P_2) \right] \end{aligned}$$

This produces the following first-order conditions:

$$(B.3) \quad \begin{aligned} \text{FOC w.r.t. } P_1 \quad & \mu_1 \left[\int_{P_1}^b [f(P_1 - P_2) + f(2x - P_1 - P_2)] dG(x) - \int_a^{P_1} [f(P_1 - P_2) + f(2x - P_1 - P_2)] dG(x) \right] \\ & - \mu_2 \left[\int_c^{P_2} [f(P_2 - P_1) - f(2x - P_1 - P_2)] dH(x) + \int_{P_2}^d [f(P_2 - P_1) - f(2x - P_1 - P_2)] dH(x) \right] \\ & + \frac{1}{2} (1 - \mu_1 - \mu_2) i\left(\frac{P_1 + P_2}{2}\right) [\mathbf{1}(P_1 < P_2) - \mathbf{1}(P_1 > P_2)] \end{aligned}$$

$$(B.4) \quad \begin{aligned} \text{FOC w.r.t. } P_2 \quad & \mu_1 \left[\int_{P_1}^b [f(P_1 - P_2) - f(2x - P_1 - P_2)] dG(x) - \int_a^{P_1} [f(P_1 - P_2) - f(2x - P_1 - P_2)] dG(x) \right] \\ & + \mu_2 \left[\int_c^{P_2} [f(P_2 - P_1) + f(2x - P_1 - P_2)] dH(x) - \int_{P_2}^d [f(P_2 - P_1) + f(2x - P_1 - P_2)] dH(x) \right] \\ & + \frac{1}{2} (1 - \mu_1 - \mu_2) i\left(\frac{P_1 + P_2}{2}\right) [\mathbf{1}(P_1 < P_2) - \mathbf{1}(P_1 > P_2)] \end{aligned}$$

which leads to the following extension of Lemma 1:

Lemma B.1: Any pair of policy platforms (P_1, P_2) represent a political equilibria iff:

1. $\mu_1[\int_{P_1}^b dG(x) - \int_a^{P_1} dG(x)] = \mu_2[\int_c^{P_2} dH(x) - \int_{P_2}^d dH(x)]$
2. $\mu_1[\int_a^{P_1} f(2x - P_1 - P_2)dG(x) - \int_{P_1}^b f(2x - P_1 - P_2)dG(x)]$
 $= \mu_2[\int_c^{P_2} f(2x - P_1 - P_2)dH(x) - \int_{P_2}^d f(2x - P_1 - P_2)dH(x)]$
 $+ \frac{1}{2}(1 - \mu_1 - \mu_2)i(\frac{P_1 + P_2}{2})[\mathbf{1}(P_1 < P_2) - \mathbf{1}(P_1 > P_2)]$

Comparing to Lemma 1, we see that A.1.1 remains unchanged. Recall that this equations comes from the asymmetry in incentives driven by the potentially differing bases. Therefore, both parties must maintain the same balance relative to their own respective bases. The only change occurs in the second equation, where we now find that there is less asymmetry between the parties due to the existence of some voters who are in the bases of both, therefore representing voters of the form of a normal Hotelling-Downs model.

Result B.1: Consider a voting game Π , defined by $\{a, b, c, d, \mu, F, G, H\}$ as described above. Now consider the corresponding Π' that adds a group of size $(1 - \mu_1 - \mu_2)$ distributed I from e to j s.t. $I_{med} = x_{med}$ from the previous game:

The equilibrium of Π' will feature p'_1 and p'_2 s.t. $p_1 \leq p'_1 \leq p'_2 \leq p_2$.

Proof) Take the equilibrium form Π , as defined by Lemma 1.

Note that the equilibrium still satisfies B.1.1; however, it now violates A.5.2.

In order to raise the LHS of B.1.2 to reach equality, P_1 must increase.

Note that this involves moving towards the median voter.

However, in order to maintain B.1.1, the two parties must maintain symmetry around the median voter.

Therefore, P_2 must decrease (move towards the median voter).

However, note this only work is $P_2 > P_1$, and therefore the two parties must remain to the left and right respectively of the median voter.

