H I E R

Harvard Institute of Economic Research

Strategic Extremism:
Why Republicans and Democrats Divide on Religious Values

by

Edward L. Glaeser, Giacomo A. M. Ponzetto
and Jesse M. Shapiro

October 2004

Harvard University
Cambridge, Massachusetts

Strategic Extremism:
Why Republicans and Democrats Divide on Religious Values

Edward L. Glaeser*
Harvard University and NBER

Giacomo A. M. Ponzetto
Harvard University

Jesse M. Shapiro
Harvard University

September 27, 2004

Abstract

Party platforms differ sharply from one another, especially on issues with religious content, such as abortion or gay marriage. Religious extremism in the U.S. appears to be strategically targeted to win elections, since party platforms diverge significantly, while policy outcomes like abortion rates are not affected by changes in the governing party. Given the high returns from attracting the median voter, why do vote-maximizing politicians veer off into extremism? In this paper, we find that strategic extremism depends on an important intensive margin where politicians want to induce their core constituents to vote (or make donations) and the ability to target political messages towards those core constituents. Our model predicts that the political relevance of religious issues is highest when around one-half of the voting population attends church regularly. Using data from across the world and within the U.S., we indeed find a non-monotonic relationship between religious extremism and religious attendance.

JEL classification: D72, D78, Z12

Keywords: voting, turnout, extremism

1

1 Introduction

In the United States, support for political parties divides on religious attitudes at least as much as on income. Figure 1 shows the share of white males between 20 and 55 who voted Republican in the last election by income bracket using data from the National Election Study.[1] Figure 2 shows the share of respondents who voted Republican by religious attendance using the same selection rule. Attending church is a better predictor of supporting George Bush than is being rich. Figure 3 shows the strikingly negative relationship between state per capita income and the share of the state that voted for Bush in 2000. As figure 4 shows, the correlation between income and party affiliation has been roughly constant since the 1960s, but the correlation between religious attendance and party affiliation has risen over this period. This trend has accompanied a rise of extremism on religious issues in party platforms. In 2004, the Republican party platform says that "the unborn child has a fundamental right to life that cannot be infringed," while the Democrats in the same year "stand proudly for a woman's right to choose."

Traditional median voter results (e.g., Hotelling, 1929) make this extremism difficult to understand. After all, vote-maximizing politicians are supposed to cater to the middle, not the edges of the distribution (Downs, 1957; Becker, 1958), at least in majoritarian systems (Cox, 1990). Political theorists who try to explain extremism emphasize the ideological preferences of the candidate or party leaders (as in Alesina, 1988).[2] While such preferences may exist, in this paper we explain why extremism occurs for purely strategic, i.e. vote-maximizing, reasons in a majoritarian system.

Extreme political platforms that deviate sharply from the median voter's preferences can be vote-maximizing only when there are two electoral margins: an extensive margin

---

[1]Women and minorities are much more likely to vote Democratic, quite possibly for non-economic reasons and they are on average less rich, so including them shows a stronger connection between income and Republicanism.

[2]But see Murphy and Shleifer (2004) and Kirchgässner (2003) for recent exceptions. Murphy and Shleifer (2004) show that the effects of social networks on beliefs can lead to an incentive to move away from the political center. Fiorina (1999) reviews the political science literature on polarization.

where a politician competes for voters from the other party and an intensive margin where the politician attempts to bring his own voters into the voting booth or elicit financial contributions (Downs, 1957). If there is no intensive margin and therefore no reason to cater to the party faithful, politicians will lose votes if they move their policies away from the preferences of the median voter. But the existence of strategic extremism requires more than just a second intensive margin. Politicians who deviate from the median gain more support from their own supporters, but they also energize their opponents' supporters. For extremism to be an equilibrium, a move from the center must increase turnout among a politician's supporters more than among his opponent's supporters.

In this paper, we present an information-based theory of strategic extremism. The key assumption in this theory is that awareness of a politician's policies is higher among the politician's supporters than among his opponent's supporters. This asymmetry means that when a politician's policies deviate from the median, he energizes his own supporters (who are more likely to be aware of this deviation) more than he energizes his opponent's supporters (who are less likely to be aware of this deviation). That politicians have an edge in communicating with their own supporters is reflected in recent efforts by the Bush and Kerry campaigns. For example, during the Republican National Convention there was a "closed, invitation-only Bush campaign rally for Christian conservatives" at which "Senator Sam Brownback of Kansas called for a broad social conservative agenda notably different from the televised presentations at the Republican convention" (Kirkpatrick, 2004b). The Kerry campaign also appears to be targeting public appearances to reach loyal Democrats (Borsuk, 2004).

In the model with only one political issue, extremism increases with the variance of voter preferences, the informational asymmetry between a politician's supporters and his opponent's, and with the ability of politicians to target political messages to their supporters. When voters differ along two attributes, such as the desire for income redistribution and abortion-related policies, extremism is more likely along the issue where there is greater

3

heterogeneity of preferences (Mullainathan and Shleifer, 2003; Irmen and Thisse, 1998). Extremism is also more likely along the issues that determine informational groups. If a candidate's audience is particularly defined by religion, then the candidate will be more extreme in her religious positions. If the candidate's audience is defined by economics, the economic policies will become more important. These findings echo Murphy and Shleifer's (2004) discussion of the role of social groups in politics.[3]

After presenting the model, we present evidence on the importance of social organizations in forming political divisions. A key difference between a model of strategic extremism and a model in which extremism reflects politicians' preferences is that, when extremism is strategic, politicians' policies will be more moderate than their messages. When extremism reflects leaders' preferences, policies will be more extreme than political messages. We examine policies and platforms on tax policy and abortion over the last 25 years to examine these implications. The economic messages in platforms are extremely moderate, but there are big differences in mean tax rates between Democratic and Republican regimes. Conversely, political messages about abortion tend to extremes, but abortion rates are independent of the party in power. These results suggest that differences in economic policies between the parties reflect preferences of party leaders but that differences in abortion and other religion-related policies reflect political strategy.

We then briefly address the issue of voter turnout. Is it plausible that politicians can be tailoring their messages to increase turnout? Anecdotes about how Democrats and Republicans target specific groups and broadcast more extreme messages to those groups abound. Both politicians and commentators argue that this targeting reflects a desire to increase turnout. Voter turnout can increase in races where the politicians differ along key dimensions. For example, between 1976, when Republicans and Democrats barely differed on religion-related issues, and 1984, when Reagan faced Mondale in a race with much starker

---

[3]We use the term social group to distinguish these groups from pressure groups and interest groups, which have received considerable attention in the literature (Becker, 1983; Dixit and Londregan, 1995; Grossman and Helpman, 2002).

divisions, turnout among the highly religious rose by seven percentage points.

We end the paper by examining a key prediction of the model: religious determinants of political orientation will be maximized when about 50 percent of the population attends church regularly and that economic determinants of political orientation will be maximized when about 50 percent of the population is in labor unions. We find a strong non-monotonic relationship across countries between church attendance at the national level and the extent to which religion determines right-wing orientation. In countries with very low levels of church attendance, such as Norway or Russia, religion is uncorrelated with political preferences. The same fact is true in those countries, like the Philippines, with very high levels of religious attendance. The countries with very strong connections between religion and being right-wing are all those, like the United States, where about one-half of the population attends church once per month or more.

After looking across countries, we turn to U.S. states. While the problem with unionization across countries was that there were few countries where more than one-half of the population is unionized, the problem with data on U.S. states is that there are few states where much less than one-half of the population attends church regularly. As such, we focused on whether church attendance increases the extent that religious determines voting Republican. In states like Mississippi, with high attendance levels (close to 100 percent), there is little connection between religion and political orientation. In states like California (where closer to 50 percent of the population attends church once per month or more), there is much more connection between religiosity and being a Republican.

Finally, we use our parameter estimates to see whether declining church attendance and declining union membership can explain the rising importance of religion in American politics. Unions have declined from being close to 35 percent of non-agricultural employment in the 1950s to being less than 15 percent of the labor force today (Freeman, 1998; Hirsch, MacPherson and Vroman, 2001). Church attendance has declined from 1968 when 57 percent of the population went church once per month or more to 2000 the point when 47 percent of

the population went to church this often. Both of these declines should have caused politics to focus less on economics and more on religion. Using our parameter estimates from the cross-state and the cross-country regressions, we estimate that these changes can explain about one-third of the increasing importance of religion and the decreasing importance of economics in explaining voting patterns.

2   Model

In this model two parties, labeled L and R, compete for votes. Each party chooses a policy proposal to maximize the difference between its votes and the votes of its opponent.[4] After a party chooses its policy proposal, a set of voters see that proposal; other voters do not see the proposal. Citizens who don't see the proposal base their votes on their prior beliefs about the politicians' proposals. We don't require these beliefs to be rational, but certainly they can be. Citizens receive utility from voting for politicians whose proposed policies are close enough to their own, and utility from voting against politicians whose policies are different enough from their own. There are also costs of voting, which differ among individuals, and people vote only when these utility gains from voting outweigh the costs of voting. By assuming that people get utility directly from voting, we are deliberately sidestepping the thorny issue of why people vote. The most straightforward justification for this assumption is that voters' decisions are emotional, not based on any estimation of how their votes will impact government policy (Schuessler, 2000). It is also possible to interpret the changes to voter utility from voting as reflecting voters' utility under different policy regimes. However, this interpretation requires policy proposals to be have predictive content (or be binding)

---

[4]This specification would be consistent with politicians maximizing the probability of victory if, for example, each party's vote totals were affected by exogenous shocks whose sum is uniformly distributed. We therefore take it to be an approximation of politicians' objective function in a deterministic context such as the one we analyze.

and that individuals don't understand the irrelevance of their vote to political outcomes.[5]

There exists an n-dimensional policy space represented by $[-1, 1]^n$. We refer to policies with negative values as left-wing positions and policies with positive values as right-wing positions. Voters have preferences represented by an ideal point in this space, and parties propose policies or platforms which are also points in this space. The distribution of citizens' ideal policies along each dimension of the policy space is independent and symmetric around the origin. As the origin represents the preferred policy of the median voter, we will refer to the distance between the origin and parties' proposed policies as the extent of extremism. Each citizen is also characterized by a cost of voting, c, which is independent of policy preferences and characterized by cumulative distribution function $Z(c)$. The timing of the model is the following:

1. All citizens have a belief $\bar{\mathbf{x}}^P \in [-1, 1]^n$ about the political proposal of each party P, and potentially an informational affiliation with one or both parties.

2. Parties simultaneously choose platforms in the space $\mathbf{x}^P \in [-1, 1]^n$ to maximize their margin of victory, and these platforms are observed by a fraction $\bar{\theta}$ of the party's affiliates, and by a fraction $\theta \leq \bar{\theta}$ of non-affiliates. Citizens who do not observe the platform maintain their initial beliefs $\bar{\mathbf{x}}^P$ about party platforms.

3. The election is held and citizens decide whether to vote and if so for which party.

We solve the model recursively starting at period 3 and the voting decision. Each citizen receives utility from voting for party P equal to

$$S_P = B - \sum_{i=1}^{n} \mu_i M\left(\left|\hat{x}_i^P - x_i^*\right|\right) \text{ where } \sum_{i=1}^{n} \mu_i = 1$$

where B measures the psychological gain from expressing support for one's favorite policy vector $\mathbf{x}^*$, the weights μ_i represents the salience of each dimension of the policy space in

---

[5]The phenomenon of voters basing turnout on the intensity of their preference for the preferred candidate has sometimes been called "abstention from indifference" (Guttman, Hilger and Shachmurove, 1994).

citizen's minds, and $M(.)$ is an increasing, convex and bounded function. The function $M(.)$ captures the fact that citizens will receive less utility if they vote for a candidate whose perceived policy proposals, $\hat{\mathbf{x}}^P$, differ from their own ideal proposals. People also receive utility from voting against party P equal to $O_P = -\gamma S_P$, where $\gamma \geq 0$. If γ equals one, then individuals compare the losses from each of the two candidates. Different values of γ will reflect particularly caring about whom you vote for or vote against depending on whether γ is less than or greater than one. The benefit from voting equals:

$$V\left(\mathbf{x}^*, \hat{\mathbf{x}}^L, \hat{\mathbf{x}}^R\right) = (1-\gamma)B + \max \left\{ \begin{array}{l} \gamma \sum_{i=1}^n \mu_i M\left(\left|\hat{x}_i^L - x_i^*\right|\right) - \sum_{i=1}^n \mu_i M\left(\left|\hat{x}_i^R - x_i^*\right|\right), \\ \gamma \sum_{i=1}^n \mu_i M\left(\left|\hat{x}_i^R - x_i^*\right|\right) - \sum_{i=1}^n \mu_i M\left(\left|\hat{x}_i^L - x_i^*\right|\right) \end{array} \right\}$$

which is less than an upper bound $\bar{V}$. Conditional on voting, people will support the party that is closer to their ideal platform, and as the act of voting imposes a cost c, people will vote when

$$c \leq V\left(\mathbf{x}^*, \hat{\mathbf{x}}^L, \hat{\mathbf{x}}^R\right).$$

We assume that c has full support on the interval $\left[0, \bar{V}\right]$, so that for every possible vector of voter preferences there are always voters who abstain. Parties anticipate that voters will make their decisions in this manner and, in period 2 of the game, they select their proposed policy to maximize their votes minus their opponent's votes, taking the opponent's proposed policy as given. We assumed that citizens have prior beliefs about the parties' platforms, but we have not assumed that these beliefs are inaccurate and it may be particularly appealing to assume that citizens share a common prior that is accurate. Our results do not rely on any assumptions about the values of these prior beliefs; we require only the possibility that individuals who observe a politician's platform can have their beliefs changed.

We make two assumptions about functional forms that greatly simplify calculations:

Assumption 1 *The loss function is quadratic*

$$M\left(\left|\hat{x}_i^P - x_i^*\right|\right) = \left(\hat{x}_i^P - x_i^*\right)^2$$

Assumption 2 *The distribution of the cost of voting is uniform*

$$
Z\left(c\right) = \begin{cases} 0 & if & c < 0 \\ \frac{c}{\bar{V}} & if & c \in \left[0, \bar{V}\right] \\ 1 & if & c > \bar{V} \end{cases}
$$

The quadratic loss function implies that $\bar{V} = \left(1 - \gamma\right) B + 4\gamma$. Assuming functional forms for preferences is costly because we have lost the ability to consider the empirical implications of different forms of loss functions. As loss functions are not directly observable, this problem may not be too severe. The assumption of a uniform distribution of voting costs is also restrictive, and in proposition 4 below we consider the implications of a more general distribution.

We now consider when parties' platforms diverge from the preferences of the median voter.

2.1 Political extremism as a result of voters' preferences

We begin by considering preference-based explanations for political extremism. To simplify matters, we consider a one-dimensional policy space and perfect information so $\theta = \bar{\theta} = 1$. Given these assumptions, extremism can still result if $\gamma < 1$, so individuals care more about the policies of the candidate that they vote for than the policies of the candidate that they vote against. In this case, parties are going to have an incentive to diverge from the median policy because of the turnout margin. By becoming more extreme, the party's platform comes nearer to the preferred policies of some inframarginal voters and these voters are increasingly induced to vote. Moving away from the center has two negative effects: losing voters in the middle and increasing turnout among the opponent's supporters. However since $\gamma < 1$, the impact on the opposing party's turnout will be less than the impact on the party's own turnout. The asymmetric impact on one's own supporters and the supporters of one's opponents makes extremism a possibility (all proofs are in the appendix):

Proposition 1 *If $\theta = \bar{\theta} = 1$, $0 < \gamma < 1$, then the equilibrium platform choices of the two parties will diverge from the origin and each other whenever voters' preferences are sufficiently heterogeneous. The distance between the two policies and the origin is monotonically decreasing in γ and increasing in the heterogeneity of voters' preferences (as measured by the mean deviation of their distribution).*

This proposition emphasizes that voter preferences are enough to generate political extremism and that this model generates a natural comparative static showing that extremism is linked to voter heterogeneity. There are, however, three reasons for doubting the usefulness of this approach. First, there is no particular reason why $\gamma < 1$ is a natural assumption. The benchmark assumption in the rational choice literature has always been that voters compare the two politicians and consider the relative benefits of voting for one or the other, which would be captured by assuming that $\gamma = 1$. Indeed, the current rash of popular literature saying that many Democrats are motivated more by hatred of Bush than by fondness for Kerry suggests that for at least these voters $\gamma > 1$. Second, this assumption generates the testable prediction that voter turnout will be highest for voters whose policies exactly match those of the candidates, not for voters in the extremes tails of the ideological distribution. Yet Figure 5 shows that voter turnout is monotonically rising with self-described ideological extremism (data from the General Social Survey, 1972-2000). Third, the degree of heterogeneity required for the proposition to hold is actually fairly extreme. For example:

Corollary 1 *When $\bar{\theta} = \theta$, extremism cannot emerge if the density of voters holding a given preference is monotone non-increasing away from the origin ($x \leq y \Leftrightarrow f(x) \geq f(y) \, \forall x, y \in [0, 1]$).*

This corollary says that extremism is only possible if densities rise away from the origin. Extremism is incompatible with any single-peaked distribution and with the uniform distribution. Explanations for political extremism based solely on voters' preferences do exist, but the simple example that we have sketched here illustrates some of the difficulties with

10

these explanations. We now turn to explaining political extremism as the result of voters' information.

2.2 Political extremism as a result of voters' information

Extremism requires that immoderation has a stronger positive impact on one's own supporters than it has a negative impact on the opponent's voters. In the previous section, differential impact occurs because people care more about the characteristics of the candidate they support than the characteristics of the candidate they oppose. In this section, differential impact occurs because the politician's supporters are more aware of changes in the politician's policies. This differential awareness might occur because individuals pay more attention to their own candidate, but it can also result from politicians strategically targeting where they broadcast their messages. Examples of this targeting include direct mailing, television ads in particular markets, speeches to the party faithful, and the use of allies, like religious leaders or unions, to broadcast information to a particular group. The model requires politicians to have some ability to provide their own supporters with more information about proposed policies.

We no longer assume that $\theta = \bar{\theta} = 1$ so that we can consider information differences and hereafter assume that $\gamma = 1$. This assumption means that voters care only about relative positions. In this section, we also assume that there is only one policy dimension. As such, a voter whose ideal policy is x_i^* and who perceives the two parties as offering policies $\hat{x}_i^R$ and $\hat{x}_i^L$, will perceive net benefits of $2x_i^* \left(\hat{x}_i^R - \hat{x}_i^L \right) - \left(\hat{x}_i^R \right)^2 + \left(\hat{x}_i^L \right)^2 - c$ of voting for the right-wing candidate R. When $\hat{x}_i^R > 0 > \hat{x}_i^L$, then benefits of voting for the right-wing candidate rise always rise with x_i^* and the net benefits of voting for the L candidate always falls with x_i^*, so these assumptions predict the pattern of voter turnout seen in Figure 5.

