Specialized Learning and Political Polarization

Sevgi Yuksel

UC Santa Barbara

August 2018

Abstract

When citizens are differentiated by how much they care about different issues informing policy, specialization allows citizens to focus their learning on the issues that are most important for them. However, as different citizens focus on different issues the electorate becomes less responsive to party platforms. In particular, equilibrium policies polarize more in fractionalized societies where there is greater disagreement about which issues matter the most. When the learning technology allows for more specialization, it effectively transforms the society into a more fractionalized one without changing the underlying preferences of the voters, thereby increasing polarization.

JEL Classification Numbers: D72, D80, Z13

Keywords: Polarization, Specialization, Voting behavior, Elections, Learning

I would like to thank Anmol Bhandari, Micael Castanheira, Catherine Hafer, Dimitri Landa, Jacopo Perego, Howard Rosenthal, Tobias Salz, Jesse Shapiro, Ken Shotts, Séverine Toussaert, and participants at seminars at UC Santa Barbara, USC, UC Davis, UC Santa Cruz, NYU, Stanford GSB, WUSTL, HKS, McGill, EUI, TSU, HEC Paris for helpful comments and suggestions. I would like to thank Guillaume Fréchette and Alessandro Lizzeri for their advice and support. I'm deeply indebted to Debraj Ray for his constant guidance at each stage of this project. All remaining errors are mine. Please address all correspondence to sevgi.yuksel@ucsb.edu.

1 Introduction

Political polarization in the US has traced a U-shaped pattern in the last century.[1] There were high levels of polarization in the early 20th century, which subsided after 1930. However, polarization has been increasing in the last thirty years, and now appears to be at an all-time high. At the same time, there is limited evidence to suggest that the distribution of voter preferences on a liberal-conservative scale has changed significantly in the last thirty years.[2]

The period in which polarization has increased coincides with a period in which there were also significant developments in learning technologies. The internet is becoming a primary information source for political news.[3] How the shift from traditional news sources to digital information affects citizens' engagement in the political process is not fully understood. However, there are many indicators suggesting that online news consumers gather information in fundamentally different ways. They are skeptical of the integrity of news organizations, rely more on local information and social media. Importantly, the internet allows citizens to easily target the type of information that is most relevant to them.

In some ways, things are not so different from the local, specialized pamphlets or newspapers of a century ago. As documented by Gentzkow et al. (2014), the end of the 19th and the beginning of the 20th century is also marked by a rapid growth in the number of local newspapers. However, circulation of most of these newspapers declined sharply within 50-100 miles of the publishing center. Thus, the content covered was highly specialized to the local readership. These local newspapers died out in the midcentury as national newspapers, radio and television took their place. As the news market consolidated, specialization became more difficult.

This paper presents a simple model to illustrate how optimal specialization in the type of information gathered by individual citizens, allowed by new learning technologies, can increase political polarization in the absence of any changes in the preferences of the citizens. Contrary to the previous literature on the topic, the model abstracts away from possible ideological biases in the news media or behavioral biases that might distort the decisions of citizens. The main results of the paper build on the following insight. When citizens are differentiated by how much

[1] This is measured as the distance (on a liberal-conservative scale) between the median policy position of a Democrat and that of a Republican. (McCarthy et al. (2006))

[2] This is a highly debated issue as changes in the distribution of voter preferences are harder to track. Fiorina & Abrams (2008) and Abramowitz and Saunders (2008) provide arguments on either side.

[3] Fifty percent of the Americans in 2013 stated internet as a main source for news compared to thirteen percent in 2001. (PEW Research Center)

they care about different aspects of ideological policies, specialization in information allows for differentiation in learning strategies adopted by the electorate. Specialization in this model takes place not on the broad ideological dimension of left and right as it is often assumed, but across different subdimensions that can inform policy. Citizens focus their learning on the issues that are most relevant for them in determining their position. However, as different citizens focus on different issues, the electorate becomes less responsiveness to party platforms, enabling ideologically motivated parties to propose more extreme policies.

A key starting point for the model is that citizens care differently about different aspects of policy. Some might care more about healthcare, while others care more about the environment or taxes. Moreover, any issue can be broken down into smaller subissues: issues relating to gender equality can include women's reproductive rights and discrimination in the workplace. When citizens have diverse views on how important different issues are, and there is uncertainty on ideal policy response to these issues, party platforms in equilibrium depend on the learning strategies adopted by the citizens. The link between the learning strategies and equilibrium policies arises as follows. Parties are ideologically motivated and receive rents from office. This creates a basic tension. The right party would like to implement right leaning policies conditional on being elected, but has an incentive to propose policies close to the position of the representative citizen to increase the probability of election. The way this tension is resolved depends on how responsive citizens are to the choice of the platform. As demonstrated by the model, that responsiveness depends on the characteristics of the learning technology.

The paper identifies several channels through which characteristics of the learning technology interact with the distribution of preferences in the electorate to impact responsiveness to policies. First, when the learning technology allows for specialization, in *more fractionalized* societies, where there is greater disagreement among citizens about which issues matter the most, equilibrium policies polarize more. For example, a society consisting of citizens who care only about economic or social issues, even if there are equal number of each type, would be defined as more fractionalized relative to another one, where all citizens care equally about both issues. Observe that more fractionalized societies, while allowing for higher polarization, look, in aggregate, the same as less fractionalized societies, putting equal weight on both issues.