Note, the voting game Π' is simply the original game in terms of noise, voter bases, and the median voter of the entire population. Therefore, the only change from the addition of a set of voters who are in the base of both parties is that there will be relatively less divergence. This is consistent with the standard Hotelling-Downs framework, and tells us that the closer we get to a world with perfect observability, the closer we get to convergence to the median voter.

B.2 Voters in the base of neither party

Next, let's consider voters who are in the base of neither party. These voters are unable to perfectly observe the policy platform of either party. Again, let CDF $I(\cdot) \sim [e, j]$, with associated PDF i , represent the distribution of preferences for those within the base of neither party. Therefore, voter i within this group will vote for candidate 1 iff:

$$(B.5) \quad \begin{aligned} \text{If } \epsilon_2 > x_i - P_2, \quad \epsilon_1 &\in [2x_i - P_1 - P_2 - \epsilon_2, P_2 + \epsilon_2 - P_1] \\ \text{If } \epsilon_2 < x_i - P_2, \quad \epsilon_1 &\in [P_2 + \epsilon_2 - P_1, 2x_i - P_1 - P_2 - \epsilon_2] \end{aligned}$$

Thus, if we take μ_1 voters to be in the base of candidate 1 and μ_2 voters to be in the base of candidate 2, we can re-write the 1's vote share winning as:

$$(B.6) \quad \begin{aligned} \pi = & \mu_1 \left[\int_a^{P_1} [F(2x - P_1 - P_2) + 1 - F(P_1 - P_2)] dG(x) + \int_{P_1}^b [F(P_1 - P_2) + 1 - F(2x - P_1 - P_2)] dG(x) \right] \\ & + \mu_2 \left[\int_c^{P_2} [F(P_2 - P_1) - F(2x - P_1 - P_2)] dH(x) + \int_{P_2}^d [F(2x - P_1 - P_2) - F(P_2 - P_1)] dH(x) \right] \\ & + (1 - \mu_1 - \mu_2) \left[\int_e^j \left[\int_{-\infty}^{x_i - P_2} [F(2x_i - P_1 - P_2 - \epsilon_2) - F(P_2 + \epsilon_2 - P_1)] dF(\epsilon_2) \right. \right. \\ & \left. \left. + \int_{x_i - P_2}^{\infty} [F(P_2 + \epsilon_2 - P_1) - F(2x_i - P_1 - P_2 - \epsilon_2)] dF(\epsilon_2) \right] dI(x) \right] \end{aligned}$$

This produces the following first-order conditions:

$$(B.7) \quad \begin{aligned} \text{FOC w.r.t. } P_1 \quad & \mu_1 \left[\int_{P_1}^b [f(P_1 - P_2) + f(2x - P_1 - P_2)] dG(x) - \int_a^{P_1} [f(P_1 - P_2) + f(2x - P_1 - P_2)] dG(x) \right] \\ & - \mu_2 \left[\int_c^{P_2} [f(P_2 - P_1) - f(2x - P_1 - P_2)] dH(x) + \int_{P_2}^d [f(P_2 - P_1) - f(2x - P_1 - P_2)] dH(x) \right] \\ & + (1 - \mu_1 - \mu_2) \left[\int_e^j \left[\int_{-\infty}^{x_i - P_2} [f(P_2 + \epsilon_2 - P_1) - f(2x_i - P_1 - P_2 - \epsilon_2)] dF(\epsilon_2) \right. \right. \\ & \left. \left. + \int_{x_i - P_2}^{\infty} [f(2x_i - P_1 - P_2 - \epsilon_2) - f(P_2 + \epsilon_2 - P_1)] dF(\epsilon_2) \right] dI(x) \right] \end{aligned}$$