Given these assumptions about voting behavior, if a party has some affiliates who are more likely to be aware of the party platform, and if those affiliates do not have views that

perfectly mirror those of society as a whole, then the median voter result vanishes:

Proposition 2 *A party with a positive measure of affiliates will adopt a platform that coincides with the position of the median voter if and only if there is no informational difference between affiliates and non-affiliates ($\bar{\theta} = \theta$) or there is no difference between the ideal policies of the average affiliate and the average voter.*

The proposition tells us that any informational difference among voters is enough to break down the rush to the center. We now consider the platform decision of a party with right-wing affiliates and the determinants of the extremism of this party's platform. A natural measure of extremism is the value of x^R which captures the distance between the party's platform and the ideal platform of the median voter. Needless to say, perfectly symmetric results apply to party L when its average supporter is less conservative than the average voter.

Proposition 3 *If party affiliates are on average better informed about the party platform ($\bar{\theta} > \theta$) and more conservative than non-affiliates the party will adopt a right-wing platform so $x^R > 0$.*

The party's extremism (i.e. the value of x^R) is increasing in its ability to convey information to its affiliates ($\bar{\theta}$), and decreasing in its tendency to convey information to non-affiliates (θ). The value of x^R increases as the number of party affiliates increases (holding their average ideal policy constant) and increases as the average ideal policy within party affiliates becomes more conservative (holding the number of party affiliates constant).

If right-wing party affiliates include all citizens i for whom $x_i^* > 0$ and no others, then extremism is increasing in the heterogeneity of voters' preferences (as measured by the mean deviation of their distribution).

This proposition has three elements. First, it shows that any party whose affiliates are more conservative than the national norm will tend to choose a conservative platform.

This result will be true even if both parties have an information advantage in reaching conservative voters. In that case, both parties will choose conservative platforms, so the first result highlights that policies will cater to groups that are aware.

The second result in the proposition yields comparative statics on the extent of extremism. Extremism is more likely when the information asymmetry between affiliates and non-affiliates is greater. If the news media rapidly ensures that any speech given to the party faithful is broadcast universally then this will reduce extremism relative to a world in which these speeches are kept private. Extremism is also a function of the number of party affiliates. When there are few party affiliates then it makes little sense to cater to them by taking an extreme position. As the number of party affiliates rises, the gains from appealing to them also rise. The amount of extremism will also rise with the amount of sorting into the affiliate group. As the group that is particularly aware of changes to party platforms becomes more extreme, then policies will also become more extreme. The last part of the proposition makes the further assumption that everyone whose views are right of center is a party affiliate. In that case, extremism rises with the heterogeneity of voters' preferences. As there is a greater range of preferences in the population, party platforms will get more extreme.

Further comparative-statics results can be obtained by generalizing the distribution of the cost of voting: specifically, we modify the uniform distribution by adding a point mass in the origin, representing a group of people who always turn out to vote. The larger this group, the closer the outcome is to the median-voter result:

Proposition 4 *If the distribution of the cost of voting is generalized to include a point mass $z_0 \in [0, 1]$ of voters with zero cost of voting as well as a uniform density of voters $z(c) = \frac{1 - z_0}{\bar{V}} \forall c \in [0, \bar{V}]$, extremism can only emerge if $z_0 < 1$ and its extent (i.e. the value of x^R) is monotone decreasing in z_0.*

2.3 Political competition with two issues

In this section, we focus on the case with two issues A and T where issue T carries weight $\mu \in (0, 1)$ in voters' utility function. These letters might stand for taxes and abortion for example. An individual who perceives the two parties as having platforms $\left(\hat{T}_L, \hat{A}_L\right), \left(\hat{T}_R, \hat{A}_R\right)$ and whose ideal policies are (t, a) receives benefits from voting for the R candidate equal to two times:

$$\Delta\left(t, a; \Omega\right) = \mu\left(\hat{T}_R - \hat{T}_L\right)\left(t - \frac{\hat{T}_L + \hat{T}_R}{2}\right) + (1 - \mu)\left(\hat{A}_R - \hat{A}_L\right)\left(a - \frac{\hat{A}_L + \hat{A}_R}{2}\right)$$

where Ω is the belief set $\left(\left(\hat{T}_L, \hat{A}_L\right), \left(\hat{T}_R, \hat{A}_R\right)\right)$. The benefits of voting for the L candidate are -1 times that amount. The following lemma then characterizes voter behavior:

Lemma 1 *A voter votes for L if $\Delta\left(t, a; \Omega\right) \leq -\frac{c}{2}$, abstains if $-\frac{c}{2} < \Delta < \frac{c}{2}$, and votes for R if $\Delta\left(t, a; \Omega\right) \geq \frac{c}{2}$. In other words, among all voters characterized by $(t, a; \Omega)$, a fraction $Z\left(-2\Delta\left(t, a; \Omega\right)\right)$ vote for L, a fraction $Z\left(2\Delta\left(t, a; \Omega\right)\right)$ vote for R (where at least one of these fractions is nil), and the rest abstain.*

Figure 6 illustrates voting behavior in one potential example. The horizontal axis depicts the range of policies along dimension T and the vertical axis depicts the range of policies along dimension A. The beliefs about the two parties platforms are described by a small circle for the right-wing candidate and a small box for the left-wing candidate. The line running through the origin is the line described by the values of t and a for which the benefit function, $\Delta\left(t, a; \Omega\right) = 0$. All citizens with zero voting costs will support the R candidate if their ideal policies (i.e. (a, t) values) are above that line. All citizens with zero voting costs will support the L candidate if their ideal policies are below that line.

The line that runs above the origin reflects the values of t and a for which the benefit function, $\Delta\left(t, a; \Omega\right)$ equals some positive value $\frac{c}{2}$. The line below the origin reflects the values of t and a for which the negative of the benefit function equals $\frac{c}{2}$. These two lines describe the

voting behavior of an individual whose cost of voting equals c. If the individual's preferences lie above the upper line, then that citizen will vote for the R candidate. If the individual's preferences lie below the lower line, then that citizen will vote for the L candidate. If the individual's preferences lie between the lines, then the individual will abstain. As in the section immediately before this one, voters with more extreme preferences are more likely to go to the polls.

We consider the optimal policy choice for party R; the choice for party L remains symmetric. The probability of a voter being an affiliate is described by the function $\pi_R(t, a)$ of his ideal policies. This function is first treated as being exogenous; parties are endowed with affiliates. Extremism in the T dimension equals T_R and extremism in the A dimension is A_R, i.e. the difference between the policy platforms and the preferences of the median voter. The core results of the previous section extend to two dimensions:

Proposition 5 *A party with a positive measure of affiliates will adopt a platform that coincides with the position of the median voter if and only if there is no informational difference between affiliates and non-affiliates ($\bar{\theta} = \theta$) or there is no difference between the ideal policies of the average affiliate and the average voter along either policy dimension. If party affiliates are on average better informed about the party platform ($\bar{\theta} > \theta$) and more conservative than non-affiliates along both dimensions, then the party will choose a platform that is right-wing along both dimensions.*

The party's extremism along both dimensions will increase with its ability to convey information to its affiliates ($\bar{\theta}$) and to withhold it from non-affiliates ($-\theta$); moreover it will increase as the number of party affiliates increases (keeping their average ideal policy constant), or their average conservatism increases (keeping their number constant).

This proposition directly repeats the core results of the previous section. Extremism occurs whenever there is some ability to target information to a group whose preferences differ from the preferences of the nation as a whole. Again, extremism rises with the degree

to which information can be targeted and also with the size and bias of the group of party affiliates. The next proposition addresses the question of which dimensions will tend to dominate party politics. Will parties differ from the median voter primarily along dimension T or dimension A?

Proposition 6 *The party platform will be more extreme on the issue where the average preference of affiliates is more different from the median-voter position ($T_R \lesseqgtr A_R \Leftrightarrow E(t|R) \lesseqgtr E(a|R)$).*

If all voters for whom $t + a > 0$ are party affiliates, then a monotonic increase in the heterogeneity of voters' preferences on one issue increases extremism on that issue and reduces it on the other.

Parties' extremism is determined by the preferences of their core affiliates. If their affiliates are particularly chosen on the basis of dimension A then the parties will divide more clearly over that dimension. As such, if religious views exert a stronger push towards political involvement, then we might expect religious issues to divide parties. This result, naturally, only pushes the puzzle back one level, since we have not addressed the determinants of party affiliation, but we will turn to that issue in the next two sections.

The second result in the proposition emphasizes heterogeneity of beliefs in the population as a whole. This proposition suggests that differences of opinion, or of real circumstances, will tend to increase extremism along one issue and decrease extremism along the other issue. If heterogeneity of opinion over religion-related issues has risen substantially over the past 30 years, then this might explain some part of the rising political differentiation based on religion. Presumably, rising income inequality would tend to create further differentiation on economic grounds, although this doesn't seem to have happened over the last 30 years. There is however a case that rising inequality during the late 19th century led to a political shift where parties, such as the Democrats with William Jennings Bryan, began to become more extreme in economic platforms (as opposed to classic divisions based on the Civil War,

16

Prohibition, and religion or "Rum, Romany, and Rebellion").

2.4   Social organizations and party affiliation

In our model, party affiliation means differential access to information about the party platforms. One way in which parties may gain the ability to broadcast messages is to have access to a selected subgroup of the population, such as a Church or a Union. In this section, we take the existence of such subgroups as given, and ask how their existence influences the choice of political platform. We assume that there is no way to target voters other than to have access to these particular groups. We assume that there are two groups, whose membership is deterministic and not exclusive nor exhaustive. The first, which we label "the church", comprises all voters with a sufficiently conservative preference with respect to issue a; the second, which we label "the union", comprises all voters with a sufficiently progressive preference with respect to issue t. If we further assume that party R has a relationship with the church and can broadcast its platform to church members, and party L can broadcast its message particularly to union members, then it follows that:

Proposition 7 *If party R has access, through the church, to all voters for whom $a > \alpha$ and party L has access, through the union, to all voters for whom $t < \tau$, then both parties will move away from the median-voter position whenever $\bar{\theta} > \theta$. The R party platform will differ from zero only along the a dimension and the L party platform will differ from zero only along the t dimension.*

R-party extremism is maximized at a value of α in the interval $\left(0, \frac{\bar{\theta}-\theta}{\bar{\theta}+\theta}\right)$, and L party extremism is maximized at a value of τ in the interval $\left(-\frac{\bar{\theta}-\theta}{\bar{\theta}+\theta}, 0\right)$. The value of α that maximizes R party extremism is increasing in the party's ability to convey information to its affiliates ($\bar{\theta}$) and to withhold it from non-affiliates ($-\theta$); the value of τ that maximizes L-party extremism is decreasing in the same parameters.

17

Since the value of α that maximizes extremism is in the interior of the interval $\left(0, \frac{\bar{\theta}-\underline{\theta}}{\bar{\theta}+\underline{\theta}}\right)$, the relationship between extremism and group size is necessarily non-monotonic. Moreover, extremism is maximized when less than one-half of the population is in the group. As group size decreases from its extremism-maximizing level, the gains from extremism fall because the group becomes increasingly small and politically marginalized. As group size increases from its extremism-maximizing level, the gains from extremism also fall because the group itself becomes increasing moderate and representative of the country as a whole. When we make stronger assumptions about preferences, the relationship between extremism and group size becomes more straightforward. Figure 7 shows the connection between R party policy along the a dimension and the size of the a-based group that this party can target, assuming a uniform distribution of preferences on the interval $(-1, 1)$ and that $\frac{\bar{\theta}-\underline{\theta}}{\bar{\theta}+\underline{\theta}} = 0.2$.

The assumption that $\frac{\bar{\theta}-\underline{\theta}}{\bar{\theta}+\underline{\theta}} = 0.2$ means that the probability that a church member learns of the R party platform is $\frac{3}{2}$ of the probability that a non-church member learns of that platform. Given this assumption, extremism is maximized when the church contains 45 percent of the population. When $\frac{\bar{\theta}-\underline{\theta}}{\bar{\theta}+\underline{\theta}} < 0.2$, so there is less information asymmetry, then the extremism-maximizing church size lies between .45 and .5. When $\frac{\bar{\theta}-\underline{\theta}}{\bar{\theta}+\underline{\theta}} = 0.3$, the extremism-maximizing church size is .42 and when $\frac{\bar{\theta}-\underline{\theta}}{\bar{\theta}+\underline{\theta}} = 0.1$, the extremism-maximizing church size is .47.

While formally we assume that the connection between a party and a pre-existing social organization is exogenous, we can justify this assumption with a sketch of a more complex model. When a party has no affiliates, its optimal position is the origin. If voters originally believe both party platforms to be located at the origin, then when both parties have no affiliates, these parties will present identical platforms, receive no votes, and the election will presumably be settled by the toss of a coin. By gaining access to an organization, a party can improve its electoral chances. With access to a social group, a party will shift its policy away from the origin and towards the preferences of the members of the organization. If the opposing party is prevented from also gaining access to the organization, this shift yields an

increase in the margin of victory. Given this power, politicians should be willing to pay for exclusive access to social groups (presumably through political benefits) and the political equilibrium will reflect the outcome of this first stage in which social groups sell access to their members. Groups could sell exclusive access to their members to the highest bidding party.

If parties have the same amount of funds available to bid, we might expect each party to acquire one group, trading off the benefits of extra votes with the cost of buying access to the group. When the political relevance of the two groups is equal, the election results in a draw when each party caters to one group. More generally, the value of a group does not solely depend on group size, but also on the salience of the two issues and the distribution of preferences.

2.5   Endogenous affiliation

We have so far assumed that groups are formed exogenously and then used by politicians. In this section, we allow citizens to choose their own group in an initial stage of the model. This stage will be based on initial beliefs about party platforms and must occur before the parties broadcast their platforms. A natural basis for affiliation could be the original position of the parties: specifically, we suggest that the electorate is initially split among right-wingers and left-wingers based on their relative preference for the platforms $\bar{\mathbf{x}}^L$ and $\bar{\mathbf{x}}^R$ that they originally believe the parties to be adopting. The benefits from affiliating with a party equals:

$$J_P\left(\mathbf{x}^*\right) = \Gamma - \sum_{i=1}^{n} \lambda_i \left(\bar{x}_i^P - x_i^*\right)^2 \ \text{ with } \sum_{i=1}^{n} \lambda_i = 1$$

We further assume that this is always positive, which means that everyone will always affiliate with one of the two parties.[6] The assumption will tend to reduce extremism since many moderates will avoid party affiliation if affiliation can be costly. If moderates are

---

[6]We assume that affiliation is randomized with equal probabilities if $J_L = J_R$.

independent, then the incentives for extremism will rise. A particularly natural assumption is that the values of λ_i, that determine affiliation, are the same as the parameters, μ_i, that determine voting, but there may be interesting cases where these parameters are not identical.

In this section, we will assume that people's initial beliefs about party platforms are rational in the sense that they correctly anticipate the parties' platforms; or more formally:

Definition 1 *A point $\boldsymbol{\xi} \geq 0$ in $\mathbb{R}^N$ constitutes a* symmetric political equilibrium *if the parties' optimal platforms are $\mathbf{x}^L = -\boldsymbol{\xi}$ and $\mathbf{x}^R = \boldsymbol{\xi}$ when voters initially hold rational beliefs $\bar{\mathbf{x}}^L = -\boldsymbol{\xi}$ and $\bar{\mathbf{x}}^R = \boldsymbol{\xi}$.*

We start by considering the symmetric political equilibrium with only one policy dimension.

It is possible to have an equilibrium where both parties are expected to choose the median position, affiliation with the two parties is completely random, and parties do therefore continue to choose the median policy. However, this equilibrium is highly unstable in the sense that it would break down if there is any deviation in affiliation, so that both parties don't have ideologically identical affiliates. As soon as there is any deviation in the distribution of affiliates, the parties have an incentive to choose extreme policies.

We do not find this equilibrium interesting because of this instability and we assume that there is some differentiation in initial beliefs about the parties' platform. As long as there is some difference in initial beliefs, even if that difference is arbitrarily close to zero, citizens affiliate with party R if their ideal policy is $x^* > 0$ and with party L if $x^* < 0$. Symmetry then means that both final policies and initial beliefs about policies will be symmetric around the origin. This is a scenario we have considered previously, so the following proposition follows:

Proposition 8 *The one-dimensional political game with endogenous affiliation has a unique symmetric political equilibrium $\boldsymbol{\xi} = \frac{\bar{\theta}-\theta}{\bar{\theta}+\theta} E\left(|x|\right)$.*

Extremism (or ξ) increases with $\bar{\theta}$, decreases with θ, and increases with the heterogeneity of voters' preferences (as measured by the mean deviation of their distribution).

In weighing a deviation in party platform away from the center, party R weighs the benefit of increasing turnout among party R's affiliates against the cost of alienating moderates (who may not turn out) and energizing the opposition. A move to the center will instead reduce the turnout of left-wingers and cause some moderates to switch from party L to party R. At the equilibrium level, costs and benefits must balance out. Crucially, because R's affiliates are more likely to notice a change in R's platform than are L's affiliates, convergence to the ideal policy of the median voter is not an equilibrium.

We now turn to the bidimensional case in which $\lambda, \mu \in (0,1)$. Party affiliation is then determined by the following equation:

Lemma 2 *All individuals whose values of t and a imply that:*

$$\lambda \left(\bar{T}_R - \bar{T}_L \right) \left(t - \frac{\bar{T}_L + \bar{T}_R}{2} \right) + (1 - \lambda) \left(\bar{A}_R - \bar{A}_L \right) \left(a - \frac{\bar{A}_L + \bar{A}_R}{2} \right) > 0$$

affiliate with party R, while individuals whose values of t and a cause this inequality to fail strictly will affiliate with party L.

As λ rises, the T dimension becomes more important in determining voters' affiliation, and extremism along this dimension becomes stronger. As λ approaches 1, the affiliation divide becomes the A axis: those who prefer positive values of T affiliate with party R and those who prefer negative values with party L. In this case parties will compete only on one issue, and converge to the median on the other. If initial platforms are such that the parties are differentiated only on issue A, and voters' initial affiliations will be determined only by their preferences on issue A, then any deviation from the median position on issue T would be observed equally often by own-party and other-party affiliates, producing no net gain in votes. Therefore neither party has an incentive to deviate on issue T. Applying our earlier results to this setting yields:

Proposition 9 *There always exist two symmetric political equilibria, one where $\tau = \frac{\bar{\theta} - \theta}{\bar{\theta} + \theta} E\left(|t| \right)$ and $\alpha = 0$ and the other where $\tau = 0$ and $\alpha = \frac{\bar{\theta} - \theta}{\bar{\theta} + \theta} E\left(|a| \right)$.*

The extent of differentiation among the parties in these equilibria is increasing in the fraction of own affiliates that would notice a change, decreasing in the fraction of the opponent's affiliates that would notice a change, and increasing in the heterogeneity of voters' preferences (as measured by the mean deviation of their distribution).