A second issue of interest is greater specialization in learning, or the "fineness" of the learning technology. Increasing the *depth of learning*, which captures how fine the learning technology is in terms of the subissues a voter is able to specialize in, also increases polarization. For an electorate whose preferences remain constant, such refinements in the learning technology create further differentiation in terms of the actual learning strategies adopted by the electorate. The insight this paper builds on is that when citizens are uncertain about their position on a policy, the degree to which the underlying heterogeneity in the preferences of the citizens impacts elections depend on the level of specialization allowed by the learning technology. As different citizens take positions determined by different issues, responsiveness to party platforms declines, consequently increasing polarization. The main result in this respect also draws a connection between specialization allowed by learning technologies and fractionalization of a society. As the depth of the learning technology increases, a society operates in a more fractionalized manner. Although the distribution of citizens remain the same, changes in the learning technology increase the level of fractionalization manifested in how they respond to policies.

Finally, to highlight how a more informed electorate could restrain political polarization, an alternative change to learning technologies that allow for easier *access* to information is considered. Keeping the level of specialization constant, citizens are assumed to simply receive more precise signals. On the individual level, this has the same effect as greater ability to specialize: the probability that each citizen supports the party which is closer to her ideal position increases. However, in contrast to specialization, this has the impact of correlating citizens decisions, increasing their responsiveness to policies which decreases polarization.

This paper is motivated by the larger public debate on how the new information landscape governed by social media and digital news reshapes the political process. These technologies dramatically reduce the cost of providing and acquiring information, but they also endow citizens with remarkable tools to receive specialized information. The goal of this paper is to demonstrate that while both channels enable citizens to learn more effectively, they can have radically different consequences on political outcomes.

This paper contributes to a growing literature studying political polarization. Linking political polarization among ideologically motivated parties to uncertainty about the ideal points of the electorate goes back a few decades.[4] Recently, McCarty et al. (2018) provides empirical evidence that is suggestive of this causal link.[5] This paper builds on this insight by endogenizing the

[4]Duggan (2005) classifies this as the *stochastic preference model with policy motivation.* Hansson & Stuart (1984), Wittman (1983, 1990), Calvert (1985) and Roemer (1994) are some early papers that make this connection.

[5]Using roll-call voting behavior among state legislators, they show that the ideological distance between Democrats and Republicans from a district is correlated with the ideological heterogeneity of the electorate from that district. They interpret this to be supportive of a model where intra-district ideological polarization causes candidates to be uncertain about the ideological location of the median voter, thereby reducing their incentives for platform moderation.

uncertainty using a multi-dimensional framework with the aim of studying how changes in the learning stage can affect polarization.

Papers studying political polarization have predominantly focused on the role of media bias where bias can broadly be considered as partial information disclosure or slanted reporting.[6] Galeotti and Mattozzi (2011) and Murphy and Shleifer (2004) study the interaction of network structure and polarization. Levy & Razin (2014) provides a model in which the distribution of voters' ideal points polarize due to *correlation neglect* in learning, defined as the failure to take into account correlation in information sources. Other explanations include specialized candidates (Krasa & Polborn (2010, 2012)), interaction between valence and ideology (Polborn & Snyder Jr (2017), Eyster & Kittsteiner (2007), Groseclose (2001) and Ashworth & de Mesquita (2009)), informational asymmetries between voters and parties (Martinelli (2001)), and political entry deterrence when there is preference heterogeneity (Callander (2005)). This paper also contributes to a literature studying how preference heterogeneity affects public good provision.[7] Relatedly, Matějka & Tabellini (2017) studies policy choice when voters are rationally inattentive. Complementing results in this paper, they find that divisive issues attract the most attention by voters, and that this can create inefficiency in public good provision.

This paper is most closely related Perego & Yuksel (2018) which studies the competitive provision of political information. The critical insight put forward in that paper is that competition forces information providers to specialize by becoming relatively less informative on issues that are important from a social point of view, amplifying social disagreement. Focus on specialization of information in a framework with multi-dimensional preferences link the the two papers. However, while Perego & Yuksel (2018) studies competition among information providers, this paper takes changes in the learning environment as given, focusing on the policy choices of the parties.

The paper is organized as follows. Section 2 presents the model and discusses its key features. Section 3 defines and solves for the equilibrium. Section 4 presents the main results, and Section 5 concludes.

[6] See Besley & Prat (2006), Baron (2006*a*), Duggan & Martinelli (2010), Gul & Pesendorfer (2012) for prominent examples. Bernhardt et al. (2008), Mullainathan & Shleifer (2005), Gentzkow & Shapiro (2006) study endogenously supplied information and it's impact on electoral outcomes when there is demand for biased news. Gentzkow et al. (2015) provides a recent review of the theoretical literature.

[7] See Banerjee & Pande (2007), Alesina et al. (1999), Lizzeri & Persico (2001) and Fernandez & Levy (2008).

2 Model

Two parties indexed by $i \in \{L, R\}$ compete in an election. Each party commits to a policy y_i that will be implemented upon election. Let $y := (y_R, y_L)$ denote the policy choices of both parties. There are $\mathcal{K} := \{1, ..., K\}$ issues that could influence a citizen's position on this policy. There is a continuum of citizens with diverse views on how different issues should be considered to form a position. Each citizen is characterized by a vector $w := (w^1, ..., w^K) \in \Delta^K$ which denotes how much weight she puts on each issue in evaluating her position. $\theta := (\theta^1, ..., \theta^K)$ represents relevant facts that are considered in relation to each of these issues, such that $w \cdot \theta = \sum_{k \in \mathcal{K}} w^k \theta^k$ describes a citizen's ideal *position* on the policy.