$$(B.8) \quad \begin{aligned} \text{FOC w.r.t. } P_2 \quad & \mu_1 \left[\int_{P_1}^b [f(P_1 - P_2) - f(2x - P_1 - P_2)] dG(x) - \int_a^{P_1} [f(P_1 - P_2) - f(2x - P_1 - P_2)] dG(x) \right] \\ & + \mu_2 \left[\int_c^{P_2} [f(P_2 - P_1) + f(2x - P_1 - P_2)] dH(x) - \int_{P_2}^d [f(P_2 - P_1) + f(2x - P_1 - P_2)] dH(x) \right] \\ & + (1 - \mu_1 - \mu_2) \left[\int_e^j \left[\int_{-\infty}^{x_i - P_2} [-f(2x_i - P_1 - P_2 - \epsilon_2) - f(P_2 + \epsilon_2 - P_1)] dF(\epsilon_2) \right. \right. \\ & \left. \left. + \int_{x_i - P_2}^{\infty} [f(P_2 + \epsilon_2 - P_1) + f(2x_i - P_1 - P_2 - \epsilon_2)] dF(\epsilon_2) \right] dI(x) \right] \end{aligned}$$

which leads to the following extension of Lemma 1:

Lemma B.2: Any pair of policy platforms (P_1, P_2) represent a political equilibria iff:

1. $\mu_1[\int_{P_1}^b dG(x) - \int_a^{P_1} dG(x)]$
 $+ (1 - \mu_1 - \mu_2)[\int_e^j [\int_{-\infty}^{x_i - P_2} \frac{f(P_2 + \epsilon_2 - P_1)}{f(P_2 - P_1)} dF(\epsilon_2) - \int_{x_i - P_2}^{\infty} \frac{f(P_2 + \epsilon_2 - P_1)}{f(P_2 - P_1)} dF(\epsilon_2)] dI(x)]$
 $= \mu_2[\int_c^{P_2} dH(x) - \int_{P_2}^d dH(x)]$
2. $\mu_1[\int_a^{P_1} f(2x - P_1 - P_2) dG(x) - \int_{P_1}^b f(2x - P_1 - P_2) dG(x)]$
 $+ (1 - \mu_1 - \mu_2)[\int_e^j \int_{-\infty}^{x_i - P_2} f(2x_i - P_1 - P_2 - \epsilon_2) dF(\epsilon_2) dI(x)]$
 $= \mu_2[\int_c^{P_2} f(2x - P_1 - P_2) dH(x) - \int_{P_2}^d f(2x - P_1 - P_2) dH(x)]$
 $+ (1 - \mu_1 - \mu_2)[\int_e^j \int_{x_i - P_2}^{\infty} f(2x_i - P_1 - P_2 - \epsilon_2) dF(\epsilon_2) dI(x)]$

Therefore, we can see that having a group of voters who are within the base of neither party changes candidate incentives; however, they do so in a way that is very similar to the noise in the base of the other party. On the one hand, there is now more incentive for convergence, since the voters who observe both parties with noise act as a probabilistic version of voters who observe both parties with certainty, as above, and therefore there is competitive pressure to move towards the other party and capture more of that group. On the other hand, these voters are still worth less all else being equal than voters in your own base, since you need to move closer to this group to get them to support you with the same probability. This leads to the following result:

Result A.2: Consider a voting game Π , defined by $\{a, b, c, d, \mu, F, G, H\}$ as described above. Now consider the corresponding Π' that adds a group of size $(1 - \mu_1 - \mu_2)$ distributed I from e to j s.t. $I_{med} = x_{med}$ from the previous game:

1. The equilibrium of Π' will feature p'_1 and p'_2 s.t. $|p'_1 - p'_2| \leq |p_1 - p_2|$.
2. If $|x_{med} - p_1| \geq |x_{med} - p_2|$, $p'_1 > p_1$; if $|x_{med} - p_1| \leq |x_{med} - p_2|$, $p'_2 < p_2$.

Proof) Take the equilibrium form Π , as defined by Lemma 1.

Looking at 1.2, note that to make B.11.2 hold, it will be necessary for $|p'_1 - p'_2| \leq |p_1 - p_2|$, giving result B.12.1.

In order for convergence to occur, at least one party must move towards the median voter.

WLOG, let $|x_{med} - p_1| \geq |x_{med} - p_2|$.

Now suppose, for a contradiction, that $p'_1 \leq p_1$.

Comparing 1.1 to B.11.1, the LHS of B.11.1 would be larger than the LHS of 1.1.