The stability of these equilibria depends on the independence of the two attributes. Once a party's affiliates are determined by one dimension, independence ensures that the party has affiliates with far-ranging preferences along the other dimension. Heterogeneity among affiliates along the other dimension ensures that the party does not want to become extreme along that dimension, and remains extreme only along the dimension that drives its affiliation. These two equilibria universally exist, and in many cases, they are the only stable equilibria. For example, when preferences are uniformly distributed:

Proposition 10 *When voters' preferences are uniformly distributed on both issues, the only stable equilibria are those where differentiation is restricted to only one of the two issues.*

Parties can divide only along dimension A when $\lambda < .4$ and they can divide only along dimension T when $\lambda > .6$.

Using the results from the previous proposition, in the uniform case, the distance between the two party platforms equals $\frac{\bar{\theta} - \theta}{\bar{\theta} + \theta}$, so extremism depends entirely on the informational differences between the two groups. The second part of the proposition tells us that both equilibria only exist when λ takes on values close to .5. When one of the issues becomes particularly important in driving the affiliation decision, then political division and extremism must center on this issue.

2.5.1 A generalization

To consider the role of heterogeneity of preferences, we now extend our previous results on endogenous affiliation beyond the uniform distribution. We assume a generalization of

the uniform distribution for preferences with density $h(t,a) = \frac{1+\pi+\rho}{4} - \frac{\rho}{2}|t| - \frac{\pi}{2}|a|$ with $\pi, \rho \in \left[-\frac{1}{2}, \frac{1}{2}\right]$. The parameters ρ and π can be thought of as measuring the concentration of preferences on issues T and A, respectively. Higher values of these parameters imply greater concentrations of voters in the center of the preference distribution. The distribution has the characteristic shape of a roof when they are both positive, and its mirror image when they are negative; of course, preferences can also be concentrated on one issue and polarized on the other. Given this functional form, we no longer need the assumption that preferences are independent along the two issues.

Using Proposition 8 above, when preferences are distributed with density $f(x) = \frac{1+\psi}{2} - \psi|x|, \psi \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ the unique symmetric equilibrium is $\xi = \frac{\bar{\theta}-\theta}{\theta+\theta}E(|x|) = \frac{\bar{\theta}-\theta}{\theta+\theta}\frac{3-\psi}{6}$. In the two-dimensional case, differentiation is more likely to occur on an issue that is more salient and on which preferences are less concentrated. Therefore, to restrict the space of possible equilibria, it is convenient to assume that neither issue has at the same time a low salience and a high concentration of preferences. Given this restriction, it follows that:

Proposition 11 *If preferences are distributed according to the density function $h(t,a)$, if $\pi > 0$ then $\frac{\lambda}{1-\lambda} \leq \frac{1}{2}\left(\frac{2+2\pi-\rho}{3\pi}\right)^2$, if $\rho > 0$ then $\frac{\lambda}{1-\lambda} \geq 2\left(\frac{3\rho}{2-\pi+2\rho}\right)^2$, and $\max\{4\pi + \rho, \pi + 4\rho\} \leq 2$ then the only stable equilibria are those where differentiation is restricted to one only of the two issues.*

Differentiation can occur on issue A if and only if $\frac{\lambda}{1-\lambda} < \frac{3-\pi}{2+2\pi-\rho}$, and on issue T if and only if $\frac{\lambda}{1-\lambda} > \frac{2+2\rho-\pi}{3-\rho}$.

The assumptions needed for this proposition require that if preferences are concentrated on issue A ($\pi > 0$) then this issue is at least somewhat salient in the affiliation decision and likewise if preferences are concentrated on issue T ($\rho > 0$) then this issue is also somewhat salient. The assumption $\max\{4\pi + \rho, \pi + 4\rho\} \leq 2$ ensures that there is not too much concentration on both issues at the same time.

The proposition tells us that differentiation is more likely to occur along one issue the

more salient it is, the lower the concentration of voters' preferences on it, and the higher the concentration of voters' preferences on the other issue. The equilibrium is identical to the one that would obtain if the other issue were completely irrelevant in voters' preferences, as described in Proposition 9 above. This confirms the intuition that when more voters take extreme positions on an issue candidates are more likely to be differentiated on that issue.

2.6   Incorporating policy outcomes

Thus far we have focused our attention on party platforms, rather than policy outcomes. But the preferences of politicians and parties may well affect political outcomes, and a complete theory of the electoral process requires some accounting of how these preferences weigh off against the considerations we have modeled above. In this subsection, we address the policy decisions of politicians after they have been elected. We assume that politicians weigh their desire to fulfill their own policy preferences against their desire not to contradict their stated party platforms. We assume that politicians' choices in the election are not impacted by these ex-post considerations. This assumption can be justified if politicians' preferences are lexicographic and they care about winning more than anything else.

The party that wins the election subsequently decides which policy to enact by weighing two considerations: the personal preferences of the politician in office, and the political cost of being seen to deviate from one's electoral platform. Specifically, we assume that

$$U_P\left(\mathbf{x}\right) = G - \beta \sum_{i=1}^{n} \nu_i L\left(|x_i - x_i^*|\right) - (1-\beta) \sum_{i=1}^{n} \nu_i L\left(\left|x_i - x_i^P\right|\right) \text{ where } \sum_{i=1}^{n} \nu_i = 1$$

where $\mathbf{x}^*$ denotes the preferences of the politician and $\mathbf{x}^P$ his electoral platform, and $L\left(.\right)$ is an increasing and strictly convex function on $\mathbb{R}^+$ such that $L\left(0\right) = 0$. We assume that R-party politicians have preferences that are equally conservative on all issues $\mathbf{x}^* : x_i^* = x_R^* \geq 0 \forall i = 1...n$ while analogously L-party politicians have preferences that are equally progressive on all issues $\mathbf{x}^* : x_i^* = x_L^* \leq 0 \forall i = 1...n$.

For simplicity, suppose that $L(.)$ is quadratic, so that $L(z) = z^2$ for all z. Then considering the standard case where party R's platform is unambiguously right-wing and party L's unambiguously left-wing ($\mathbf{x}^R \geq 0 \wedge \mathbf{x}^L \leq 0$), the first-order condition for a maximum immediately proves the following:

Proposition 12 *The enacted policy is a linear combination of the electoral platform and the elected politician's ideal policy*

$$\mathbf{x} = \beta \mathbf{x}_P^* + (1 - \beta) \mathbf{x}^P \forall i = 1...n$$

where P denotes the party of the electoral winner. Therefore, the enacted policy is more extreme on one issue than on another if and only if the electoral platform is.

This provides some intuition about the issues on which policy extremism ought to exceed platform extremism, and vice versa. In particular, consider a two-dimensional case in which party R wins the election. Let (T_R, A_R) denote the party's platform, and x_R^* its ideal policy on both issues. It then follows immediately that

Corollary 2 *The enacted policy of party R is more conservative than its platform on issue T and less conservative than its platform on issue A if and only if $T_R < x_R^* < A_R$*

This corollary leads us immediately to the next section. If extremism is driven by strategic concerns, not preferences, then we should expect to see extreme platforms, and less difference in actual policies. If extremism is driven by preferences, then platforms will be more moderate than eventual policies.

3   Strategic vs. Preference Extremism

The most significant existing theory of political extremism that differs from our model is that this extremism reflects the preferences of party leaders. The previous section suggested

that this model can be tested against our model by comparing the relative extremity of platforms and policies. If extremism is based on supplier preferences, then platforms will be less extreme than policies. If extremism comes from trying to appeal to voters, then platforms will be more extreme than policies. We now examine these implications.

Over the last three decades, Republican and Democratic platforms have moved definitively away from the center on religious values. Consider, for example, trends in party platform statements about abortion. Between 1976, the first year in which either party's platform mentioned abortion, and 2004, the Republican party moved to the right on this issue:

> We protest the Supreme Court's intrusion into the family structure through its denial of the parents' obligation and right to guide their minor children. The Republican Party favors a continuance of the public dialogue on abortion and supports the efforts of those who seek enactment of a constitutional amendment to restore protection of the right to life for unborn children. (Republican Party Platform, 1976).

> As a country, we must keep our pledge to the first guarantee of the Declaration of Independence. That is why we say the unborn child has a fundamental individual right to life which cannot be infringed. We support a human life amendment to the Constitution and we endorse legislation to make it clear that the Fourteenth Amendment's protections apply to unborn children. Our purpose is to have legislative and judicial protection of that right against those who perform abortions. We oppose using public revenues for abortion and will not fund organizations which advocate it. We support the appointment of judges who respect traditional family values and the sanctity of innocent human life. (Republican Party Platform, 2004).

At the same time, the Democratic platforms trended leftward:

> We fully recognize the religious and ethical nature of the concerns which many Americans have on the subject of abortion. We feel, however, that it is undesirable to attempt to amend the U.S. Constitution to overturn the Supreme Court decision in this area. (Democratic National Platform, 1976).

> We will defend the dignity of all Americans against those who would undermine it. Because we believe in the privacy and equality of women, we stand

proudly for a woman's right to choose, consistent with Roe v. Wade, and regardless of her ability to pay. We stand firmly against Republican efforts to undermine that right. At the same time, we strongly support family planning and adoption incentives. Abortion should be safe, legal, and rare. (Democratic National Platform, 2004).

At the same time, much of the verbiage in both platforms on economic issues seems quite moderate, and similar across platforms. Indeed, both parties appear to be sending the message that they will reduce the tax burden on American businesses:

We believe that the private sector, not government is the engine of economic growth and job creation......Under John Kerry and John Edwards, 99 percent of American businesses will pay less in taxes than they do today. (Democratic Party Platform, 2004)

Small business .... deserve far better treatment from government than they have received. We will provide it through many of the initiatives explained elsewhere in this platform: lower tax rates, ending the death tax, cutting through red tape... (Republican Party Platform, 2004)

We now turn to the reality of policy implementation and ask whether, as the preference-based model would suggest, policy outcomes are more extreme than platforms. To look at abortion policy, we focus on one particularly critical abortion outcome: the number of abortions per 1000 live births. The data, which come from the CDC, are available from 1970 to 2000. To look at economic policies, we look at the mean tax rates as a share of GDP over the fiscal years 1970-2000. Table 1 shows the results of a simple comparison of mean abortion rates and tax rates between years in which the White House was Democratic and years in which it was Republican. The difference in mean abortion rates is small and statistically indistinguishable from zero ($p = 0.4710$). However, in the case of taxes the data show more divergence in policy than in rhetoric. As Table 1 shows, mean tax rates as a share of GDP differed considerably between Republican and Democratic presidential administrations during fiscal years 1970-2000. The difference is statistically distinguishable

from zero ($p = 0.0002$) and large, representing about one standard deviation in the tax share over this time period.

The results of this section present a mixed message. Economic policy divergence is accompanied by platform uniformity. This suggests that economic extremism is driven by political preferences not strategic considerations. Conversely, divergence of platforms on abortion is accompanied by no observable difference in abortion rates during the years of Democratic and Republican presidents. This combination is consistent with the view that abortion extremism is driven by strategic considerations, not political preferences.

4   Is there a Turnout Margin?

Our model depends on the idea that party leaders believe that they can increase turnout by targeting specific populations. There is abundant anecdotal information that supports the view that politicians target specific populations and tailor their messages to increase enthusiasm in that population. Platforms, and citizens' information about these platforms, can meaningfully affect voter turnout. Current mobilization efforts make it clear that politicians are trying to increase turnout by using organizations, such as churches, to communicate platforms and stimulate voter participation. At the West County Assembly of God outside of St. Louis,

> They hold open meetings for parishioners each month. They inform church members about socially conservative electoral issues. They register them to vote at stands outside the sanctuary on designated "voter registration" Sundays. Last week, the "moral action team" even drove church members to the polls, and they plan to do the same for this fall's general election as well...
> According to campaign memorandums, [the Bush campaign] has asked "people of faith team leaders" to help identify thousands of "friendly congregations" around the country. It asked religious outreach volunteers to petition their pastors to hold voter registration drives, and to speak on behalf of the campaign to Bible studies and church groups. (Kirkpatrick, 2004a)

These coordination activities are part of a more general strategy which puts "top priority on maximizing voter turnout among conservative constituencies already disposed to back the president" (Calmes and Harwood, 2004). Such anecdotes support the idea that political messages sent to these groups are more extreme than the messages broadcast to the general public. While the Republican and Democratic conventions both tried to present a mainstream message, there were abundant reports of much more extreme discussions behind closed doors.

Are politicians right? Is turnout an important margin that can shape presidential elections? Certainly, there have been examples where turnout changed significantly at least in part due to more extreme political platforms. The Republican (and Democratic) move towards extremes on abortion was discussed above. If turnout responds to these more extreme views, then we should expect to see increased turnout among particularly pro-choice or pro-life voters during this period. To examine this possibility, we looked at changes in turnout between 1976 and 1984 by religious attendance. In 1976, both Carter and Ford were relatively centrist on religious issues. In 1984, there was a big gap between the platforms of Reagan and Mondale.

As Figure 8 illustrates, the move away from the center on religious issues appears to have had a disproportionately positive impact on turnout among more religious Americans. Among those who attend church more than once a week, the probability of an eligible citizen voting increased by some seven percentage points. Reagan's more extreme policies on abortion and other issues appears to have had a significant impact on the level of voter turnout among religious Americans. Evidence such as this suggests that voter turnout does create incentives for candidates to move their platforms away from the political center.

5 Evidence on Organization Size and Political Extremism

The previous section emphasized that strategic extremism is a function of social organizations and underlying belief heterogeneity. In our empirical work we will focus on social organizations, largely because organization size is more straightforward to measure in a way that is comparable across space and time. Additionally, heterogeneity in beliefs seems more likely to itself be caused by political strategies, which seems a less significant concern in the case of organization strength.

We will focus on two of the developed world's most important social organizations: trade unions and churches. These two types of organizations are without peer for their combination of size, significance to their members and history. Within the United States, eighty-nine percent of respondents in the General Social Survey since 1990 report being a member of a religious group. Unions are far less significant in America today, although they were more important in the past and they remain more important in other countries.

Churches and unions are also particularly natural topics of investigation because we know that they regularly connect with politics. The political involvement of unions is a historical fact. They endorse candidates and in many European countries the connection between one of the major parties and the labor unions has sometimes been almost seamless (e.g. the pre-Blair Labour party in the U.K.). The political influence of the church is only somewhat less subtle. In many European countries, while the Church does not explicitly endorse candidates, in many cases church resources and preaching have backed Christian Democratic candidates. Within the U.S. as well, right-wing candidates today regularly seek support from church leaders and the opportunity to connect to the church faithful.

Churches and unions are natural social organizations also because they each have obvious policy domains. Unions inherently include people with a common set of economic interests. They do not particularly sort on religious views or foreign policy opinions. As such, the model

predicts the union density will be particularly connected with polarization along economic lines. Likewise, churches do not naturally sort along economic or foreign policy lines, and as such, we should particularly expect to see links between the size of religious organizations and the degree to which politics is polarized along religious grounds.

We will focus on the key, and perhaps somewhat surprising, implication of the model about the role of social groups that the link between group size and political polarization along the issue related to that group is non-monotonic. This leads us to three related tests of the hypothesis. First, in situations where group size varies between zero and one, we will test for the existence of a non-monotonic relationship between group size and the relevance of the group's issue. This will be the test that we use for the relationship between religious attendance and the political importance of religion across countries. Second, in situations where the group size is mostly less than one-half of the population, we will look for an increasing relationship between group size and the relevance of the group's core issues. This is the relevant case for considering heterogeneity in union density across countries or across states within the U.S. Third, in situations where the group size is mostly greater than one-half of the population we will look for a decreasing relationship between group size and the political relevance of the group's core issue.

Our basic measure of polarization along economic or religious grounds is the extent that income or religious attendance predicts supporting the right-wing candidate. In almost no cases is there a reversal where religious people or rich people are more likely to support the left. If there is a tight connection between income and supporting the right-wing candidate, we will consider a place to be highly polarized along economic grounds. If there is a tight connection between religion and supporting the right-wing candidate, we will consider a place to be highly polarized along religious grounds.

As such, for our cross-country work on church size, our core regression will take the form:

$$\text{Support for the Right} \quad = \quad b\text{*Religious Belief}+c\text{*Religious Belief*National Church Size}$$

$$+d\text{*Religious Belief* (National Church Size)}^2$$

$$+\text{Controls} +\text{Country and Year Fixed Effects}$$

The model predicts that c will be positive and d will be negative and that the maximum impact of religious belief occurs in countries where around one-half of respondents attend church regularly.

In our cross-country or cross-state work on unions, our core regression is:

$$\text{Support for the Right} \quad = \quad b\text{*Income}+c\text{*Income*Union Size}$$

$$+\text{Controls}+\text{Country and Year Fixed effects}$$

The model predicts that c will be positive. In our cross-state work on religion, we will estimate:

$$\text{Support for the Right} \quad = \quad b\text{*Religious Beliefs}+c\text{*Religious Beliefs*Church Size}$$

$$+ \text{Controls} + \text{State and Year Fixed Effects}$$

The model predicts that c will be negative in this case, since more than one-half of the population are generally religious adherents. Before turning to the regressions, we discuss our data sources and core stylized facts.

5.1 Data Sources: The World Values Survey and the General Social Survey

Our work uses two primary data sets: the World Values Survey (Inglehart et al, 2000) and the General Social Survey (Davis, Smith and Marsden, 2003). The General Social Survey (GSS)

is an annual survey taken of a random sample of U.S. residents. It contains questions on a host of demographics as well as religion. Religious attendance is scored on a nine-category scale ranging from never attending a religious institution to attending more than once per day. The key political variable will be whether the respondent voted for a Republican in the last election. We will eliminate those observations where individuals either did not vote or voted for an independent.

The World Values Survey (WVS) was loosely modeled on the General Social Survey and its questions are quite similar. The survey was conducted in three major waves in 1981-84, 1990-93 and 1995-97, and we will use data from all three waves. Like the General Social Survey, the World Values Survey has questions on many basic demographic variables and a categorical variable on church attendance. Instead of using votes in the last election, since electoral differences across countries make this quite problematic, we will use respondents' self-reported political orientation as our measure of support for the right. The World Values Survey asks people to self-report their political orientation (left vs. right) on a 10 point scale which we have standardized by subtracting the mean for each country-wave pair and dividing by the standard deviation. Our measure of support for the right will be a person's value on this scale.

We also require measures of union and church size. To measure union size for our cross-national regressions, we will rely Blanchflower's (1996) data on union density. For U.S. data on unions, we take data from Hirsch and Macpherson's (2003) database. In both cases, we believe that these are credible measures of the prevalence of unions although the heterogeneity in the structure of unions across space makes even these measures less than perfect. We take attending once per month or more as our measure of membership in a religious organization. For many purposes, it makes sense to rely on different cutoff values, and our results are robust to slightly different definitions, but we believe that monthly attendance comes closest to the spirit of the model. People who attend less than once per month are unlikely to receive much information either from the pulpit or from other church-

related activities. People who attend once per month or more can be thought of as regularly receiving some form of information through the church.