Distribution of citizen types:

Citizens are distributed with cdf $F(w)$ on Δ^K. The assumed symmetry guarantees that the electorate puts equal weight on each issue in aggregate, but marginal distributions can vary with k. The policy position associated with the representative citizen is denoted as $y_m := w_m \cdot \theta$.[8]

Assumption 1. $F(w)$ is smooth and symmetric in aggregate, i.e. $w_m := \int w \, dF(w) = (\frac{1}{K}, ..., \frac{1}{K})$.

Preferences of the citizens:

Individual citizens prefer policies that are closer to their position. That is, conditional on θ, the payoff to citizen of type w if policy y_i is implemented can be written as:

$$-(w \cdot \theta - y_i)^2. \tag{1}$$

Individuals take a binary action $a \in \{-1, 1\}$, which could be interpreted as voting, to show support for either party R ($a = 1$) or L ($a = -1$). Citizens are assumed to receive utility directly from supporting the party whose proposed policy gives them a higher payoff.[9]

Uncertainty and learning:

A citizen's decision to support either party depends on y and her ideal policy $w \cdot \theta$, but while y is publicly observed, citizens have imperfect information on θ. The state of the world on each

[8] Equivalence of y_m with the median position in the society requires stronger symmetry assumptions.

[9] Since no individual has an impact on the policy outcome, a direct utility associated with expressing one's preferences (perhaps rising from a sense of civic responsibility) is the most straightforward and possibly realistic assumption in this context. Others, such as Eraslan & Ozerturk (2017), Stromberg (2004), Gentzkow & Shapiro (2006), Baron (2006*b*) have argued that political news can have instrumental value for citizens beyond their impact on voting informing private actions such as financial, educational, or labor supply decisions.

dimension, θ^k, is assumed to be independently drawn from a normal distribution with mean 0 and precision ρ^k (with ρ^k the same on all dimensions).

Each citizen has access to a learning technology $\mathcal{L} := (\eta, \mathcal{P})$. The learning technology is characterized by a *precision* η and a *partition* $\mathcal{P}$. Formally, $\mathcal{P}$ is a partition of $\mathcal{K}$, which implies the following conditions on $\mathcal{P}$: (1) $\bigcup\limits_{A \in \mathcal{P}} A = \mathcal{K}$; (2) if $A, B \in \mathcal{P}$ such that $A \neq B$ then $A \bigcap B = \emptyset$.

Each citizen receives a costless signal $s = w_l \cdot \theta \; + \; \epsilon$ from the learning technology where w_l is the learning strategy chosen by the citizen, with the subscript l included to differentiate it from the citizen's type w. The *precision* of the learning technology determines the distribution of the noise term, $\epsilon \sim \mathcal{N}(0, \frac{1}{\eta})$, which is independent across citizens. The *partition* of the learning technology imposes constraints on the learning strategies available to the citizens. Formally, $w_l \in \Delta_{\mathcal{P}}$, where

$$\Delta_{\mathcal{P}} := \{w \in \Delta^K \mid w^k = w^l \text{ for all } k, l \in A \text{ for some } A \in \mathcal{P}\}.$$

Note that each cell of the partition corresponds to a subset of issues in $\mathcal{K}$. Each citizen chooses a learning strategy w_l (which issues to focus on) to maximize the probability of supporting the party whose proposed policy is closest to her position.

Preferences of the parties:

Before committing to policies, parties gather information and form a prior on the position of the representative citizen: $\mu \sim \mathcal{N}(y_m, \frac{1}{\rho^0})$.[10] ρ^0 captures how much uncertainty parties face in terms of the position of the representative citizen when they commit to policies. Both the prior on θ and the information structure available to the citizens to learn about θ are common knowledge among the parties and the citizens.

The probability a party gets elected is assumed to be directly linked to how much support it has from the public. Let $\Gamma(y, \theta) := \int\limits_{w} z_w(y|\theta) dF(w)$ denote the support party R receives where $z_w(y|\theta)$ is a random variable that denotes the action taken by a citizen of type w given policy choices y and state of the world θ.

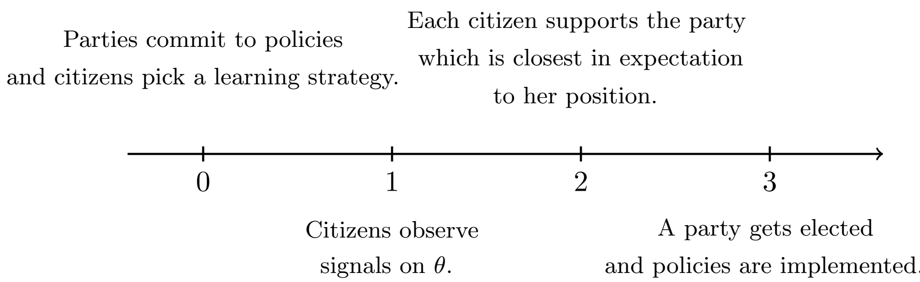
Parties are ideologically motivated, but also receive rents from being in office. The payoff to party i if policy y is implemented can be written as:

$$-L(y_i^\star - y) + \mathbf{1}_i \Lambda \tag{2}$$

[10]For simplicity, the prior is assumed to be the same for both parties.

where $y_i^\star := \mu + b_i$, the preferred policy for party i, is assumed to be biased either towards the right or left relative to the expected position of the representative citizen. This implies that parties put equal weight on each issue.[11] L is a loss function and is assumed to be symmetric with $L'(0) = 0$, $L'(x) > 0$ and $L''(x) > 0$ for $x > 0$. Fixing the loss function, the underlying ideological polarization in party preferences can be captured by $b_R = -b_L$. $\Lambda > 0$ represents rents from office. $\mathbf{1}_i$ is the indicator function which equals 1 when party i wins the election.