Therefore, since 1.1 held with p_1 and p_2 , the RHS would need to increase as well.

However, this would involve *both* parties diverging further from the median voter, violating B.12.1.

Therefore, the party which was more divergent in Π must move closer to the median voter in Π' , giving B.12.2.

Therefore, the addition of a median-preserving set of voters who observe neither party perfectly leads to a situation with more convergence between the two parties relative to Π , and at least one party moving closer to the median voter. This comes from the fact that you can rank the “priority” of voters: your own party’s base first, followed by the voters who observe both with noise, followed by the voters in the other base. Therefore, the party who put more weight on his own base relative to the other party will always now be pulled towards the median voter by the relative addition of voters to the other side who are not in the other party’s base.

B.3 Noise within Cultural Bases

Suppose that, instead of cultural bases having perfect information, they instead simply had a more accurate signal. In particular, let members of the cultural base of party 1 have beliefs $\hat{p}_i^1 = P_1 + \epsilon_1$, where $\epsilon_1 \sim \hat{F}$ and F is a mean-preserving spread of \hat{F} . Therefore, voter i within 1’s base will vote for candidate 1 iff:

$$(B.9) \quad \begin{aligned} \text{If } \epsilon > x_i - P_2, \quad \epsilon_1 &\in [2x_i - P_1 - P_2 - \epsilon, P_2 + \epsilon - P_1] \\ \text{If } \epsilon < x_i - P_2, \quad \epsilon_1 &\in [P_2 + \epsilon - P_1, 2x_i - P_1 - P_2 - \epsilon] \end{aligned}$$

Thus, we can re-write the 1’s vote share as:

$$(B.10) \quad \begin{aligned} \pi = & \mu \left[\int_a^b \left[\int_{-\infty}^{x_i - P_2} [\hat{F}(2x_i - P_1 - P_2 - \epsilon) - \hat{F}(P_2 + \epsilon - P_1)] dF(\epsilon) \right. \right. \\ & + \left. \int_{x_i - P_2}^{\infty} [\hat{F}(P_2 + \epsilon - P_1) - \hat{F}(2x_i - P_1 - P_2 - \epsilon)] dF(\epsilon) \right] dG(x) \\ & + (1 - \mu) \left[\int_c^d \left[\int_{-\infty}^{x_i - P_2} [F(2x_i - P_1 - P_2 - \epsilon) - F(P_2 + \epsilon - P_1)] d\hat{F}(\epsilon) \right. \right. \\ & + \left. \left. \int_{x_i - P_2}^{\infty} [F(P_2 + \epsilon - P_1) - F(2x_i - P_1 - P_2 - \epsilon)] d\hat{F}(\epsilon) \right] dH(x) \right] \end{aligned}$$

and then the F.O.C. for the candidates:

$$\begin{aligned}
(B.11) \quad FOC \text{ w.r.t. } P_1 \quad & \mu \left[\int_a^b \left[\int_{-\infty}^{x_i - P_2} [\hat{f}(P_2 + \epsilon - P_1) - \hat{f}(2x_i - P_1 - P_2 - \epsilon)] dF(\epsilon) \right. \right. \\
& \left. \left. - \int_{x_i - P_2}^{\infty} [\hat{f}(P_2 + \epsilon - P_1) - \hat{f}(2x_i - P_1 - P_2 - \epsilon)] dF(\epsilon) \right] dG(x) \right] \\
& + (1 - \mu) \left[\int_c^d \left[\int_{-\infty}^{x_i - P_2} [f(P_2 + \epsilon - P_1) - f(2x_i - P_1 - P_2 - \epsilon)] d\hat{F}(\epsilon) \right. \right. \\
& \left. \left. - \int_{x_i - P_2}^{\infty} [f(P_2 + \epsilon - P_1) - f(2x_i - P_1 - P_2 - \epsilon)] d\hat{F}(\epsilon) \right] dH(x) \right]
\end{aligned}$$