All of our regressions will include country or state fixed effects to capture the extent that political support differs over space. We also allow income and religion to have different effects in different elections. In our fixed effect regressions, we also include an interaction between country (in the World Values Survey) or state (in the General Social Survey) and the key variable of interest (i.e. religiosity or income).[7]

5.2 Cross-National Evidence on Church Size and Unions

Our first regressions look at the connection between church size and religious polarization in politics. There is remarkable heterogeneity in the World Values Survey in the amount of church attendance across countries. In some places, such as the Scandinavian countries or the Russian Federation, ten percent or less of the population attends church once per month or more. In the United States in wave 3 of the survey around 60 percent of respondents attend church once per month or more. In the Philippines 90 percent attend church once per month or more. As such, cross-national investigation of church attendance gives us our best chance of measuring the connection between organizational size and polarization for a wide range of organizational size.

There are many different measures of religious belief that we can use. In the regressions in Table 2, we use self-reported answers to the question: how important is God in your life? This question is scored on a 10 point scale ranging from "not at all" to "very". We have standardized this measure by subtracting the mean for the country-wave and dividing by the standard deviation, so that a unit increase can be interpreted as an increase of one

---

[7] As the World Values Survey and the General Social Survey contain many observations with missing data, especially for income, and since these missing observations do not appear to be randomly selected, we include observations that are missing one or more of the control variables. We then give that observation a value for missing variable equal to the mean in the control and include a set of dummy variables each of which takes on a value of one when a particular control variable is missing. We never impute values for any of the variables that are reported in the table.

standard deviation from the mean. Our results are generally robust to alternative measures of religious belief. Our regressions include controls for gender, income, age, age squared, and years of completed schooling.

In the first column of Table 2, we show the basic connection between identification with the political right and people's opinions about the importance of God. The coefficient on the importance of God is .16 meaning that as the importance of God to the respondent rises by one standard deviation, the respondent's tendency to support the right rises by .16 standard deviations. This coefficient is extremely significant statistically, and it also seems economically significant to us.

This pooled regression masks the considerable heterogeneity that exists within countries. Figure 9 shows the connection coefficients from country-level versions of regression (1) where the tendency to identify with the right is regressed on self-reported importance of God. In some countries, such as India or Turkey, the correlation is more than double our pooled estimated (.36 and .32 respectively). The estimated coefficient in the United States is slightly higher than that for the world as a whole. Then there are many countries such as Japan, Norway, Bangladesh and the Philippines were the coefficient is essentially zero.

The model predicts a non-monotonic relationship and indeed that is exactly what we see. With the exception of Bulgaria, there is no country with religious attendance below 25 percent that has a particularly strong connection between religious identification and right-wing status. There is also no country with religious attendance above 60 percent that has an above-average connection between right-wing status and religiosity. All of the countries with extremely tight connections between religion and political orientation have monthly church attendance values between .3 and .7.

Perhaps it is unsurprising that the countries with little religion also have little connection between religion and political orientation. It seems to us somewhat more counterintuitive that there is no connection between religion and politics in those countries with extremely high levels of religious attendance, although this is exactly what is predicted by the model.

35

The countries that don't conform well with the theory (apart from Bulgaria) are four countries with church attendance that includes between 45 and 55 percent of the population and still have very little political polarization along religious grounds: Chile, Brazil, the Dominican Republic and Venezuela. One potential explanation for these observations is the history of liberation theology and left-wing clerics opposing military regimes.

To test for the non-monotonic impact of church membership on the religion-right-wing connection, in regression (2) of Table 2, we interact the self-report of the importance of God with the share of the country's respondents attending church more than once per month and the square of that variable. We also include the raw values of the church attendance variable in the regression and allow importance of God to differ during the different waves of World Value Survey. Regression (2) shows both a strong positive interaction between national church attendance and the impact of the importance of God for low levels of national church attendance and a strong negative quadratic term. Both the linear and quadratic terms are highly significant statistically. The connection between right-wing status and self-reported importance of God appears to be maximized when about 45 percent of the population attends church.

In the third regression of the table, we repeat regression (2) only for those countries with Polity III democracy scores averaging at least 5 from 1970-1995. We believe that political competition also occurs in countries that aren't democracies and certainly people identify with the left and the right even when they don't have a chance to vote. Nonetheless, our model is specifically based on political competition in a democratic setting, so it makes sense to ensure that these results are robust to excluding non-democracies for the sample. In this case, the linear term remains significantly positive and the quadratic term remains significantly negative. The importance of religion on being right-wing is now maximized when 48 percent of the population attends church at least once per month.

In regression (4), we return to our entire sample of countries and allow each country to have a different inherent connection between religiosity and right-wing status. This specifi-

cation is akin to a country fixed effects specification where the country fixed effect is not in the level of right-wing status (we always have that fixed effect) but rather in the connection between religious intensity and being right-wing. In this case, all of our identification comes from changes in church attendance over time. In this specification, we again find a significant positive linear impact of national church attendance on the importance of God coefficient and a significant negative quadratic impact of national church attendance on the same coefficient. In this case, the impact of importance of God is maximized when 37 percent of the population attends church once per month or more.

In regression (5), we turn to the impact of income on right-wing identification and differences in union density across countries. Our information on union density (from Blanchflower, 1996) is available only for OECD countries and as such our sample is considerably smaller. We also include only those observations for which income is not missing. The income variable runs from 1 to 10 and reflects the person's income relative to the country average. For example, 1 reflects being in the bottom decile of the income distribution and 10 in the top decile of that distribution. We have normalized this variable by subtracting the country-wave mean and dividing by the country-wave standard deviation to make it comparable with the religion variable.

Regression (5) in Table 2 reports the overall coefficient on income which is .09. This variable includes the same basic set of age and gender controls, as well as country and year fixed effects. This variable is both statistically significant and economically meaningful. A one standard deviation increase in the income variable is associated with a .09 standard deviation increase in right-wing status. Of course, it is hardly surprising that richer people are more right-wing.

Just as in the case of religiosity, there is considerable heterogeneity across countries in the connection between income and right-wing status. Figure 10 shows the relationship between this variable and trade union density for wave 2 of the World Values Survey. In some countries, such as France and Austria, the coefficient is less than .05 (the U.S. coefficient

is also quite low). In other places such as the Netherlands or the United Kingdom, the coefficient is much larger and there is a strong connection between income and right-wing status. The figure shows a strong tendency of the coefficient on income to rise with trade union density for trade union density levels below 40 percent and then little pattern after that. In the case of trade unions there is no clear non-monotonic relationship.

One explanation for this absence is that we only have four countries with levels of trade union density above 50 percent and these countries are all in Scandinavia. Moreover, church membership is actually based on self-reported attendance to church, union membership is often based on administrative rules that reflect little about true connection to the union. As such, these countries with very high trade union membership may have large portions of the population with little actual political connection to the union so that these places may actually be closer to have an effective union membership around 50 percent.

In regression (6) of Table 2, we interact the coefficient on income with national trade union density and its square. We also allow this coefficient to differ by survey wave. The impact of income on right-wing status appears to rise with trade union density. Because we have so few countries with extremely high levels of unionization, our attempt to estimate a quadratic effect of union density resulted in standard errors too large to permit meaningful inferences. We also do not have enough time-series variation within these countries to duplicate the fixed effects methodology of regression (4). These trade union density regressions show that the connection between income and being right-wing rises with trade union density across countries, but there is little evidence either to confirm the view that this impact disappears for sufficiently high levels of union density.

Overall, the country-level regressions suggest that the importance of religion is likely to rise as religious attendance declines towards 50 percent of the population and that the importance of economics is likely to fall as union membership declines from some number below 50 percent of the population towards zero. These international results suggest that the switch within the U.S. from economic polarization to religious polarization might be related

to the simultaneous decline of both unions and church attendance.

5.3   United States Evidence on Church Attendance

For our United States evidence we will focus entirely on church size. Unfortunately, the range of union density levels across states within the U.S. is actually quite modest. Ninety percent of the respondents in our GSS sample live in states with union density levels between 6 and 26 percent in the 1990s; 50 percent live in states with levels between 8 and 20 percent. These low level of union density make it essentially impossible to estimate any compelling relationship between union density and political polarization based on income within the United States. As such, we focus our attention on church attendance in these regressions.

In Table 3, we regress voting for a Republican in the last election on church attendance, and controls for income, education, age, age squared, and gender. We include state and year fixed effects. Reported coefficients are marginal effects from a probits evaluated at sample means. The first specification looks at the impact of going to church at least once per month on voting Republican. There is a strong positive effect. People who attend church once per month are ten percent more likely to vote Republican than people who attend church less frequently than that. This basic result can be found using a large number of different measures of religiosity such as the continuous measure of church attendance or qualitative variables about the importance of God.

This basic coefficient on religious attendance obscures the considerable variation that exists within states. Figure 11 shows the coefficients from regressions like regression (1) that were run separately for each state. In some places, like Washington or Oregon, the coefficient is more than .2. In other places, such as the Carolinas, the coefficient is less than .05. Figure 11 shows the relationship between these estimated coefficients and the share of respondents in the state that attend church at least once per month. Since these shares are generally more than 50 percent across states (and almost always more than 40 percent), the

model predicts a negative relationship and that is indeed what we observe in the graph. The highly religious states have less connection between religion and voting Republican than the moderately religious states.

In regression (2) of Table 3, we estimate the interaction between the share in the state attending church at least once per month and the individual's own religious attendance. We also include an interaction between the year of the last election and religiosity to allow for the fact that the connection between religion and voting Republican has changed over time. The impact of religion on voting republican decreases sharply with the share of the state that is itself religious.

In regression (3), we include interactions between state dummies and individual religious attendance. This procedure duplicates our methodology in the previous table where we allowed each country to have its own connection between religion and right-wing orientation. This estimation relies on changes in religious attendance within the state, not on variation across states. The estimate of the cross effect between average religious attendance in the state is still negative and not statistically distinguishable from the estimate in regression (2). However, the standard error on this estimate is so large that it is also not statistically distinguishable from zero. Our variation in state church attendance over time is extremely noisy and as such it is unsurprising that are results are not precise.

In regression (4), we repeat regression (2), using the categorical variation for church attendance as our measure of religiosity. The raw effect of this variable on voting Republican is quite positive, and the effect gets smaller as the state becomes sufficiently religious. The coefficients are in this case quite statistically significant. We have duplicated this regression using a wide variety of different measures of religiosity and confirmed that these results are quite robust to different ways of capturing religion. In all of these methods, those states with more religious attendance have a weaker connection between religion and voting Republican.

Table 4 reproduces the specifications of Table 3 using the American National Election Studies (Sapiro and Rosenstone, 2002). The National Election Studies (NES) have been

40

conducted since 1948, and are designed to track the determinants of voter behavior over time. As in Table 3, our main interest will be in asking whether church attendance matters less for voting behavior in highly religious states.

In specification (1) of Table 4, we reproduce the basic finding that monthly churchgoers are significantly more likely to vote Republican during the period 1972-2000. In specification (2), we interact the monthly attendance dummy with a measure of the share of people who attend church monthly or more in the respondent's state and election year. As in Table 3, we find a statistically significant negative interaction: church attendance has less of an effect on the probability of voting Republican in highly religious states. The coefficient is similar in magnitude to the analogous coefficient in Table 3. Column (3) of Table 4 repeats the specification of column (2), allowing the effect of monthly attendance to differ freely by state. Again we find a significant negative interaction, although the standard errors are larger due to the inclusion of state fixed effects interacted with monthly attendance. Finally, column (4) shows that the results are robust to using a categorical measure of church attendance rather than a binary variable for monthly attendance.

5.4 Can Our Results Explain the Rise of Religious Polarization within the United States?

Can the broad secular trend in religious adherence explain the rise of religion as a determinant of political affiliation in the United States? For the purposes of this exercise, we will take this trend as being exogenous, even though political factors interact with religion. This exercise is best thought of as a calibration where we ask whether the observed changes in religion are large enough, given our parameter estimates, to explain the changing pattern of U.S. politics.

In column 1 of Table 5, we show the impact of religion and income on voting Republican in 1968 and 1992. These are marginal effects estimated from probit regressions where other

variables are held constant. In column 1, we show that the coefficient on religion rose from 0.07 to 0.16 over this period. In column 2 of the table, we show that monthly religious attendance in the United States declined from 57 percent to 49 percent over this same time period. This change is not continuous and is concentrated in the early years of the period. In general, our framework seems to offer more possibility of understanding the broad trend rather than of understanding the exact timing of changes in politics. Using the coefficient of -.35 estimated from the state regressions, this predicts an increase in the importance of religion of 2.9% in voting Republican over this time period or 33 percent of the total change.

These results suggest that changing membership can explain some, but not all, of the rise of religion in politics. Hopefully future work will focus on other causes of this shift. One natural candidate explanation is the rise of technologies, such as direct mail, that facilitate sending targeted messages to a party's supporters.

Turning to the political importance of income, our estimates from international regressions predict that the decline in unionization from 26.9 percent in 1968 to 15 percent in 1992 would produce a decline in the effect of a one standard deviation increase in income on right-wing orientation of about .02 standard deviations. Though difficult to compare with our estimates for the US, this relatively small decline does not seem inconsistent with the mild time trends in the effects of income (as compared with the effects of religion) presented in Figure 4.

6   Conclusion

Even in majoritarian systems, there are often major differences in policies between political candidates. There is rarely complete convergence by candidates to the views of the median voter. In the U.S. today, the gulf between the Republican and Democratic parties on religiously-oriented issues like abortion is quite significant. In this paper, we have explored the reasons why political platforms might be extreme rather than moderate.

Our core model emphasizes the role of the intensive margin of getting voters into the voting booth and informational asymmetries in creating extremism. If politicians care only about attracting the median voter and have little interest in inducing their core supporters to show up, then there will be little extremism. Even with an important voter turnout margin, a second asymmetry is required so that when politicians deviate from the center they benefit more from attracting their own supporters than they lose from alienating their opponent's supporters. The heart of our model is that when a politician deviates from the center, his own supporters are more aware of this deviation than his opponent's supporters. With a single relevant policy dimension, extremism is more likely when the information asymmetry is more extreme and when informational groupings are more closely tied to political tastes. Extremism is also, unsurprisingly, more common when underlying preferences are more extreme.

The model with two issues predicts that the relevance of organizations to political extremes is non-monotonic in the size of the organization. When the organization is quite small, it is politically irrelevant. When the organization is very large (i.e. more than 50 percent of the population), it no longer creates the opportunity for the transmission of targeted messages. Looking across countries in the world, we confirm this non-monotonicity in the case of religion. Countries with either very high or very low levels of church attendance have little connection between religiosity and right-wing orientation. Extreme connections between religion and politics occur only for countries where church attendance is around 50 percent. We also find a positive association between union density and the connection between income and conservation orientation. Across U.S. states, we find a negative relationship where more church decreases the connection between religiosity and voting Republican. These parameter estimates suggest that declining church attendance in the U.S. can explain one third of the rise in the connection between religious attendance and voting republic between 1972 and 2000.

In this paper, we have avoided discussing institutions that might tend to further or

43

modify extremism, but this is clearly an important topic for future research. Some of these institutional links have already been examined by Cox (1990) and others, but this paper suggests further directions for research on constitutional design. For example, if geography creates a natural ability for politicians to target their messages, i.e. if people in your own area are more aware of your policies than people elsewhere, then national systems are always at risk of politicians proposing extreme regional policies. One way to counter this tendency is to have voting systems, like the U.S. electoral college, that limit the value of getting extra votes in one particular geographic area.

References

[1] Alberto Alesina. Credibility and policy convergence in a two-party system with rational voters. *American Economic Review*, 78(4):796–805, September 1988 1988.

[2] Gary S. Becker. Competition and democracy. *Journal of Law and Economics*, 1:105–109, October 1958.

[3] Gary S. Becker. A theory of competition among pressure groups for political influence. *Quarterly Journal of Economics*, 98(3):371–400, August 1983.

[4] David G. Blanchflower. The role and influence of trade unions in the OECD. *Report to the Bureau of International Labor Affairs, U.S. Department of Labor*, August 1996.

[5] Alan J. Borsuk. Bush, Kerry preach to faithful; non-supporters unlikely to get close to candidates. *Milwaukee Journal Sentinel*, page A1, August 23 2004.

[6] Jackie Calmes and John Harwood. Campaign '04–mobilization plan: Bush's big priority: Energize conservative Christian base–unusual strategy plays down importance of swing vote as demographics shift–a coordinator in each church. *Wall Street Journal*, page A1, August 30 2004.

[7] Gary W. Cox. Centripetal and centrifugal incentives in electoral systems. *American Journal of Political Science*, 34(4):903–935, November 1990.

[8] James A. Davis, Tom W. Smith, and Peter V. Marsden. *General Social Surveys, 1972-2002*, volume 3728. ICPSR, Ann Arbor, MI, 2003.

[9] Avinash Dixit and John Londregan. Redistributive politics and economic efficiency. *American Political Science Review*, 89(4):856–866, December 1995.

[10] Anthony Downs. An economic theory of political action in a democracy. *Journal of Political Economy*, 65(2):135–150, April 1957.

[11] Morris P. Fiorina. Whatever happened to the median voter? *Stanford University Mimeograph*, 1999.

[12] Richard B. Freeman. Spurts in union growth: Defining moments and social processes. In Michael D. Bordo, Claudia Goldin, and Eugene N. White, editors, *The Defining Moment: The Great Depression and the American Economy in the Twentieth Century*, chapter 8. The University of Chicago Press, 1998.

[13] Gene M. Grossman and Elhanan Helpman. *Special Interest Politics*. MIT Press, Cambridge, MA, August 2002.

[14] Joel M. Guttman, Naftali Hilger, and Yochanan Schachmurove. Voting as investment vs. voting as consumption: New evidence. *Kyklos*, 47(2):197–207, 1994.

[15] Barry T. Hirsch and David A. Macpherson. Union membership and coverage database from the current population survey: Note. *Industrial and Labor Relations Review*, 56(2):349–354, January 2003.

[16] Barry T. Hirsch, David A. Macpherson, and Wayne G. Vroman. Estimates of union density by state. *Monthly Labor Review*, pages 51–55, July 2001.

[17] Harold Hotelling. Stability in competition. *Economic Journal*, 39(153):41–57, March 1929.

[18] Ronald Inglehart et al. *World Values Surveys and European Values Surveys, 1981-84, 1990-93, and 1995-97*, volume 2790. ICPSR, Ann Arbor, MI, 2000.

[19] Andreas Irmen and Jean-Francois Thisse. Competition in multi-characteristics spaces: Hotelling was almost right. *Journal of Economic Theory*, 78:76–102, 1998.

[20] Gebhard Kirchgässner. Abstention because of indifference and alienation, and its consequences for party competition: A simple psychological model. *University of St. Gallen Discussion Paper*, 2003.

[21] David D. Kirkpatrick. Churches see an election role and spread the word on bush. *New York Times*, August 9 2004a.

[22] David D. Kirkpatrick. The Republicans: The convention in New York–the religion issue. *New York Times*, page P1, September 1 2004b.