Timing:



Discussion of the Model:

This section revisits the key modeling assumptions and their implications. As highlighted above the main focus of the paper is to study how the learning environment impacts the distribution of citizens' choices and consequently the policies proposed by the parties.

Multi-dimensional uncertainty. It is reasonable to assume that different individuals evaluate the many issues that might inform their position on a policy differently. This assumption also provides a natural framework to study the implications of specialized learning.

Learning environment. The learning technology, with partition $\mathcal{P}$, effectively limits the degree to which citizens can differentially get in informed on separate dimensions of the state, and with precision η, puts constraints on how much citizens are able to learn from the signal overall. In a model with multi-dimensional uncertainty comparative statics on the partition $\mathcal{P}$ can be used to study how the ability to specialize affects polarization. Focusing on the precision η, on the other hand, provides insight on how greater access to information (holding constant the level of specialization) affects polarization. Note that all citizens are assumed to have access to the same learning technology and their choice of w_l can be interpreted as solving an attention allocation problem subject to the constraints of the learning environment.[12]

[11] This is a natural assumption to make if we take the parties to be representative of their support base.

[12] One could, alternatively, consider a model with costly information acquisition. The model presented here, by abstracting from this decision for the citizens, contrasts how different aspects of the learning technology affects the

Policy choice. Parties are assumed to commit to their policy choices, and are not able to shift policy in response to the state of the world. This contributes to the uncertainty with respect to the probability of winning for each party. While there are certainly many instances where policies shift in response to aggregate shocks, a fully flexible shift in policy positions to respond to the state of the world is invariably difficult.[13] Furthermore, even when parties are perfectly informed on the position of the representative citizen, remaining uncertainty about the distribution of positions is sufficient to generate polarization. Also, the preferred policy for each party is assumed to depend on the *expected* position of the representative citizen, μ, not y_m. Since parties are allowed to be informed about y_m with arbitrarily precision (with $\rho^0 \gg \rho^k$), results would generalize to the case where preferred policies depend directly on y_m. Note that this can create additional incentives for polarization. Hence focusing on μ provides the most transparent demonstration of the mechanism driving the main results.

3 Equilibrium with Learning

Policy choices y and learning strategy w_l for each type of citizen constitutes an equilibrium if each party maximizes its expected payoff (given the learning strategy of the citizens) holding constant the other party's policy choice; and citizens choose their learning strategy to maximize the probability they support the party whose proposed policy is closest to their position. An equilibrium is referred to as *symmetric* if parties polarize equally to the left and right of the ex-ante position of the representative citizen, i.e. $y_R = \mu + \beta$ and $y_R = \mu - \beta$ for some $\beta > 0$.

The following definition will be useful in describing the learning strategies of the citizens.

Definition 1. $w_{\mathcal{P}}$ is a *projection* of w on partition $\mathcal{P}$, if $w_{\mathcal{P}} \in \Delta_{\mathcal{P}}$ and $\sum_{k \in A} w_{\mathcal{P}}^k = \sum_{k \in A} w^k$ for every $A \in \mathcal{P}$.

Lemma 1. *In a symmetric equilibrium, the optimal learning strategy is to choose* $w_l = w_{\mathcal{P}}$.

Lemma 1 states that a citizen, following his optimal learning strategy, concentrates attention on the issues that he cares most about.[14] The degree to which such specialization can take place is constrained by the partition associated with the learning technology. More importantly, the

political process in the most transparent way.

[13] For instance, a party that has called for deregulation of the financial market for years, which has chosen a candidate who symbolizes this position, cannot easily shift gears in response to economic shocks that increase support in the electorate for more regulation.

[14] All formal proofs (including those for the Remarks) are provided in the Appendix.

Lemma formalizes the natural intuition that in a society where there is heterogeneity with respect to which issues are more relevant for determining the optimal policy, there will consequently be heterogeneity with respect to the learning strategies adopted.

Remark 1. Given policy choices y and the learning strategy of the citizens in a symmetric equilibrium, the probability party R is elected can be written as:

$$s(y) := \mathbb{E}(\Gamma(y,\theta)) = \int \Phi\left(\frac{-\delta}{\sigma_{\mathcal{L}}(w)}\right) dF(w), \tag{3}$$

where $\delta = \frac{y_R+y_L}{2} - \mu = 0$ and $\sigma^2_{\mathcal{L}}(w) = \sum_k \left(w^k_{\mathcal{P}} - w^k_m\right)^2 \frac{1}{\rho^k} + \frac{1}{\rho^0} + \frac{1}{\eta}$.

Remark 1 demonstrates how learning strategies can potentially impact the election. Aggregate uncertainty with respect to winning probability is an equilibrium object. $\sigma_{\mathcal{L}}(w)$ is different for each type w and a function of the learning technology.

Definition 2. Given policy choices y, let $\pi(y) := (\pi_R(y), \pi_L(y))$ stand for the marginal cost of deviations on policies in terms of probability of winning. This is referred as *responsiveness* to policies (evaluated at y).

In a symmetric equilibrium, there are two forces that dictate the positions of the parties. The optimality condition written for party R, for a policy strictly to the right of the ex-ante position of the representative citizen, reveals these two opposing forces:

$$-\pi_R(y)\left[L(y^\star_R - y_R) - L(y^\star_R - y_L) + \Lambda\right] + s(y)L'(y^\star_R - y_R) = 0 \tag{4}$$

Conditional on winning the election parties prefer policies that are closer to their ideological position (second component of Equation 4). This can be considered as a *policy distortion effect.* At the same time, deviating from the equilibrium policy by moving closer to the party's ideal position decreases the probability with which the party wins the election (first component of Equation 4). This can be considered as a *fear of loss effect.* The force of this effect increases with the level of office motivation and polarization in policy platforms but decreases with responsiveness to the policy choices. There is a unique symmetric equilibrium where these opposing forces are in balance.