$$\begin{aligned}
(B.12) \quad FOC \text{ w.r.t. } P_2 \quad & \mu \left[\int_a^b \left[\int_{-\infty}^{x_i - P_2} [-\hat{f}(2x_i - P_1 - P_2 - \epsilon) - \hat{f}(P_2 + \epsilon - P_1)] dF(\epsilon) - \int_{x_i - P_2}^{\infty} [\hat{f}(P_2 + \epsilon - P_1) \right. \right. \\
& \left. \left. - \hat{f}(2x_i - P_1 - P_2 - \epsilon)] dF(\epsilon) \right] dG(x) \right] \\
& + (1 - \mu) \left[\int_c^d \left[\int_{-\infty}^{x_i - P_2} [-f(2x_i - P_1 - P_2 - \epsilon) - f(P_2 + \epsilon - P_1)] d\hat{F}(\epsilon) - \int_{x_i - P_2}^{\infty} [f(P_2 + \epsilon - P_1) \right. \right. \\
& \left. \left. + f(2x_i - P_1 - P_2 - \epsilon)] d\hat{F}(\epsilon) \right] dH(x) \right]
\end{aligned}$$

which leads to the following extension of Lemma 1:

Lemma B.3: Any pair of policy platforms (P_1, P_2) represent a political equilibria iff:

1. $\mu \left[\int_a^b \left[\int_{-\infty}^{x - P_2} \hat{f}(P_2 + \epsilon - P_1) \partial F(\epsilon) - \int_{x - P_2}^{\infty} \hat{f}(P_2 + \epsilon - P_1) \partial F(\epsilon) \right] \partial G(x) \right]$
 $= (1 - \mu) \left[\int_c^d \left[\int_{-\infty}^{x - P_2} f(P_2 + \epsilon - P_1) \partial \hat{F}(\epsilon) - \int_{x - P_2}^{\infty} f(P_2 + \epsilon - P_1) \partial \hat{F}(\epsilon) \right] \partial H(x) \right]$
2. $\mu \left[\int_a^b \left[\int_{x - P_2}^{\infty} \hat{f}(2x - P_1 - P_2 - \epsilon) \partial F(\epsilon) - \int_{-\infty}^{x - P_2} \hat{f}(2x - P_1 - P_2 - \epsilon) \partial F(\epsilon) \right] \partial G(x) \right]$
 $= (1 - \mu) \left[\int_c^d \left[\int_{-\infty}^{x - P_2} f(2x - P_1 - P_2 - \epsilon) \partial \hat{F}(\epsilon) - \int_{x - P_2}^{\infty} f(2x - P_1 - P_2 - \epsilon) \partial \hat{F}(\epsilon) \right] \partial H(x) \right]$

Looking at Lemma B.3, one observes that B.3 simply represents a noisier version of Lemma 1, with expectations on the probability of voter i perceiving candidates 1 and 2 (respectively) being to her right and left. Therefore, all proofs can be re-created with equations B.3.1 and B.3.2 in place of equations 1 and 2 from Lemma 1.

C Polarization Comparative Statics

In order to test the model against other voting models which generate political polarization, there need to be additional testable implications. For example, let us consider one party's base becoming more extreme:

Proposition 2: Suppose f' is convex. If H shifts to the right (i.e. H' strictly first-order stochastically dominates H) in a way that preserves x_{med} , both P_1 and P_2 diverge from the median voter, with party 2 diverging by a greater amount.

The same is true if G shifts to the left in a similar fashion.

Proof) Without loss of generality, let $x_{med} = 0$ and begin from a Theorem 1 equilibrium.

Lemma 1.2 gives us $\int_c^{H^{-1}(1-G(P_1))} f(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dH(x) - \int_{H^{-1}(1-G(P_1))}^d f(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dH(x) = \int_a^{P_1} f(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dG(x) - \int_{P_1}^b f(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dG(x)$

This can be re-written as $E_h[f(\frac{2x-P_1-P_2}{\delta})|x < P_2]P_h[x < p_2] - E_h[f(\frac{2x-P_1-P_2}{\delta})|x > P_2]P_h[x > p_2] = E_g[f(\frac{2x-P_1-P_2}{\delta})|x < P_1]P_g[x < p_1] - E_g[f(\frac{2x-P_1-P_2}{\delta})|x > P_1]P_g[x > p_1]$

Consider h' s.t. x_{med} still is 0, but $h'_{med} > h_{med}$. In words, h has shifted to the right, but only to the right of the median.