[23] Sendhil Mullainathan and Andrei Shleifer. The market for news. December 2003. Harvard University mimeo.

[24] Kevin M. Murphy and Andrei Shleifer. Persuasion in politics. *American Economic Review*, 94(2):435–439, May 2004.

[25] Virginia Sapiro and Steven J. Rosenstone. *American National Election Studies Cumulative Data File*, volume 8475. ICPSR, Ann Arbor, MI, 2002.

[26] Alexander A. Schuessler. *A logic of expressive choice*. Princeton University Press, Princeton, NJ, 2000.
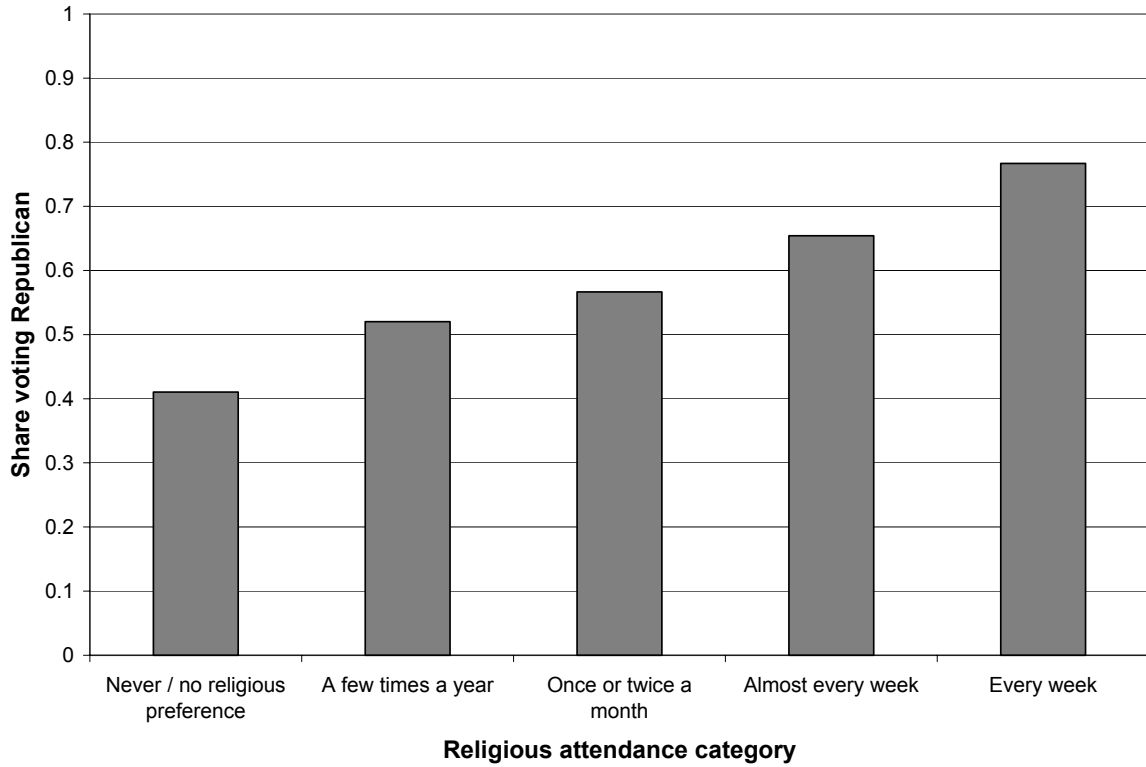
Figure 1 *Income and Republicanism*
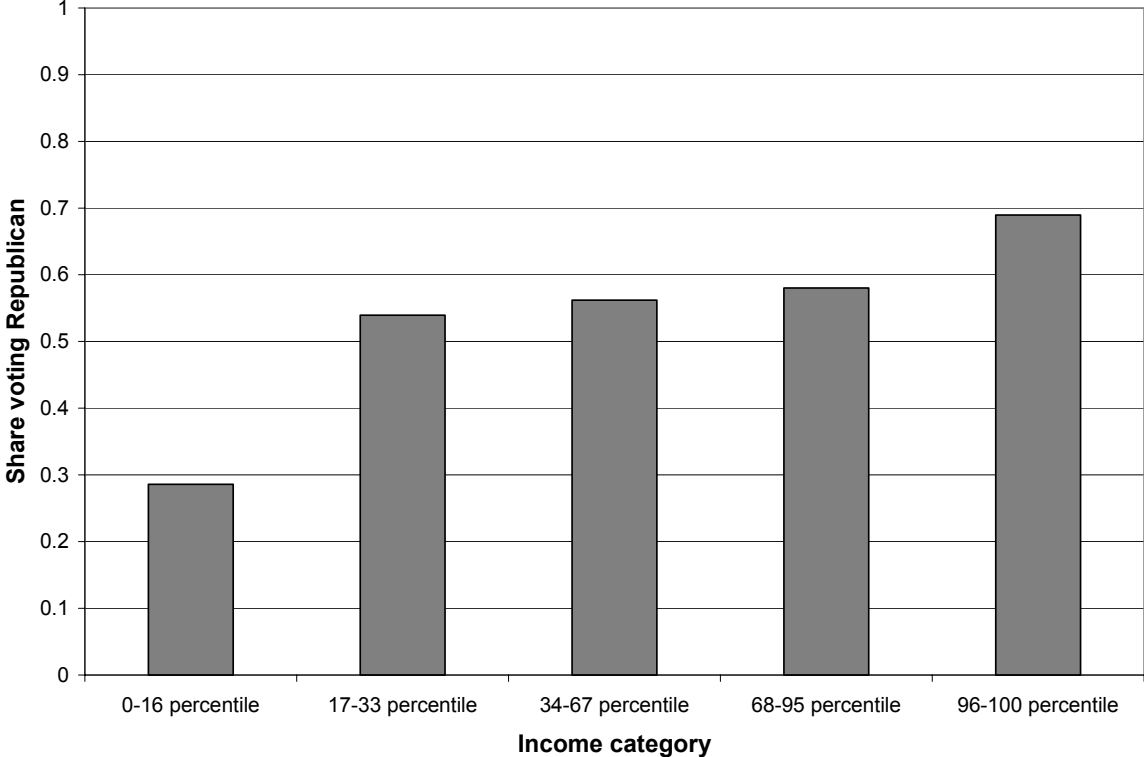
Notes: Data are from National Election Study Cumulative File. Data shown are for years 1988-2000 and include only white males between the ages of 20 and 55 who voted for a Democratic or Republican candidate in the previous presidential election.
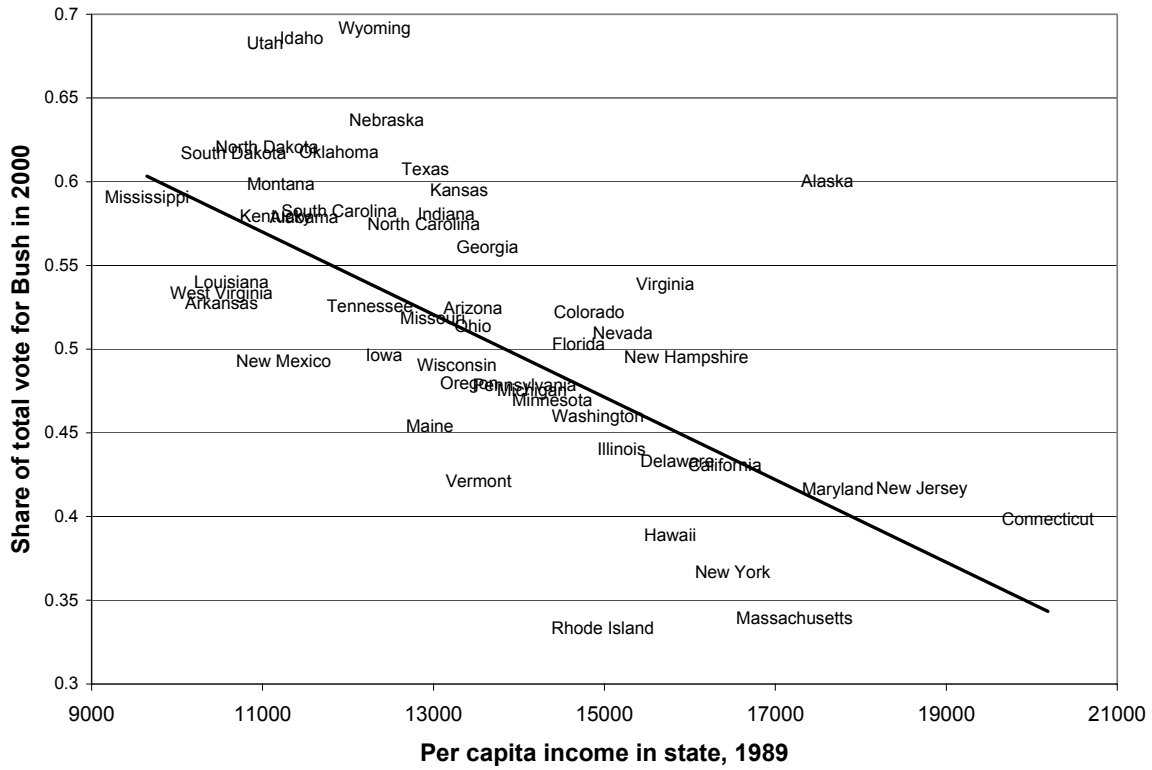
**Figure 2** *Church attendance and Republicanism*

Notes: Data are from National Election Study Cumulative File. Data shown are for years 1988-2000 and include only white males between the ages of 20 and 55 who voted for a Democratic or Republican candidate in the previous presidential election..

**Figure 3** *State income and Republicanism in the 2000 election*

Notes: Data on per capita income are from the US Census. Figure excludes Washington, DC.

Figure 4 *Trends in the determinants of voting Republican*
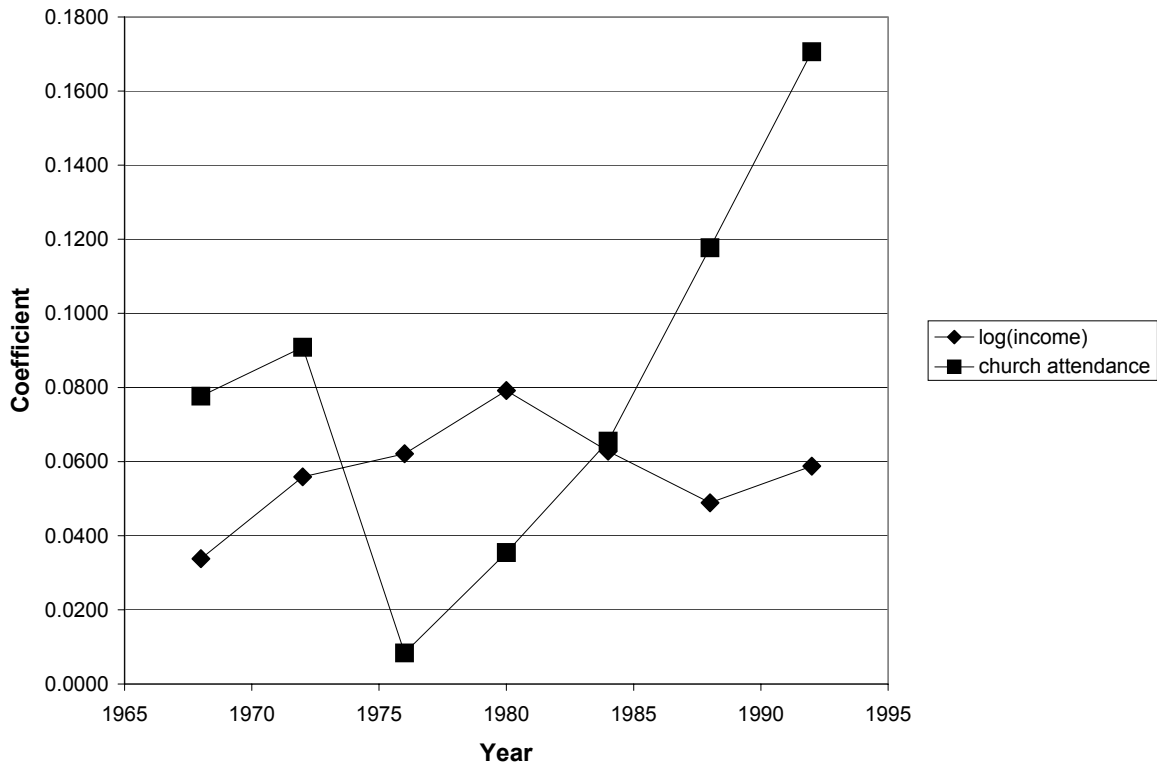
Notes: Data are from the General Social Survey. Data reflect marginal effects evaluated at sample means from probit models of the propensity to vote Republican. Sample includes respondents who voted for a Democratic or Republican candidate in the previous presidential election. All regressions include controls for years of schooling, age, age^2, race, gender, and a dummy for missing income data.

Figure 5 *Ideology and voter turnout*
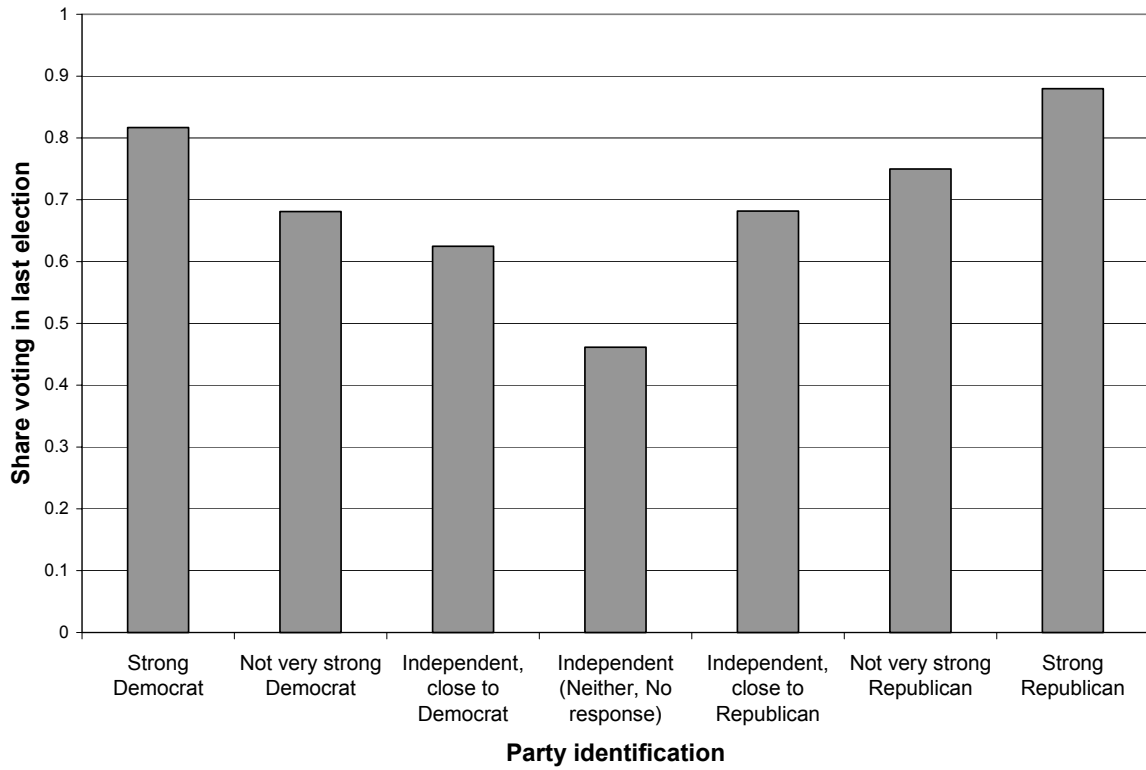
Notes: Data are from the General Social Survey 1972-2000 cumulative file. Data reflect shares of respondents in each group participating in the most recent presidential election.

Figure 6 *The voting decision*

Notes: The horizontal axis depicts the range of policies along dimension T and the vertical axis depicts the range of policies along dimension A. The beliefs about the two parties platforms are described by a small circle for the right-wing candidate and a small box for the left-wing candidate. The line running through the origin is the line described by the values of t and a for which the benefit function, $\Delta(t, a; \Omega) = 0$. All citizens with zero voting costs will support the R candidate if their ideal policies (i.e. (a, t) values) are above that line. All citizens with zero voting costs will support the L candidate if their ideal policies are below that line. The line that runs above the origin reflects the values of t and a for which the benefit function, $\Delta(t, a; \Omega)$ equals some positive value $\frac{c}{2}$. The line below the origin reflects the values of t and a for which the negative of the benefit function equals $\frac{c}{2}$. These two lines describe the voting behavior of an individual whose cost of voting equals c. If the individual's preferences lie above the upper line, then that citizen will vote for the R candidate. If the individual's preferences lie below the lower line, then that citizen will vote for the L candidate. If the individual's preferences lie between the lines, then the individual will abstain.