Proposition 1. *There is a unique symmetric equilibrium.*

A key insight is that polarization is closely linked to responsiveness to policies. The upcoming sections investigate how changes in the distribution of citizens preferences as well as changes in the learning technology can affect equilibrium policies through this channel.

Fractionalized societies

The following definition introduces a partial ranking on how specialized citizens are.

Definition 3. Let i and j be two different citizens with associated weight vectors w_i and w_j. i is a *more specialized citizen* than j relative to partition $\mathcal{P}$ if for $\tilde{w}_{i_\mathcal{P}}$ and $\tilde{w}_{j_\mathcal{P}}$, referring to permutations of projections $w_{i_\mathcal{P}}$ and $w_{j_\mathcal{P}}$ such that weights are in increasing order:

$$\tilde{w}_{i_\mathcal{P}} \geq_{FOSD} \tilde{w}_{j_\mathcal{P}} \tag{5}$$

According to the definition, voter i is a *more specialized citizen* than j (relative to partition $\mathcal{P}$) if the projection of i's weight vector on $\mathcal{P}$ first order stochastically dominates j's when the weights put on different issues are reshuffled such that weights are in increasing order. The definition captures the idea that weights are more concentrated for voter j than i. As an example, consider a policy (changes to the tax code or healthcare) which will impact society in many different ways. Assume the policy can be evaluated broadly by separately considering it's economic and social consequences. Compare a citizen who puts equal weight on both of these aspects to a citizen who determines his position by solely considering the economic aspects. The corresponding weight vectors can be represented as $(\frac{1}{2}, \frac{1}{2})$ and $(0, 1)$ where the second component corresponds to weight put on economic aspects. Since $(0, 1) \geq_{FOSD} (\frac{1}{2}, \frac{1}{2})$, the second citizen is defined to be more specialized. Actually, for such a partition, where the underlying issues can only be separated into social and economics ones, $(\frac{1}{2}, \frac{1}{2})$ represents the *least* specialized citizen, and $(0, 1)$, $(1, 0)$ represent the *most* specialized citizens.

The following lemma demonstrates why comparing citizens in terms of how specialized they are is important.

Lemma 2. When all citizens choose their learning strategy optimally, more specialized citizens (relative to partition $\mathcal{P}$) are less responsive to policy choices.

The key issue is that more specialized citizens prefer more specialized learning strategies. Relative to others, they focus on a smaller subset of issues. This allows them to respond more directly to aspects of the state of the world that are most relevant for them, which are aspects that differentiate them from the representative citizen. This, in return, decreases their responsiveness to policies.

Definition 4 extends the previous ranking of specialization across citizens to societies in a natural way. A society is identified to be more fractionalized than another one if it is generated by a transformation of the latter one where each citizen is mapped to a more specialized voter.

Definition 4. Let S and $\tilde{S}$ be two societies where the type distribution is captured by F and $\tilde{F}$. $\tilde{S}$ is a *more fractionalized society* than S relative to partition $\mathcal{P}$ if there exists a mapping g from $\Delta^K \to \Delta^K$ such that for any w, $g(w)$ describes a more specialized citizen and $\tilde{F}$ corresponds to the transformed distribution using g from F.

Proposition 2. *If society $\tilde{S}$ is more fractionalized than society S relative to partition $\mathcal{P}$, there is higher polarization in proposed policies in $\tilde{S}$.*

Proposition 2 highlights how fractionalization can generate further polarization in a society. The result relies on Lemma 2 which states that more specialized citizens are less responsive to policy choices. As the population overall gets more fractionalized, the marginal cost (in terms of the decrease in probability of winning) of deviating to more extreme policies declines, while conditional on winning, the marginal benefit of implementing policies closer to the party's ideal point remains constant. This force pushes proposed policies in equilibrium away from the position of the representative citizen.

An important note with regards to Proposition 2 is that more fractionalized societies do not look any different in aggregate. Symmetry assumptions on distribution F impose the position of the representative citizen to remain constant. In terms of the weights put on different issues, in aggregate, the society puts equal weight on all issues regardless of how fractionalized it is. More fractionalized societies refer to ones where there is greater disagreement among the citizens about which issues matter the most for determining optimal policy. Going back to the example where the partition allows for comparison along two dimensions (consisting of economic and social issues), Proposition 2 implies that a society consisting of citizens who only care about either the economic or social issues (with equal share of both) would generate higher polarization relative to a society where everyone cares equally about both issues.

4 Changes in the learning technology

The characterization of the learning environment lends itself to two natural comparative static exercises, one with respect to the precision η and one with respect to the partition $\mathcal{P}$. An increase η can be interpreted as easier access to information. Considering improvements in learning technologies that allow citizens to get specialized information about issues that are most relevant for them requires focusing on changes in the partition $\mathcal{P}$.

Definition 5. Learning technology $\tilde{\mathcal{L}}$ *allows for more specialization* than $\mathcal{L}$ if the associated $\tilde{\mathcal{P}}$ is a finer partition of $\mathcal{K}$ than $\tilde{\mathcal{P}}$, i.e. for any $A \in \mathcal{P}$, $\exists \tilde{A} \in \tilde{\mathcal{P}}$ such that $\tilde{A} \subset A$.