Suppose $P_1^* = P_1$. This leads to $P_2^* > P_2$, and no change to the probabilities, but only to the expected f .

By construction, the effects on $f(2x-P_1-P_2)$ of P_2 increasing cancel out on both sides (due to the initial symmetry of g and h). Therefore, the LHS is now smaller than the RHS since the only net effect is on $E_h[f(2x-P_1-P_2)]$ caused by the shift in h .

WTS: The LHS is decreasing in P_1 at a faster rate than the RHS (or increasing at a slower rate).

The net derivative wrt to P_1 is $[g(P_1)H^{-1'}(1-G(P_1))-1](\int_c^{H^{-1}(1-G(P_1))} f'(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dH(x) - \int_{H^{-1}(1-G(P_1))}^d f'(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dH(x) - \int_a^{P_1} f'(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dG(x) + \int_{P_1}^b f'(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dG(x)) - g(P_1)(2+2h^{-1}(1-G(P_1))h(H^{-1}(1-G(P_1))))f(\frac{P_1-H^{-1}(1-G(P_1))}{\delta})$.

By construction, $g(P_1)h^{-1}(1-G(P_1)) > 1$.

Therefore, it is sufficient for $\int_{H^{-1}(1-G(P_1))}^d f'(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dH(x) + \int_a^{P_1} f'(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dG(x) > \int_c^{H^{-1}(1-G(P_1))} f'(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dH(x) + \int_{P_1}^b f'(\frac{2x-P_1-H^{-1}(1-G(P_1))}{\delta})dG(x)$

Since $h(x) > g(-x)$, this is true by convexity of f' . \square

As one cultural group begins to have a more extreme base, both parties may diverge further from the median than the symmetric case of Theorem 1. For the party that has more extreme members in its base, the effect is clear as the center of gravity among the group they can communicate with more easily has moved away from the median and towards the edge. In fact, the group becoming more heterogeneous moves farther away from the median. Interestingly, the party whose base stays the same also diverges as long as the mass of the error distribution is sufficiently near 0. This seems like a natural extension of the assumption of convexity of f , as small errors in beliefs are *significantly* more likely than large errors, particularly in the age of free information. They now have less worry that moderate members of their own base

may believe the opposition is better for them through imperfect beliefs, since their errors would need to be even larger to make such a mistake. In addition, since moderate members of the other cultural base perfectly observe the party becoming more extreme, they will notice the divergence and, *ceteris paribus*, will need a smaller error in beliefs to vote for the opposite party. Therefore, even the party that has the more moderate base has an incentive to, in effect, “radicalize”, as they will prefer to gain their own outer supporters with near certainty and risk that errors in belief will allow them to gain moderates from the other side.

Furthermore, the less extreme party will also win a larger share of voters:

Property 2: *As H shifts to the right, the vote share of 1 will increase, and vice-versa for 2.*

Consider moderate members of the party with a more extreme base. They perfectly observe that their own party has radicalized. Inherently, this makes them less likely to vote for the candidate of their own social group. As stated above, the party of the other social group will also diverge farther, making it unlikely that moderate voters will be better off switching; however, they view that other candidate with some error, and therefore will potentially miss his more radical platform. Therefore, since the more homogeneous social group candidate did not diverge as much as the heterogeneous candidate with more extreme observers, in expectation he will increase his vote share from the other candidate’s base.

Indeed, if we look at the American National Economic Survey on voter preferences and compare the data to politician voting habits via DW Nominat scores, we find the results in Figures 4 and 5: the percentage of Democratic voters who describe themselves as extremely liberal and the percentage of Republican voters who describe themselves as extremely conservative tracks the rise in political divergence (Carroll et al. (2013); *American National Election Studies* (2014)). This is not polarization of the general population, but simply better sorting into the parties (Fiorina, Abrams and Pope (2005)). Figures 4 and 5 are in line with the model: as more extreme voters sort into the same political parties, both parties diverge from one another. Such an exercise is not a perfect test of the model’s predictions: ideally we’d compare party platforms *within* elections, but the model is broadly consistent with the stylized facts we observe concerning polarization and elections.