Figure 7 *Group size and extremism*

Notes: Graph presents the relationship between group size and extremism (party deviation from the center).
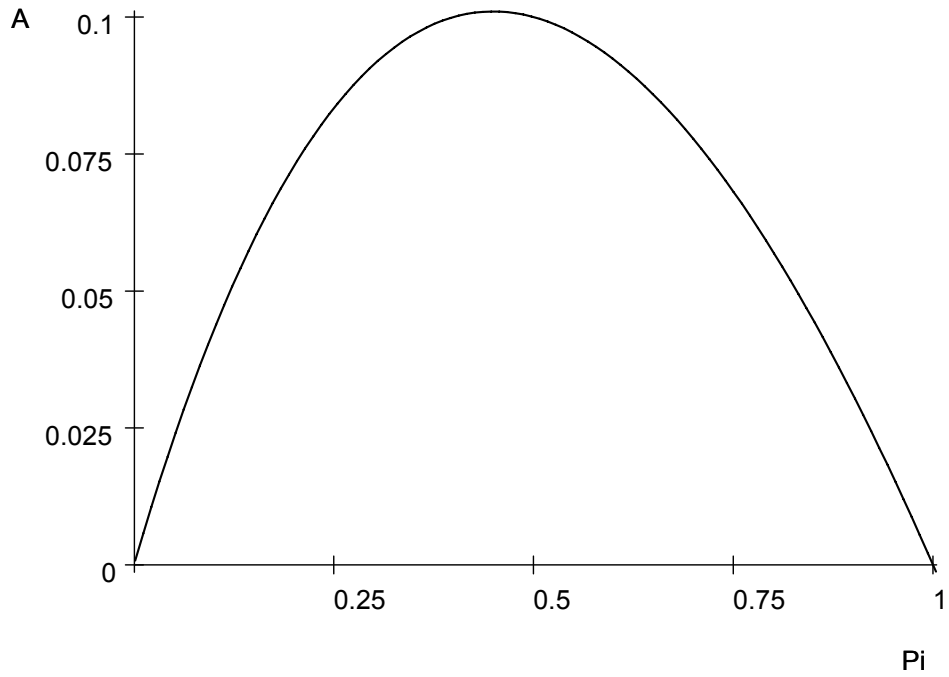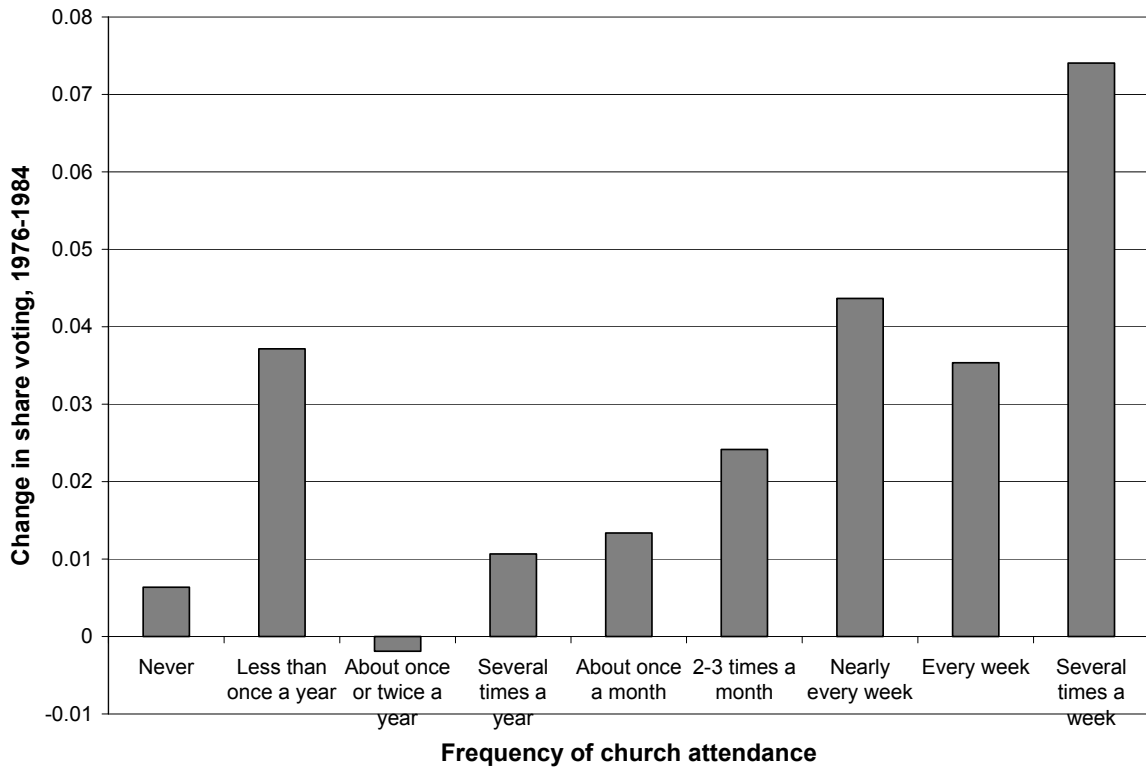
Figure 8 *Religion and voter turnout, 1976-1984*

Notes: Data are from the General Social Survey 1972-2000 cumulative file. Data reflect shares of respondents in each group participating in the most recent presidential election.

Figure 9 *The political role of religion across countries*
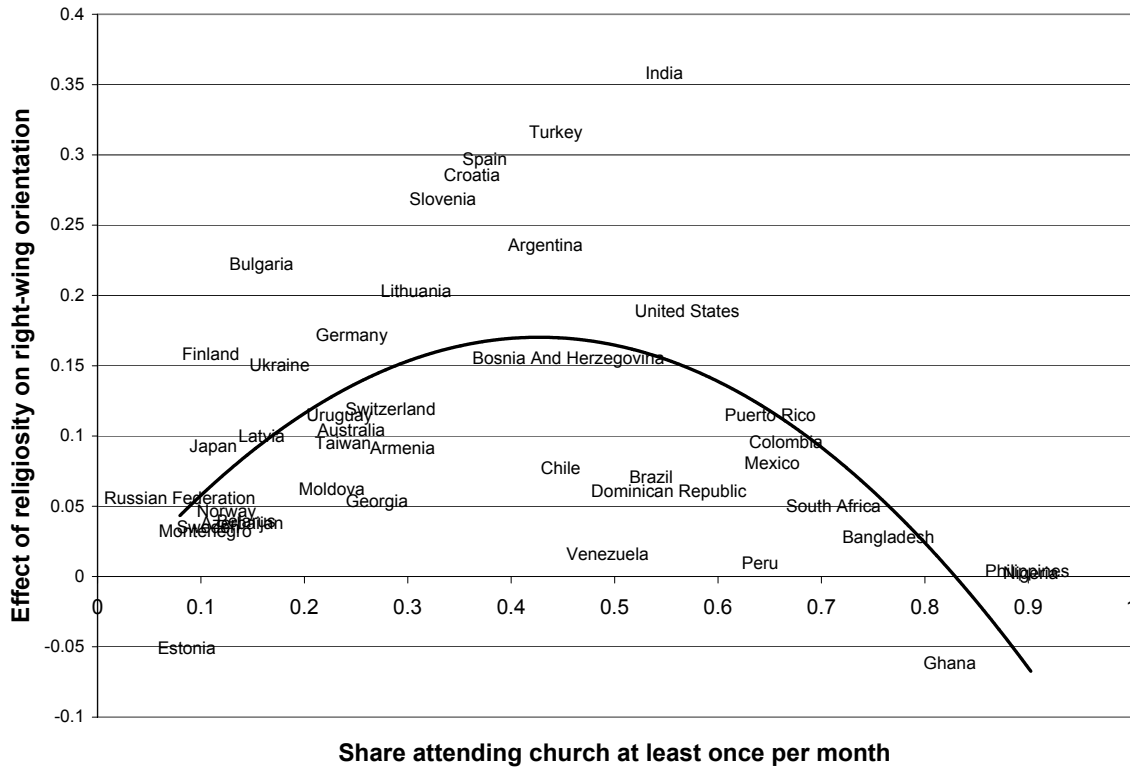
Notes: Data are from Wave 3 of the World Values Survey. Vertical axis shows the estimated effect of a one standard deviation increase in the importance of God in the respondent's life on the respondent's identification with the political right. Horizontal axis shows the share of respondents in the country who report attending church (or analogous religious institution) once a month or more.

Figure 10 *Unions and the political importance of income across countries*

Notes: Data are from Wave 2 of the World Values Survey. Vertical axis shows the estimated effect of a one standard deviation increase in income on the respondent's identification with the political right. Horizontal axis shows Blanchflower's (1996) measure of trade union density in 1994.

Figure 11 *The political role of religion across US states*

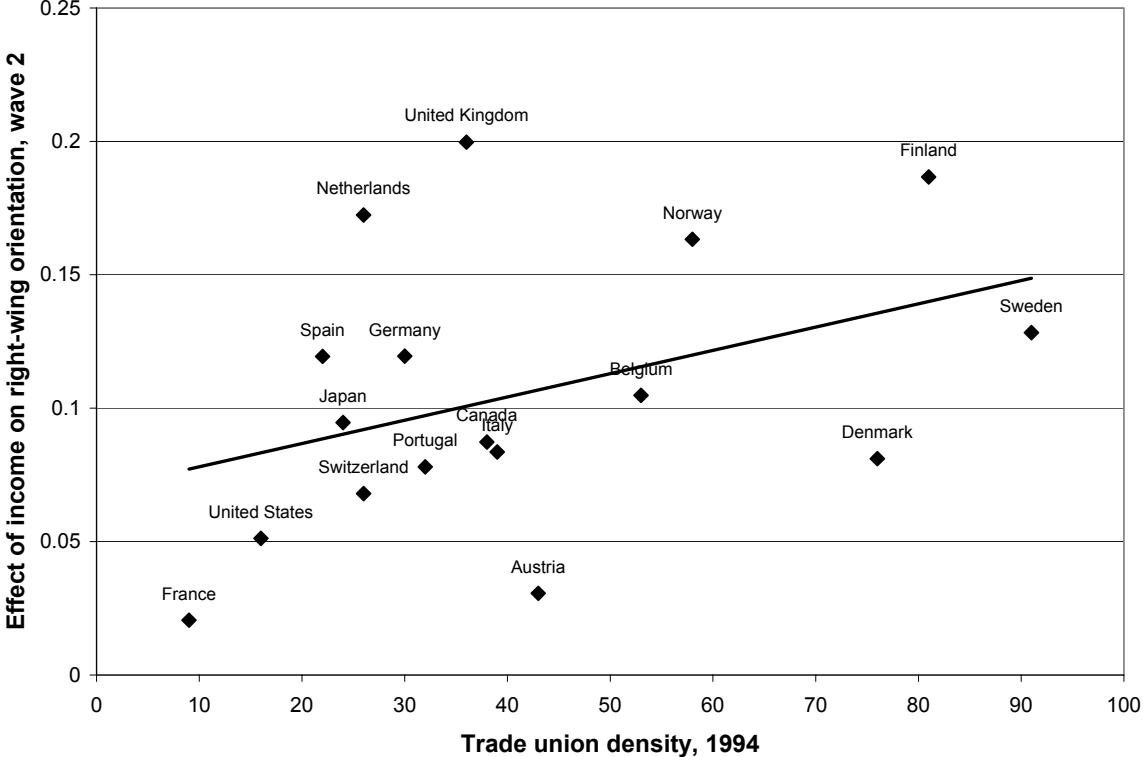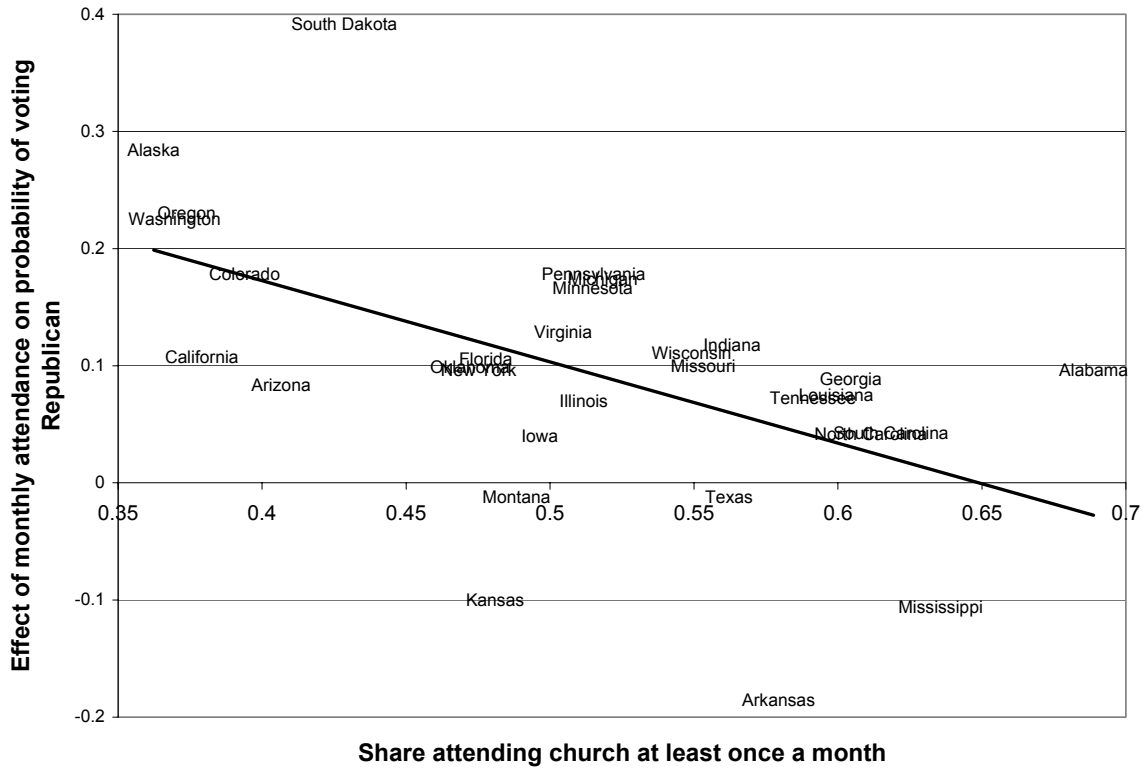Notes: Data are from General Social Survey cumulative file, 1972-2000. Only respondents voting for either Democrat or Republican in previous election have been included. Vertical axis shows the marginal effect of monthly church attendance from probit models evaluated at sample means. The dependent variable is a dummy for Republican vote in the previous presidential election and all specifications are run separately for each state, including dummies for year of most recent election, race, gender, and controls for the log of real income, age, age^2, years of completed schooling, and a dummy for missing income data. Sample excludes Utah residents.

Table 1 *Party effects on policy outcomes, 1970-2000*

| | (1) | (2) |
White House party	Abortions per 1000 live births	Income tax as % of GDP
Democrat	312.7	11.22
$(n = 12)$	(11.08)	(0.2149)
Republican	294.4	10.04
$(n = 19)$	(20.81)	(0.1285)
Difference	17.25	1.175
$(n = 31)$	(23.58)	(0.2504)

Notes: Standard errors in parentheses. Standard errors allow unequal variances by party.
Data on abortions per 1000 live births from CDC Abortion Surveillance, http://www.cdc.gov/reproductiveh
Data on income tax as % of GDP from Center on Budget and Policy Priorities, taken from
Congressional Budget Office, http://www.cbpp.org/10-21-03tax.htm.

Table 2 *Cross-country determinants of right-wing orientation*

Dependent variable: Respondent's identification with the political right (standardized)

	(1)	(2)	(3)	(4)	(5)	(6)
Importance of God in life (standardized)	0.1658 (0.0147)					
Share in country-wave attending church monthly		-0.1336 (0.1071)	-0.01431 (0.2117)	-0.1315 (0.1082)		
(Share monthly)^2		0.0262 (0.1550)	-0.1178 (0.2796)	0.0259 (0.1585)		
Importance of God $\times$ Share monthly		0.8173 (0.1885)	0.7718 (0.3235)	1.2090 (0.3082)		
Importance of God $\times$ (Share monthly)^2		-0.9030 (0.2083)	-0.8046 (0.4007)	-1.6197 (0.3662)		
Imp. of God $\times$ wave?	NO	YES	YES	YES	–	–
Imp. of God $\times$ country?	NO	NO	NO	YES	–	–
Income (standardized)					0.0921 (0.0140)	
Trade union density in country-wave						-0.0007 (0.0013)
Income $\times$ trade union density						0.0013 (0.0004)
Income $\times$ wave?	–	–	–	–	NO	YES
Income $\times$ country?	–	–	–	–	NO	NO
Sample	All	All	Democracies	All	OECD	OECD
N	114454	114454	68837	114454	37075	37075

Notes: Data are from World Values Survey cumulative file, waves 1-3. All estimates are from OLS regressions. Standard errors in parentheses are adjusted for correlation within country of residence. All specifications include dummies for country of residence, survey wave, gender, and controls for income, age, age^2, years of completed schooling, and dummies for missing data on these controls.

Table 3 *Church attendance and the role of religion in voting (GSS)*

Dependent variable: Did respondent vote for Republican candidate in last presidential election?

	(1)	(2)	(3)	(4)
Attend church at least once a month	0.0995 (0.0113)			
Share in state-election attending at least once a month		0.2951 (0.1141)	0.1958 (0.1647)	0.3944 (0.1258)
Attend at least once a month × Share attending at least once a month		-0.3511 (0.1144)	-0.1647 (0.2068)	
Church attendance frequency × Share attending at least once a month				-0.0706 (0.0205)
Attend at least once a month × election?	NO	YES	YES	YES
Attend at least once a month × state?	NO	NO	YES	NO
N	16298	16298	16298	16298

Notes: Data are from General Social Survey cumulative file, 1972-2000. Only respondents voting for either Democrat or Republican in previous election have been included. All estimates are marginal effects from probit models evaluated at sample means. Standard errors in parentheses are adjusted for correlation within state of residence. All specifications include dummies for state of residence, year of most recent election, race, gender, and controls for the log of real income, age, age^2, years of completed schooling, and dummy for missing income data. Samples exclude Utah residents.

Table 4 *Church attendance and the role of religion in voting (NES)*

Dependent variable: Did respondent vote for Republican candidate in last presidential election?

	(1)	(2)	(3)	(4)
Attend church at least once a month	0.1237 (0.0115)			
Share in state-election attending at least once a month		0.4829 (0.1383)	0.5026 (0.1426)	0.6890 (0.1911)
Attend at least once a month × Share attending at least once a month		-0.3419 (0.1464)	-0.3820 (0.1756)	
Church attendance frequency × Share attending at least once a month				-0.1303 (0.0508)
Attend at least once a month × election?	NO	YES	YES	YES
Attend at least once a month × state?	NO	NO	YES	NO
N	8882	8882	8882	8882

Notes: Data are from National Election Study cumulative file, 1948-2000. Only years 1972-2000 are used due to a change in the survey question about religious attendance beginning in 1970. Only respondents voting for either Democrat or Republican in previous presidential election have been included. All estimates are marginal effects from probit models evaluated at sample means. Standard errors in parentheses are adjusted for correlation within state of residence. All specifications include dummies for state of residence, year of most recent election, race, gender, income category (as shown in Figure 1), and controls for the age, age^2, and years of completed schooling.

Table 5 *Explaining time trends in the determinants of voting*

Election year	(1) Effect of monthly attendance on voting Republican	(2) Share attending monthly
1968	0.0740	56.8%
1992	0.1627	48.5%
Change	0.0887	-8.3%

Notes: Data are from General Social Survey cumulative file, 1972-2000. Only respondents voting for either Democrat or Republican in previous election have been included. All estimates are marginal effects from probit models evaluated at sample means. All specifications include dummies for state of residence, race, gender, and controls for the log of real income, age, age^2, years of completed schooling, and dummy for missing income data. Samples exclude Utah residents.