Definition 5 provides an intuitive partial order on how much specialization is allowed by a learning technology. Holding constant the precision of the signals received by the citizens, the partition determines the kind of signals accessible to the citizens, and consequently affects the degree to which citizens are able to learn about the specific issues that are most relevant for them in determining their position on a policy. Going back to the example on healthcare, two citizens, whose position on the policy is determined exclusively by economic considerations, could disagree on which economic measures matter the most. A learning technology that allows for more specialization provides greater *depth of learning* as it effectively increases the number of subissues a citizen is able to learn separately about.

Proposition 3. *Changes in the learning technology that increase precision or allow for more specialization increase the probability that any citizen w (holding constant y) supports the party whose proposed policy is closer to her position.*

Proposition 3 affirms that improvements in the learning technology that increase access to information or allow for specialization are beneficial to citizens on an individual level. Both types of changes allow citizens to learn more effectively, and consequently increase the probability the the action they take conditional on the information available to them – to support party R or L – matches their true preferences.

However, the results captured in Proposition 3 on how both types of changes result in more effective learning do not translate into similar equilibrium effects. The main result of the paper presented below illustrates how changes in η and P affect the policy choices of the parties in radically different ways.

Theorem 1. *For any society S, an increase in η (holding constant $\mathcal{P}$) decreases polarization in proposed policies. In contrast, changes in $\mathcal{P}$ that allow for more specialization (holding constant η) increases polarization.*

The learning technology affects support for each party through two qualitatively different channels. Citizens face uncertainty about where the parties are located relative to the position of the representative citizen; but given heterogeneity in preferences, they also face uncertainty about how their position differs from that of the representative citizen. Polarization is linked to responsiveness to policies which is a function of both types of uncertainty. The simple model presented in this paper is intended to demonstrate in the most transparent way how access to more information vs. increased opportunities for specialization can operate to impact citizen's actions through these separate channels. Easier access to information (increase in η) creates a more informed electorate where the private signals of the citizens also correlate more. Actions of the citizens more closely reflect where the parties are located relative to the position of the representative citizen, essentially making the society more responsive to the policies. In this sense, comparative statics with respect to η capture our basic intuition on how a more informed electorate should be able to monitor the parties closer, pressuring them to choose policies that are closer to the position of the representative citizen.

In contrast, increased opportunities for specialization allow individuals to focus their learning on issues where their preferences differ from that of the representative citizen. While this improves learning from an individual perspective, increasing the probability that the citizen supports the party which is closer to her ideal position, it makes the citizen's decision on which party to support less predictable ex-ante. Citizens with diverse views on how an ideal policy shoud be determined condition their decision on which party to support on different types of information. Differentiation in learning strategies makes these decisions less correlated with each other, and hence less responsive to the party platforms. In a model where parties are trading off probability of election with how extreme their policies are, this translates into higher polarization. As a counterpart to the results on the precision, comparative statics with respect to the depth of learning, highlight that a more informed electorate on the individual level does not imply reduction in political polarization.

The intuition behind this result is closely linked to the results on fractionalization in the previous section. The following example demonstrates this link. Consider switching from a very coarse learning technology where there is only one issue to learn about to another one where the underlying issues space is partitioned into economic and social issues. With the former learning technology, all voters are identical in terms of their weight vectors, namely they all put full weight on the only issue available for learning as it includes all possible subissues. Moreover, the optimal learning strategy is trivially determined as the learning technology allows for only strategy. When the issue space is partitioned into two, there will be some citizens who put equal weight on both issues. The corresponding optimal learning strategy for these types will be to divide their focus their leanring equally between the two issues. Consider a society consisting of only such citizens as a benchmark. In this society polarization would not be affected by such a change in the learning technology. However, in general, there will be other types of citizens, who do not equally care about the two subissues. The presence of such citizens implies that the society will be more fractionalized in it's learning strategies relative to the benchmark society described above. In other words, increasing the depth of learning increases the level of fractionalization in a society. By Proposition 2, this generates higher polarization.

5 Conclusion

This paper is motivated by the larger public debate on how the new information landscape governed by social media and digital news reshapes the political process. These technologies dramatically reduce the cost of providing and acquiring information, but they also endow citizens with remarkable tools to receive specialized information. As discussed in the introduction, it can be argued that the market structure of new sources in the early 20th century also allowed for a great degree of specialization. The goal of this paper is to illustrate how specialization in the type of information gathered by citizens can increase political polarization in the absence of any changes in the preferences of the electorate. Thus, the link between specialized learning and polarization provides a potential explanation for the U-shaped pattern traced by political polarization in the last century. When citizens have diverse views on how optimal policy should be determined, specialization allows individuals to focus their learning on the issues that are most relevant for them. In equilibrium, this enables citizens to learn more about how their preferences differ from that of the representative citizen. In aggregate, as different citizens focus on different issues, the electorate becomes less responsive to policies. Main results of the paper show that, when information technologies allow for specialization, equilibrium policies polarize more in fractionalized societies where there is greater disagreement about which issues matter the most. Increasing the depth of learning, which captures the degree to which the learning technology allows for specialization, also increases polarization by transforming the society into a more fractionalized one.

References

Alesina, A., Baqir, R. & Easterly, W. (1999), 'Public goods and ethnic divisions', *The Quarterly Journal of Economics* **111**(4).

Ashworth, S. & de Mesquita, E. D. (2009), 'Elections with platform and valence competition', *Games and Economic Behavior* **67**(1), 191–216.

Banerjee, A. & Pande, R. (2007), 'Parochial politics: Ethnic preferences and politician corruption', *mimeo* .

Baron, D. (2006*a*), 'Persisten media bias', *Journal of Public Economics* **90**, 1–36.

Baron, D. (2006*b*), 'Persisten media bias', *Journal of Public Economics* **90**, 1–36.