We can also consider the case where one party is a “natural majority”. This is the case where one party can communicate more easily with the majority in the popula-

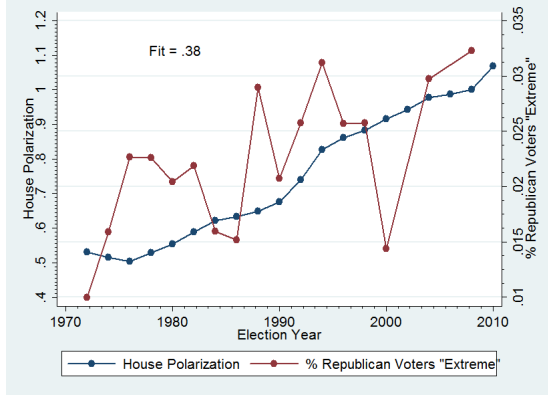


Figure 5:

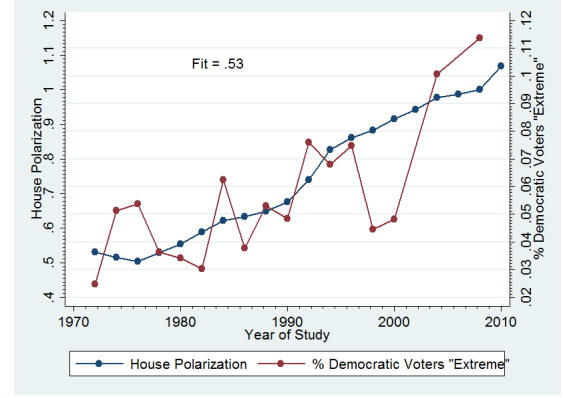


Figure 6:

tion.¹⁴ This is additionally analogous to countries where one ethnic group makes up the vast proportion of the voter population. In such cases, the median voter is the median of party 2's base, and we converge to a Theorem 2 equilibrium:

Proposition 3: *As $\mu \rightarrow 0$, both parties converge to the median voter if H is symmetric; similarly, as $\mu \rightarrow 1$ if G is symmetric.*

Otherwise, as $\mu \rightarrow 0$, party 2 converges to the median voter, but party 1 diverges; and vice-versa for $\mu \rightarrow 1$.

Proof) I will prove for $\mu \rightarrow 0$. The proof can be replicated in the other direction.

By lemma 1.1, $P_2 = H_{med} = x_{med}$ (since $\mu \rightarrow 0$)

As $\mu \rightarrow 0$, equation 1.2 can be rewritten as $\int_c^{H_{med}} f(\frac{2x-P_1-H_{med}}{\delta})dH(x) = \int_{H_{med}}^d f(\frac{2x-P_1-H_{med}}{\delta})\partial H(x)$.

By symmetry of f , the unique solution is $P_1 = H_{med}$ iff H is symmetric.

Suppose $\bar{H} > H_{med}$. If $P_1 = x_{med}$, the LHS > RHS.

By continuity and single-peakedness of f and H , there exists a unique equilibrium s.t. $P_1 > x_{med}$. \square

Intuitively, the natural majority leads to a return of convergence if and only if the natural majority has symmetric preferences. As $\mu \rightarrow 1$ or 0, it effectively returns to a case where preferences are independent from communication; the difference here is that there is still one party that has trouble communicating with what is now the general population. Therefore, we effectively recover the Theorem 2 equilibrium

¹⁴Consider a political economist running for office. Since academic economists make up a vanishingly small proportion of the population, while simultaneously speaking our own idiosyncratic language, it is reasonable that in an election between an economist and a non-economist, the non-economist would be able to better communicate with virtually everyone in the voting population. In the imperfect beliefs voting model, this is represented by $\mu \rightarrow 0$ if party 1's slew of candidates are made up of entirely economists.

where bases are identical, since the minority party's base is of no consequence.