A   Appendix

A.1   Proof of Proposition 1 and Corollary 1

Consider without loss of generality party R: letting $f(x)$ be the density of voter having ideal preference x, its margin of victory equals

$$\int_{\frac{x_L+x_R}{2}}^{1} Z\left((1-\gamma)B + \gamma(x_L-x)^2 - (x_R-x)^2\right) f(x)\,dx +$$

$$-\int_{-1}^{\frac{x_L+x_R}{2}} Z\left((1-\gamma)B + \gamma(x_R-x)^2 - (x_L-x)^2\right) f(x)\,dx$$

whose derivative with respect to x_R is

$$-2\int_{\frac{x_L+x_R}{2}}^{1} z\left((1-\gamma)B + \gamma(x_L-x)^2 - (x_R-x)^2\right)(x_R-x)f(x)\,dx +$$

$$-2\gamma\int_{-1}^{\frac{x_L+x_R}{2}} z\left((1-\gamma)B + \gamma(x_R-x)^2 - (x_L-x)^2\right)(x_R-x)f(x)\,dx +$$

$$-Z\left((1-\gamma)\left[B - \left(\frac{x_R-x_L}{2}\right)^2\right]\right) f\left(\frac{x_L+x_R}{2}\right)$$

We naturally assume that a right-wing voter will always derive a psychological benefit from supporting any right-wing candidate against any left-wing opponent, formally that

$$(1-\gamma)B + \gamma(x_L-x)^2 - (x_R-x)^2 \geq 0 \forall x \geq \frac{x_L+x_R}{2}\forall x_R > x_L$$

$$(1-\gamma)\left[B - \left(\frac{x_R-x_L}{2}\right)^2\right] \geq 0 \forall x_R > x_L$$

$$\gamma < 1 \Rightarrow B \geq 1 \wedge \gamma > 1 \Rightarrow B \leq 0$$

so that by the uniform distribution of the cost of voting this reduces to

$$-\frac{2}{\overline{V}}\int_{\frac{x_L+x_R}{2}}^{1}(x_R-x)f(x)\,dx - \frac{2\gamma}{\overline{V}}\int_{-1}^{\frac{x_L+x_R}{2}}(x_R-x)f(x)\,dx - \frac{1-\gamma}{\overline{V}}\left[B - \left(\frac{x_R-x_L}{2}\right)^2\right]f\left(\frac{x_L+x_R}{2}\right)$$

It is immediate that this is always negative if voters are mainly motivated by a taste for

opposition, namely if $\gamma > 1$ (and $B \leq 0$), since then

$$-2 \int_{\frac{x_L+x_R}{2}}^{1} (x_R - x) f(x) \, dx - 2\gamma \int_{-1}^{\frac{x_L+x_R}{2}} (x_R - x) f(x) \, dx - (1 - \gamma) \left[B - \left(\frac{x_R - x_L}{2} \right)^2 \right] f\left(\frac{x_L + x_R}{2} \right)$$

$$< -2 \int_{\frac{x_L+x_R}{2}}^{1} (x_R - x) f(x) \, dx - 2\gamma \int_{-1}^{\frac{x_L+x_R}{2}} (x_R - x) f(x) \, dx =$$

$$= -2 \left[\gamma F \left(\frac{x_L + x_R}{2} \right) + \left(1 - F \left(\frac{x_L + x_R}{2} \right) \right) \right] x_R + 2 \left[\gamma \int_{-1}^{\frac{x_L+x_R}{2}} x f(x) \, dx + \int_{\frac{x_L+x_R}{2}}^{1} x f(x) \, dx \right] \leq$$

$$\leq -2 \left[\gamma F \left(\frac{x_L + x_R}{2} \right) + \left(1 - F \left(\frac{x_L + x_R}{2} \right) \right) \right] x_R < 0 \, \forall x_R > 0$$

which implies that the optimal platform choice is represented by a corner solution, with the party converging to the position of the median voter, namely the origin; which, symmetrically, is also the equilibrium strategy for party L.

The median-voter result still obtains when $\gamma = 1$, since the first-order condition for the optimal platform choice is then

$$\int_{-1}^{1} (x_R - x) f(x) \, dx = 0 \Rightarrow x_R = E(x) = 0$$

regardless of the opponent's strategy (which will be identical).

If instead voters are keener on supporting a party than on opposing its rival, i.e. $\gamma < 1$ (and $B \geq 1$), extremism may emerge: a symmetric equilibrium such that $x_R = -x_L > 0$ is defined by the condition

$$-2 \int_{0}^{1} (x_R - x) f(x) \, dx - 2\gamma \int_{-1}^{0} (x_R - x) f(x) \, dx - (1 - \gamma) \left(B - x_R^2 \right) f(0) = 0$$

$$(1 - \gamma) f(0) x_R^2 - (1 + \gamma) x_R + (1 - \gamma) (E(|x|) - B f(0)) = 0$$

so long as voters' preferences are heterogeneous enough for the density of voters with exactly median preferences to be small compared to the mean deviation of the distribution:

$$f(0) < \frac{E(|x|)}{B} \leq E(|x|)$$

Then the derivative is a convex quadratic function $Q(x)$ such that $Q(0) > 0 \wedge Q(1) < 0$, which implies that the first-order condition for a maximum is satisfied at the unique value $x_R \in (0, 1) : Q(x_R) = 0$, namely

$$x_R = \frac{(1 + \gamma) - \sqrt{(1 + \gamma)^2 - 4(1 - \gamma)^2 f(0) (E(|x|) - B f(0))}}{2(1 - \gamma) f(0)} > 0$$

so that

$$\frac{\partial x_R}{\partial E\left(|x|\right)} = \frac{(1-\gamma)}{\sqrt{(1+\gamma)^2 - 4(1-\gamma)^2 f(0)(E(|x|)-Bf(0))}} > 0$$

while

$$\frac{\partial x_R}{\partial \gamma} = \frac{\left[1 - \frac{(1+\gamma)+4(1-\gamma)f(0)(E(|x|)-Bf(0))}{\sqrt{(1+\gamma)^2-4(1-\gamma)^2 f(0)(E(|x|)-Bf(0))}}\right](1-\gamma)}{2(1-\gamma)^2 f(0)} +$$

$$+ \frac{\left[(1+\gamma) - \sqrt{(1+\gamma)^2 - 4(1-\gamma)^2 f(0)(E(|x|)-Bf(0))}\right]}{2(1-\gamma)^2 f(0)} < 0$$

because

$$\frac{\partial x_R}{\partial \gamma} < 0 \Leftrightarrow \sqrt{(1+\gamma)^2 - 4(1-\gamma)^2 f(0)(E(|x|)-Bf(0))} < (1+\gamma)$$

which is always satisfied when $E\left(|x|\right) > Bf\left(0\right)$.

If $x \le y \Leftrightarrow f\left(x\right) \ge f\left(y\right) \forall x, y \in [0,1]$ it follows that the uniform distribution is a limiting case, hence

$$f\left(0\right) \ge \frac{1}{2} \wedge \int_0^1 x f\left(x\right) dx \le \frac{1}{4}$$

so that

$$f\left(0\right) \ge 2 \int_0^1 x f\left(x\right) dx = E\left(|x|\right)$$

and therefore the optimal policy choice will be the corner solution with the party locating in the origin.

A.2   Proof of Propositions 2, 3 and 4

Denoting as $\Omega = (\hat{x}_L, \hat{x})$ the position of the parties as perceived by a given voter and as $\Theta\left(\Omega\right)$ the distribution of beliefs, the margin of victory for party R is

$$\int_{-1}^1 \int_\Omega \left(Z\left(2\left(\hat{x}_R - \hat{x}_L\right)\left(x - \frac{\hat{x}_R + \hat{x}_L}{2}\right)\right) - Z\left(-2\left(\hat{x}_R - \hat{x}_L\right)\left(x - \frac{\hat{x}_R + \hat{x}_L}{2}\right)\right)\right) d\Theta\left(\Omega\right) f\left(x\right) dx$$

whose derivative with respect to x_R is

$$\int_{-1}^1 2\left(x - x_R\right) \int_\Omega z\left(2\left|\left(x_R - \hat{x}_L\right)\left(x - \frac{x_R + \hat{x}_L}{2}\right)\right|\right) \chi\left(\hat{x}_R = x_R\right) d\Theta\left(\Omega\right) f\left(x\right) dx$$

65

where we denote with $\chi\left(.\right)$ the characteristic or indicator function.

Thus, considering the uniform distribution of costs and the assumptions on the availability of updated information on the party platform, and letting the probability of a voter being an affiliate be described by the function $\pi\left(x\right)$ of his ideal policy, the optimal platform choice for the party, regardless of initial beliefs and of the opponent's actions, is

$$x_R = \frac{\left(\bar{\theta} - \theta\right) \int_{-1}^{1} x\pi\left(x\right) f\left(x\right) dx}{\theta + \left(\bar{\theta} - \theta\right) \int_{-1}^{1} \pi\left(x\right) f\left(x\right) dx}$$

Notice that

$$\Pi_R = \int_{-1}^{1} \pi\left(x\right) f\left(x\right) dx \in [0, 1]$$

equals the size of the group of party affiliates, which can also be interpreted as the unconditional probability of a voter being an affiliate; while the ideal policy of the average party affiliate is

$$E\left(x|R\right) = \frac{1}{\Pi_R} \int_{-1}^{1} x\pi\left(x\right) f\left(x\right) dtx \geq E\left(x\right) = 0$$

where the inequality is derived from the natural assumption that R-party affiliates are on average more conservative than voters as a whole.

Therefore

$$x_R = \frac{\left(\bar{\theta} - \theta\right) \Pi_R}{\theta + \left(\bar{\theta} - \theta\right) \Pi_R} E\left(x|R\right)$$

It follows immediately that, when $\Pi_R > 0$, the median-voter result obtains in two cases only: in the absence of informational asymmetries: $\bar{\theta} = \theta$; or in the absence of any difference in the preferences of the average affiliate and non-affiliate: $E\left(x|R\right) = E\left(x\right) = 0$. Whenever there are both an informational and an ideological difference between affiliates and non-affiliates $x_R > 0$.

Furthermore, the party's extremism is then monotone increasing in its ability to convey information to its affiliates

$$\frac{\partial x_R}{\partial \bar{\theta}} = \frac{\theta\Pi_R}{\left[\theta + \left(\bar{\theta} - \theta\right) \Pi_R\right]^2} E\left(x|R\right) \geq 0$$

and to withhold it from non-affiliates

$$\frac{\partial x_R}{\partial \theta} = \frac{-\theta\Pi_R}{\left[\theta + \left(\bar{\theta} - \theta\right) \Pi_R\right]^2} E\left(x|R\right) \leq 0$$

as well as increasing in the number of affiliates (keeping their average conservatism constant)

$$\frac{\partial x_R}{\partial \Pi_R} = \frac{\theta\left(\bar{\theta} - \theta\right)}{\left[\theta + \left(\bar{\theta} - \theta\right) \Pi_R\right]^2} E\left(x|R\right) \geq 0$$

and in the average conservatism of affiliates (keeping their number constant)

$$\frac{\partial x_R}{\partial E\left(x|R\right)} = \frac{\left(\bar{\theta} - \theta\right)\Pi_R}{\theta + \left(\bar{\theta} - \theta\right)\Pi_R} \geq 0$$

If $\pi\left(x\right) = \chi\left(x > 0\right)$, $\Pi_R = \frac{1}{2}$ and $E\left(x|R\right) = E\left(|x|\right)$, so that

$$x_R = \frac{\bar{\theta} - \theta}{\bar{\theta} + \theta}E\left(|x|\right)$$

which immediately shows that extremism is increasing in the heterogeneity of voters' preferences, measured again by the mean deviation of their distribution.

When the distribution of the cost of voting is generalized to include a point mass $z_0 \in [0, 1]$ of voters with zero cost of voting as well as a uniform density of voters $z\left(c\right) = \frac{1-z_0}{\bar{V}}\forall c \in \left[0, \bar{V}\right]$, the margin of victory for party R is

$$\frac{1 - z_0}{2}\int_{-1}^{1}\int_{\Omega}\left(\hat{x}_R - \hat{x}_L\right)\left(x - \frac{\hat{x}_R + \hat{x}_L}{2}\right)d\Theta\left(\Omega\right)f\left(x\right)dx + z_0\left(1 - 2\int_{\Omega}F\left(\frac{\hat{x}_R + \hat{x}_L}{2}\right)d\Theta\left(\Omega\right)\right)$$

whose derivative with respect to x_R is

$$\frac{1 - z_0}{2}\int_{-1}^{1}\left(x - x_R\right)\left[\theta + \left(\bar{\theta} - \theta\right)\pi\left(x\right)\right]f\left(x\right)dx - z_0\int_{\Omega}f\left(\frac{x_R + \hat{x}_L}{2}\right)\chi\left(\hat{x}_R = x_R\right)d\Theta\left(\Omega\right)$$

As shown above, the first term is a decreasing function of x_R that equals zero when

$$x_R = \frac{\left(\bar{\theta} - \theta\right)\Pi_R}{\theta + \left(\bar{\theta} - \theta\right)\Pi_R}E\left(x|R\right) > 0$$

under the weak assumption that R affiliates are on average more conservative than non-affiliates. The second term, instead, is always negative. Hence the modified distribution generates less extremism than the uniform.

In the limit, as $z_0 \to 0$ we have our standard result $x_R = \frac{\left(\bar{\theta} - \theta\right)\Pi_R}{\theta + \left(\bar{\theta} - \theta\right)\Pi_R}E\left(x|R\right)$ and as $z_0 \to 1$ convergence to the mean as a corner solution when voters' beliefs are rational.

By the second-order condition, the second derivative of the margin of victory with respect to x_R is negative; since the cross derivative with respect to z_0 is negative by inspection

$$\frac{\partial x_R}{\partial z_0} < 0$$

A.3 Proof of Propositions 5 and 6

The margin of victory for party R is

$$\int_{-1}^{1} \int_{-1}^{1} \int_{\Omega} \left(Z\left(2\Delta\left(t,a;\Omega\right)\right) - Z\left(-2\Delta\left(t,a;\Omega\right)\right) \right) d\Theta\left(\Omega\right) f\left(t,a\right) dtda$$

whose gradient with respect to the platform choice of party R, (T_R, A_R), is

$$\left(\begin{array}{c} 2\mu \int_{-1}^{1} \int_{-1}^{1} \left(t - T_R\right) \int_{\Omega} z\left(2\left|\Delta\left(t,a;\Omega\right)\right|\right) \chi\left(\hat{T}_R = T_R\right) d\Theta\left(\Omega\right) f\left(t,a\right) dtda \\ 2\left(1-\mu\right) \int_{-1}^{1} \int_{-1}^{1} \left(a - A_R\right) \int_{\Omega} z\left(2\left|\Delta\left(t,a;\Omega\right)\right|\right) \chi\left(\hat{A}_R = A_R\right) d\Theta\left(\Omega\right) f\left(t,a\right) dtda \end{array} \right)$$

Considering our usual assumptions on the distribution of costs and preferences and on the informational structure, the optimal platform choice for the party, regardless of initial beliefs and of the opponent's actions, is

$$\begin{cases} T_R = \frac{\left(\bar{\theta}-\theta\right)\Pi_R}{\theta+\left(\bar{\theta}-\theta\right)\Pi_R} E\left(t|R\right) \\ A_R = \frac{\left(\bar{\theta}-\theta\right)\Pi_R}{\theta+\left(\bar{\theta}-\theta\right)\Pi_R} E\left(a|R\right) \end{cases}$$

where

$$\Pi_R = \int_{-1}^{1} \int_{-1}^{1} \pi\left(t,a\right) f\left(t,a\right) dtda \in [0,1]$$

is the size of the group of party affiliates (or the unconditional probability of a voter being an affiliate), and by the assumption that R-party affiliates are on average at least as conservative as voters as a whole on each of the two dimensions

$$\begin{cases} E\left(t|R\right) = \frac{1}{\Pi_R} \int_{-1}^{1} \int_{-1}^{1} t\pi\left(t,a\right) f\left(t,a\right) dtda \geq 0 \\ E\left(a|R\right) = \frac{1}{\Pi_R} \int_{-1}^{1} \int_{-1}^{1} a\pi\left(t,a\right) f\left(t,a\right) dtda \geq 0 \end{cases}$$

Mirroring the one-dimensional case, this shows that convergence to the position of the median voter for a party with a positive mass of affiliates ($\Pi_R > 0$) requires the absence either of informational asymmetries ($\bar{\theta} = \theta$) or of any difference in the average preferences of affiliates and non-affiliates ($E\left(t|R\right) = E\left(t\right) = 0 \wedge E\left(a|R\right) = E\left(a\right) = 0$). The effect of the different variables is also identical; moreover

$$T_R \lesseqgtr A_R \Leftrightarrow E\left(t|R\right) \lesseqgtr E\left(a|R\right)$$

which shows that, unsurprisingly, the divergence of the party from the median-voter position is greater on the issue where the difference in preferences between affiliates and non-affiliates is on average greater.

If $\pi(t,a) = \chi(t+a > 0)$, considering that by independence $f(t,a) = f_T(t) f_A(a)$ and thus by symmetry $f(t,a) = f(-t,a) = f(t,-a) = f(-t,-a)$

$$\Pi_R = \int_{-1}^1 \int_{-t}^1 f(t,a)\,da\,dt = \int_{-1}^1 F_A(t) f_T(t)\,dt = \frac{1}{2}$$

and

$$E(t|R) = 2 \int_{-1}^1 t \int_{-t}^1 f_A(a)\,da\,f_T(t)\,dt = 2 \int_{-1}^1 t F_A(t) f_T(t)\,dt =$$
$$= 2 \int_0^1 t f_T(t)(2F_A(t) - 1)\,dt \in (0, E(|t|))$$

We say that the distribution $G(x)$ is monotonically more heterogeneous that $F(x)$ if and only if $g(x) - f(x)$ is monotone non-decreasing on $[0,1]$ (and thus monotone non-increasing on $[-1,0]$), or equivalently $G(x) - F(x)$ is convex on $[0,1]$ (and thus concave on $[-1,0]$), or equivalently (since by symmetry $F(0) = G(0) = \frac{1}{2}$ while $F(1) = G(1) = 1$) $G(x) \le F(x) \,\forall x \in [0,1]$ (and thus $G(x) \ge F(x) \,\forall x \in [-1,0]$).

Therefore, a monotonic increase in the heterogeneity of F_T (weakly) increases $E(t|R)$ because $t(2F_A(t) - 1)$ is (weakly) increasing on $[0,1]$, while it (weakly) reduces $E(a|R)$ because $t f_T(t)$ is (weakly) positive on $[0,1]$.