Bernhardt, D., Krasa, S. & Polborn, M. (2008), 'Political polarization and the electoral effects of media bias', *Journal of Public Economics* .

Besley, T. & Prat, A. (2006), 'Handcuff's for the grabbing hand? media capture and government accountability', *American Economic Review* **96**(3).

Callander, S. (2005), 'Electoral competition in heterogeneous districts', *Journal of Political Economy* **113**(5), 1116–1145.

Calvert, R. (1985), 'Robustness of the multi-dimensional voting model: Candidate motivations, uncertainty and convergence', *American Journal of Political Science* **29**.

Duggan, J. (2005), A survey of equilibrium analysis in spatial models of elections.

Duggan, J. & Martinelli, C. (2010), 'A spatial theory of media slant and voter choice', *Working paper* .

Eraslan, H. & Ozerturk, S. (2017), 'Information gatekeeping and media bias'.

Eyster, E. & Kittsteiner, T. (2007), 'Party platforms in electoral competition with heterogeneous constituencies', *Theoretical Economics* **2**, 41–70.

Fernandez, R. & Levy, G. (2008), 'Diversity and redistribution', *Journal of Public Economics* .

Fiorina, M. & Abrams, S. J. (2008), 'Political polarization in the american public', *Annu. Rev. Polit. Sci.* **11**(563-588).

Gentzkow, M. & Shapiro, J. M. (2006), 'Media bias and reputation', *Journal of political Economy* **114**(2), 280–316.

Gentzkow, M., Shapiro, J. M. & Stone, D. F. (2015), Media bias in the marketplace: Theory, *in* 'Handbook of media economics', Vol. 1, Elsevier, pp. 623–645.

Gentzkow, M., Shapiro, J. & Sinkinson, M. (2014), 'Competition and ideological diversity: Historical evidence from us newspapers', *American Economic Review* **104**(10).

Groseclose, T. (2001), 'A model of candidate location when one candidate has a valence advantage', *American Journal of Political Science* **45**(4), 862–886.

Gul, F. & Pesendorfer, W. (2012), Media and policy.

Hansson, I. & Stuart, C. (1984), 'Voting competitions with interested politicians: Platforms do not converge to the preferences of the median voter', *Public Choice* **44**(3), 431–441.

Krasa, S. & Polborn, M. (2010), 'Competition between specialized candidates', *American Political Science Review* **104**(4), 745–765.

Krasa, S. & Polborn, M. K. (2012), 'Political competition between differentiated candidates', *Games and Economic Behavior* **76**(1), 249–271.

Levy, G. & Razin, R. (2014), Does polarization of opinions lead to polarization of platforms? the case of correlation neglect. Working paper.

Lizzeri, A. & Persico, N. (2001), 'The provision of public goods under alternative electoral incentives', *American Economic Review* **91**(1).

Martinelli, C. (2001), 'Elections with privately informed parties and voters', *Public Choice* **108**(1-2), 147–167.

Matějka, F. & Tabellini, G. (2017), 'Electoral competition with rationally inattentive voters'.

McCarthy, N., Poole, K. T. & Rosenthal, H. (2006), *Polarized America: The dance of ideology and unequal riches*, Mit Press.

McCarty, N., Rodden, J., Shor, B., Tausanovitch, C. & Warshaw, C. (2018), 'Geography, uncertainty, and polarization', *Political Science Research and Methods* pp. 1–20.

Mullainathan, S. & Shleifer, A. (2005), 'The market for news', *American Economic Review* **95**(4).

Perego, J. & Yuksel, S. (2018), 'Media competition and social disagreement'.

Polborn, M. K. & Snyder Jr, J. M. (2017), 'Party polarization in legislatures with office-motivated candidates', *The Quarterly Journal of Economics* **132**(3), 1509–1550.

Roemer, J. E. (1994), 'A theory of policy differentiation in single issue electoral politics', *Social Choice and Welfare* **11**(4), 355–380.

Stromberg, D. (2004), 'Radio's impact on public spending', *Quarterly Journal of Economics* **119**, 189–221.

Wittman, D. (1983), 'Candidate motivation: A synthesis of alternative theories', *The American Political Science Review* **142-157**.

Wittman, D. (1990), 'Spatial strategies when candidates have policy preferences', *Advances in the spatial theory of voting* (66-98).

Proofs

Lemma 1.

Proof. In a symmetric equilibrium, citizen's beliefs are such that $y_i = \mu + \beta_i$ with $\beta_L = -\beta_R = -\beta$. A citizen of type w strictly prefers party R under complete information if and only if $w \cdot \theta - \hat{y} > 0$ where $\hat{y} = \frac{y_R + y_L}{2} = \mu$. Using $\mu = y_m + \epsilon_p$ and $y_m = w_m \cdot \theta$, and defining $w_d = w - w_m$, $w \cdot \theta - \mu = w_d \cdot \theta - \epsilon_p$. A learning strategy w_l provides a signal $s = w_l \cdot \theta + \epsilon$. Note that $s - \hat{y} = (w_l - w_m) \cdot \theta - \epsilon_p + \epsilon$. We can write $(w_l - w_m) \cdot \theta = w_d \cdot \theta + w_{d_l} \cdot \theta$. Without loss of generality assume that $w_d \cdot \theta - \epsilon_p > 0$. Clearly, the citizen's decision on which party to support depends on the sign of $s - \hat{y}$. Hence, the citizen supports the correct party whenever $(w_d \cdot \theta - \epsilon_p) + (w_{d_l} \cdot \theta + \epsilon) > 0$. To maximize this chance, the citizen chooses w_l to minimize the variance of $w_{d_l} \cdot \theta + \epsilon$ which is equal to $\sum (w_{d_l}^k)^2 \frac{1}{\rho^k} + \frac{1}{\eta}$. Note that $w_{d_l}^k = w_l^k - w^k$. If w_l is the optimal learning strategy for a citizen, it must be that the citizen doesn't want to change w_l^k in any $A \in \mathcal{P}$. So it must be that the derivative of the variance term is such that for any $A \in \mathcal{P}$, $\frac{2}{\rho^k} \sum (w_l^k - w^k) = 0$ which gives us the desired result. □

Remark 1.