A.4   Proof of Proposition 7

Consider without loss of generality party R, for which $\pi_R(t,a) = \chi(a > \alpha)$: we have $\Pi_R = 1 - F(\alpha)$, $E(t|R) = 0$ and $E(a|R) = \frac{1}{1-F_A(\alpha)} \int_\alpha^1 a f_A(a)\,da \ge 0$ so that

$$\begin{cases} T_R = 0 \\ A_R = \frac{1}{\frac{\theta}{\bar\theta - \underline\theta} + 1 - F(\alpha)} \int_\alpha^1 a f_A(a)\,da \ge 0 \end{cases}$$

Notice that, obviously, $\alpha = \mp 1 \Rightarrow A_R = 0$, because an organization comprising everyone or nobody has no relevance as a channel for informational differentiation; moreover

$$\frac{\partial A_R}{\partial \alpha} = \frac{f_A(\alpha)}{\left(\frac{\theta}{\bar\theta - \underline\theta} + 1 - F_A(\alpha)\right)^2} \left\{ \int_\alpha^1 a f_A(a)\,da - \alpha \left(\frac{\theta}{\bar\theta - \underline\theta} + 1 - F_A(\alpha)\right) \right\}$$

so that

$$\frac{\partial A_R}{\partial \alpha} > 0 \forall \alpha \le 0$$

and the maximum effect of the organization on extremism is attained at $\alpha^* > 0$ such that

$$1 - \int_{\alpha^*}^1 F_A(a)\,da = \alpha^* \left(\frac{\theta}{\bar\theta - \underline\theta} + 1\right)$$

69

The left-hand side of this equation is an increasing convex function $L(\alpha)$ such that $L(1) = 1$ and $L'(\alpha) = F_A(\alpha) \in \left[\frac{1}{2}, 1\right] \forall \alpha \in [0,1]$, while the right-hand side is a line through the origin with slope greater than 1, which ensures uniqueness of the maximum. Since $F_A(0) = \frac{1}{2} \Rightarrow \int_\alpha^1 F_A(a)\, da > \frac{1-\alpha}{2} \forall a \geq 0$ we also know that

$$\alpha^* < \frac{\bar{\theta} - \theta}{\bar{\theta} + \theta}$$

and indeed

$$\frac{\partial \alpha^*}{\partial \bar{\theta}} = \frac{\alpha^* \theta}{(\bar{\theta} - \theta)\left[\theta + (\bar{\theta} - \theta)(1 - F_A(\alpha^*))\right]} \geq 0$$

$$\frac{\partial \alpha^*}{\partial \theta} = -\frac{\alpha^* \bar{\theta}}{(\bar{\theta} - \theta)\left[\theta + (\bar{\theta} - \theta)(1 - F_A(\alpha^*))\right]} < 0$$

A.5 Proof of Proposition 9

Recall that the margin of victory for party R is

$$\frac{1}{2} \int_{-1}^1 \int_{-1}^1 \int_\Omega \Delta(t, a; \Omega)\, d\Theta(\Omega) f(t, a)\, da\, dt$$

whose gradient with respect to the platform choice of party R, (T_R, A_R), is

$$\left(\begin{array}{c} \frac{\mu}{2} \int_{-1}^1 \int_{-1}^1 (t - T_R) \int_\Omega \chi\left(\hat{T}_R = T_R\right) d\Theta(\Omega) f(t, a)\, dt\, da \\ \frac{1-\mu}{2} \int_{-1}^1 \int_{-1}^1 (a - A_R) \int_\Omega \chi\left(\hat{A}_R = A_R\right) d\Theta(\Omega) f(t, a)\, dt\, da \end{array} \right)$$

so that the first-order condition for a maximum is

$$\begin{cases} \int_{-1}^1 \int_{-1}^1 (t - T_R) \int_\Omega \chi\left(\hat{T}_R = T_R\right) d\Theta(\Omega) f(t, a)\, dt\, da = 0 \\ \int_{-1}^1 \int_{-1}^1 (a - A_R) \int_\Omega \chi\left(\hat{A}_R = A_R\right) d\Theta(\Omega) f(t, a)\, dt\, da = 0 \end{cases}$$

and a symmetric political equilibrium is defined by

$$\begin{cases} \int_{-1}^1 \int_{-1}^1 (t - \tau) \left[\bar{\theta}\chi\left(\lambda\tau t + (1-\lambda)\alpha a > 0\right) + \theta\chi\left(\lambda\tau t + (1-\lambda)\alpha a < 0\right)\right] f(t, a)\, dt\, da = 0 \\ \int_{-1}^1 \int_{-1}^1 (a - \alpha) \left[\bar{\theta}\chi\left(\lambda\tau t + (1-\lambda)\alpha a > 0\right) + \theta\chi\left(\lambda\tau t + (1-\lambda)\alpha a < 0\right)\right] f(t, a)\, dt\, da = 0 \end{cases}$$

If $\tau = 0 \wedge \alpha > 0$ this condition simplifies to the one-dimensional optimum

$$\alpha = \frac{\bar{\theta} - \theta}{\bar{\theta} + \theta} E(|a|)$$

and analogously if $\tau > 0 \wedge \alpha = 0$

$$\tau = \frac{\bar{\theta} - \theta}{\bar{\theta} + \theta} E(|t|)$$

Therefore, these two degenerate equilibria exist for any distribution of preferences that satisfies the assumption of independence on the two dimensions.

A.6 Proof of Proposition 10

When preferences are uniformly distributed on both issues so that $f(t,a) = \frac{1}{4}\chi\left(t, a \in [-1,1]^2\right)$ and initial beliefs on the platforms of the two parties are symmetric around the origin, the first-order condition becomes

$$\begin{cases} \int_{-1}^1 \int_{-1}^1 (t - T_R)\left[\bar\theta\chi\left(\lambda\bar T_R t + (1-\lambda)\bar A_R a > 0\right) + \underline\theta\chi\left(\lambda\bar T_R t + (1-\lambda)\bar A_R a < 0\right)\right] da\, dt = 0 \\ \int_{-1}^1 \int_{-1}^1 (a - A_R)\left[\bar\theta\chi\left(\lambda\bar T_R t + (1-\lambda)\bar A_R a > 0\right) + \underline\theta\chi\left(\lambda\bar T_R t + (1-\lambda)\bar A_R a < 0\right)\right] da\, dt = 0 \end{cases}$$

Suppose that $\lambda\bar T_R \leq (1-\lambda)\bar A_R$; then, letting $\tilde\theta = \frac{\bar\theta - \underline\theta}{\bar\theta + \underline\theta}$, the condition can be rewritten as

$$\begin{cases} T_R = \frac{\tilde\theta}{3}\frac{\lambda}{1-\lambda}\frac{\bar T_R}{\bar A_R} \\ A_R = \frac{\tilde\theta}{2}\left[1 - \frac{1}{3}\left(\frac{\lambda}{1-\lambda}\right)^2\frac{\bar T_R^2}{\bar A_R^2}\right] \end{cases}$$

This determines two political equilibria

$$(\tau,\alpha): \lambda\tau \leq (1-\lambda)\alpha \wedge \begin{cases} \tau = \frac{\tilde\theta}{3}\frac{\lambda}{1-\lambda}\frac{\tau}{\alpha} \\ \alpha = \frac{\tilde\theta}{2}\left[1 - \frac{1}{3}\left(\frac{\lambda}{1-\lambda}\right)^2\frac{\tau^2}{\alpha^2}\right] \end{cases}$$

namely

$$(\tau,\alpha) = \left(0, \frac{\tilde\theta}{2}\right) \quad \forall\lambda \in (0,1)$$

which is identical to what would obtain in the pure one-dimensional case; and

$$(\tau,\alpha) = \left(\frac{\tilde\theta}{3}\sqrt{3 - 2\frac{\lambda}{1-\lambda}}, \frac{\tilde\theta}{3}\frac{\lambda}{1-\lambda}\right) \quad \text{iff } \frac{1}{2} \leq \lambda < \frac{3}{5}$$

Symmetrically, we also have

$$(\tau,\alpha) = \left(\frac{\tilde\theta}{2}, 0\right) \quad \forall\lambda \in (0,1)$$

and

$$(\tau,\alpha) = \left(\frac{\tilde\theta}{3}\frac{1-\lambda}{\lambda}, \frac{\tilde\theta}{3}\sqrt{3 - 2\frac{1-\lambda}{\lambda}}\right) \quad \text{iff } \frac{2}{5} < \lambda \leq \frac{1}{2}$$

It is graphically immediate to distinguish two kind of equilibria: one has differentiation only on an issue whose salience is greater than $\frac{2}{5}$; the other is the continuous locus shown in the figure below (with τ the thinner, α the thicker line and $\tilde\theta = 0.2$)
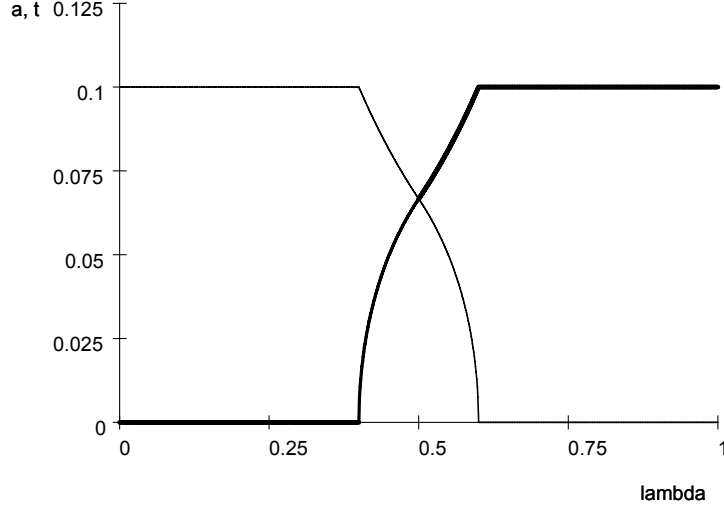
Figure A.1: Unstable equilibria

In fact, it turns out that only the first kind can be considered stable. Consider a sequential process whereby parties choose their optimal platform $(T_P, A_P)_{t+1}$ given that voters hold beliefs $(T_R, A_R)_{t-1}$ about the platform of party R and $(T_L, A_L)_{t-1} = -(T_R, A_R)_{t-1}$ about that of party L. By the first-order condition above, letting $x = \frac{T_R}{A_R}$ and $\kappa = \frac{\lambda}{1-\lambda}$ for ease of notation, we have the recursive equation

$$x_{t+1} = \frac{2\kappa x_t}{3 - \kappa^2 x_t^2} \forall x_t \leq \frac{1}{\kappa}$$

whose fixed points 0 and $\frac{\sqrt{3-2\kappa}}{\kappa}$ coincide with our equilibria. The slope of this function is

$$\frac{dx_{t+1}}{dx_t} = 2\kappa \frac{3 + \kappa^2 x_t^2}{\left(3 - \kappa^2 x_t^2\right)^2}$$

so that

$$\frac{dx_{t+1}}{dx_t}(0) = \frac{2}{3}\kappa < 1 \Leftrightarrow \lambda < \frac{3}{5}$$

while

$$\frac{dx_{t+1}}{dx_t}\left(\frac{\sqrt{3-2\kappa}}{\kappa}\right) = \frac{3-\kappa}{\kappa} > 1 \Leftrightarrow \lambda < \frac{3}{5}$$

This proves that the equilibrium with differentiation on issue A alone is stable only if $\lambda < \frac{3}{5}$, while the one with differentiation on both issues, which only exists when $\lambda < \frac{3}{5}$, is always unstable. E.g. for $\lambda = \frac{1}{2}$ both equilibria exist, but only the former is stable, as shown in the figure below (where the thinner line is the 45° line and the thicker curve the recursive
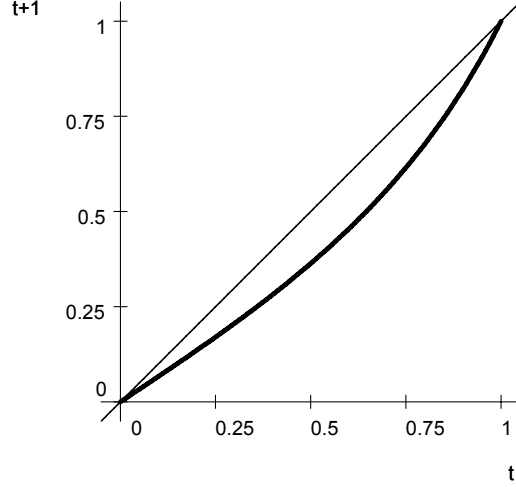
72

relationship).

Figure A.2: Stability of the equilibrium

Symmetrically, in the case where $x_t \geq \frac{1}{\kappa}$ it can be proven that the equilibrium with differentiation on issue T alone is stable only if $\lambda > \frac{2}{5}$, while the one with differentiation on both issues, which only exists when $\lambda > \frac{2}{5}$, is always unstable.

A.7   Proof of Proposition 11

When preferences are distributed with density $h(t,a) = \frac{1+\pi+\rho}{4} - \frac{\rho}{2}|t| - \frac{\pi}{2}|a|$ with $\pi, \rho \in \left[-\frac{1}{2}, \frac{1}{2}\right]$ the first-order condition becomes

$$\begin{cases} \int_{-1}^{1} \int_{-1}^{1} (t - T_R) \left(\frac{1+\pi+\rho}{4} - \frac{\rho}{2}|t| - \frac{\pi}{2}|a|\right) \begin{bmatrix} \bar{\theta}\chi\left(\lambda\bar{T}_R t + (1-\lambda)\bar{A}_R a > 0\right) + \\ +\underline{\theta}\chi\left(\lambda\bar{T}_R t + (1-\lambda)\bar{A}_R a < 0\right) \end{bmatrix} da\, dt = 0 \\ \int_{-1}^{1} \int_{-1}^{1} (a - A_R) \left(\frac{1+\pi+\rho}{4} - \frac{\rho}{2}|t| - \frac{\pi}{2}|a|\right) \begin{bmatrix} \bar{\theta}\chi\left(\lambda\bar{T}_R t + (1-\lambda)\bar{A}_R a > 0\right) + \\ +\underline{\theta}\chi\left(\lambda\bar{T}_R t + (1-\lambda)\bar{A}_R a < 0\right) \end{bmatrix} da\, dt = 0 \end{cases}$$

Suppose that $\lambda\bar{T}_R \leq (1-\lambda)\bar{A}_R$; then, letting again $\tilde{\theta} \equiv \frac{\bar{\theta}-\underline{\theta}}{\bar{\theta}+\underline{\theta}}$ and $\kappa \equiv \frac{\lambda}{1-\lambda}$, the condition can be rewritten as

$$\begin{cases} T_R = \frac{\tilde{\theta}}{12}\kappa\frac{\bar{T}_R}{\bar{A}_R}\left[4 + 4\pi - 2\rho - 3\pi\kappa\frac{\bar{T}_R}{\bar{A}_R}\right] \\ A_R = \frac{\tilde{\theta}}{6}\left[(3-\pi) - \frac{1}{2}(2 + 2\pi - \rho)\kappa^2\frac{\bar{T}_R^2}{\bar{A}_R^2} + \pi\kappa^3\frac{\bar{T}_R^3}{\bar{A}_R^3}\right] \end{cases}$$

This determines the political equilibria

$$(\tau, \alpha) : \lambda\tau \leq (1-\lambda)\alpha \wedge \begin{cases} \tau = \frac{\tilde{\theta}}{12}\kappa\frac{\tau}{\alpha}\left[4 + 4\pi - 2\rho - 3\pi\kappa\frac{\tau}{\alpha}\right] \\ \alpha = \frac{\tilde{\theta}}{6}\left[(3-\pi) - \frac{1}{2}(2 + 2\pi - \rho)\kappa^2\frac{\tau^2}{\alpha^2} + \pi\kappa^3\frac{\tau^3}{\alpha^3}\right] \end{cases}$$

73

clearly including the one found above for the one-dimensional case

$$(\tau, \alpha) = \left(0, \frac{3 - \pi}{6} \tilde{\theta}\right) \quad \forall \lambda \in (0, 1)$$

To proceed, consider the recursive equation for $x = \frac{T_R}{A_R}$ in the sequential process

$$x_{t+1} = \frac{\kappa x_t \left[2 \left(2 + 2\pi - \rho\right) - 3\pi\kappa x_t\right]}{2 \left[\left(3 - \pi\right) - \frac{1}{2} \left(2 + 2\pi - \rho\right) \kappa^2 x_t^2 + \pi\kappa^3 x_t^3\right]} \forall 0 \leq x_t \leq \frac{1}{\kappa}$$

so that

$$\frac{dx_{t+1}}{dx_t}(0) < 1 \Leftrightarrow \kappa < \frac{3 - \pi}{2 + 2\pi - \rho}$$

which proves that the one-dimensional equilibrium with differentiation on issue A is stable if and only if $\frac{\lambda}{1-\lambda} < \frac{3-\pi}{2+2\pi-\rho}$. This condition is more easily satisfied by lower values of λ, indicating greater salience of issue A; higher values of ρ, indicating greater concentration of preferences on issue T; and lower values of π, indicating lower concentration of preferences on issue A.

It remains to be shown that this is the unique stable equilibrium. Consider the equilibrium condition

$$x_{t+1} = x_t \Leftrightarrow x = 0 \lor \pi\kappa^3 x^3 - \left(1 + \pi - \frac{\rho}{2}\right) \kappa^2 x^2 + \frac{3}{2}\pi\kappa^2 x + 3 - \pi - \left(2 + 2\pi - \rho\right)\kappa = 0$$

and focus on the last equality.

Define the auxiliary function

$$g(x; \kappa, \pi, \rho) = \pi\kappa^3 x^3 - \left(1 + \pi - \frac{\rho}{2}\right) \kappa^2 x^2 + \frac{3}{2}\pi\kappa^2 x + 3 - \pi - \left(2 + 2\pi - \rho\right)\kappa$$

such that

$$g_x(x; \kappa, \pi, \rho) = 3\pi\kappa^3 x^2 - \left(2 + 2\pi - \rho\right)\kappa^2 x + \frac{3}{2}\pi\kappa^2$$

and observe that

$$g(0; \kappa, \pi, \rho) = 3 - \pi - \left(2 + 2\pi - \rho\right)\kappa > 0 \Leftrightarrow \kappa < \frac{3 - \pi}{2 + 2\pi - \rho}$$

$$g\left(\frac{1}{\kappa}; \kappa, \pi, \rho\right) = 2 - \pi + \frac{\rho}{2} - \left(2 + \frac{\pi}{2} - \rho\right)\kappa \leq 0 \Leftrightarrow \kappa \geq \frac{4 - 2\pi + \rho}{4 + \pi - 2\rho}$$

Therefore, so long as

$$g_x(x; \kappa, \pi, \rho) < 0 \forall x \in \left(0, \frac{1}{\kappa}\right]$$

74

for $0 < \kappa < \frac{4-2\pi+\rho}{4+\pi-2\rho}$ the only equilibrium such that $x \leq \frac{1}{\kappa}$ is the one-dimensional equilibrium $x = 0$, which is stable; for $\frac{4-2\pi+\rho}{4+\pi-2\rho} \leq \kappa < \frac{3-\pi}{2+2\pi-\rho}$ there are two equilibria with $x \leq \frac{1}{\kappa}$: the one-dimensional $x = 0$, which is stable, and a unique, unstable equilibrium with differentiation on both issues; if $\kappa \geq \frac{3-\pi}{2+2\pi-\rho}$ the only equilibrium such that $x \leq \frac{1}{\kappa}$ is the one-dimensional equilibrium $x = 0$, but it is unstable.

The condition required is then

$$3\pi\kappa^3 x^2 - (2 + 2\pi - \rho)\,\kappa^2 x + \frac{3}{2}\pi\kappa^2 < 0 \forall x \in \left(0, \frac{1}{\kappa}\right]$$

which is satisfied if $\pi \leq 0$, or if $\pi > 0$ and the quadratic has two real roots greater than $\frac{1}{\kappa}$, a sufficient condition for which is that

$$\kappa \leq \frac{1}{2}\left(\frac{2 + 2\pi - \rho}{3\pi}\right)^2 \wedge 4\pi + \rho \leq 2$$

The case of differentiation on issue T alone, i.e. with $\lambda \bar{T}_R \geq (1 - \lambda)\,\bar{A}_R$, follows by symmetry as usual.