Proof. Given the optimal learning strategy of the citizens ($w_l = w_{\mathcal{P}}$) , we can calculate the expected probability with which a citizen of type w supports party R given y and θ. Assume party $y_L = \mu - \beta$ and $y_R = \mu + \beta + \Delta$. Call $\delta = \frac{y_R + y_L}{2} - \mu = \frac{\Delta}{2}$. Let $w_{d_{\mathcal{P}}} = w_{\mathcal{P}} - w_m$. The citizen votes for party R whenever $s - \hat{y} = w_{d_{\mathcal{P}}} \cdot \theta - \delta - \epsilon_p + \epsilon > 0$. Conditional on θ the probability of this happening is $\Phi\left(\frac{w_{d_{\mathcal{P}}} \cdot \theta - \delta}{\sigma}\right)$ where $\sigma^2 = \frac{1}{\rho^P} + \frac{1}{\eta}$. Taking the expectation over θ and integrating over types, $\mathbb{E}(\Gamma(y, \theta)) = \int \Phi\left(-\frac{\delta}{\sigma_{\mathcal{L}}(w)}\right) dF(w)$ where $\sigma_{\mathcal{L}}^2(w) = \sum_k (w_{d_{\mathcal{L}}}^k)^2 \frac{1}{\rho^k} + \frac{1}{\rho^P} + \frac{1}{\eta}$. □

Proposition 1.

Proof. Let $\pi_R(y) = \left|\frac{\partial \mathbb{E}(\Gamma(y,\theta))}{\partial y_R}\right|$. Defining $y_R = \mu + \beta + \Delta$, the maximization problem for the right party can be written as follows.

$$\max_{\Delta} \{s(y)(-L(b - \beta - \Delta) + \Lambda) + (1 - s(y))(-L(b + \beta))\}$$

The first order conditions for optimality give us: $-\pi_R(y)(L(b - \beta - \Delta) - L(b + \beta) + \Lambda) + s(y)L'(b - \beta - \Delta) = 0$. Evaluated at the point where $\Delta = 0$, by symmetry $\mathbb{E}(\Gamma(y, \theta)) = \frac{1}{2}$. This gives us $\pi_R(y)(L(b - \beta) - L(b + \beta) + \Lambda) = -\frac{1}{2} L'(b - \beta)$. Assumptions on L ensure that the RHS is decreasing in continuously decreasing in β and positive when $\beta = 0$ and 0 when $b = \beta$. The LHS is 0 when $\beta = 0$ but is increasing in β which gives us the uniqueness of equilibrium. □

Lemma 2.

Proof. Responsiveness of a citizen is a function of $\sigma_{\mathcal{L}}(w) = \sum_k (w^k_{d_{\mathcal{L}}})^2 \frac{1}{\rho^k} + \frac{1}{\rho^P} + \frac{1}{\eta}$. The proof relies on showing that $\sum_k (w^k_{d_{\mathcal{P}}})^2$ is higher for a more specialized citizen. (Note that $w_{d_{\mathcal{P}}} = w_{\mathcal{P}} - w_m$.) Take two citizens such that the projection of their weight vectors to $\mathcal{P}$ is denoted as $y_{\mathcal{P}}$ and $z_{\mathcal{P}}$. Without loss of generality, assume that the weight vectors are ordered in an increasing fashion (so we don't need to introduce notation for the permutations) such that $z_{\mathcal{P}} \geq_{FOSD} y_{\mathcal{P}}$. This means that for any increasing vector v (where $k > l$ implies $v^k \geq v^l$), $z_{\mathcal{P}} \cdot v \geq y_{\mathcal{P}} \cdot v$. This is equivalent to $z_{d_{\mathcal{P}}} \cdot v \geq y_{d_{\mathcal{P}}} \cdot v$. We can choose $v = y_{d_{\mathcal{P}}}$ which implies $y_{d_{\mathcal{P}}} \cdot z_{d_{\mathcal{P}}} \geq ||y_{d_{\mathcal{P}}}||^2$. We can also choose $v = z_{d_{\mathcal{P}}}$ which implies $||z_{d_{\mathcal{P}}}||^2 \geq y_{d_{\mathcal{P}}} \cdot z_{d_{\mathcal{P}}}$. By transitivity, we have the result. □

Proposition 2.

Proof. Recall that $\pi_R(y) = \left|\frac{\partial \mathbb{E}(\Gamma(y,\theta))}{\partial y_R}\right|$ where $\mathbb{E}(\Gamma(y,\theta)) = \int \Phi\left(-\frac{\delta}{\sigma_{\mathcal{L}}(w)}\right) dF(w)$ with $\sigma^2_{\mathcal{L}}(w) = \sum_k (w^k_{d_{\mathcal{L}}})^2 \frac{1}{\rho^k} + \frac{1}{\rho^P} + \frac{1}{\eta}$ and $\delta = \frac{y_R + y_L}{2} - \mu$. $\pi_R(y)$ declines as F shifts distribution towards more specialized voters with $\sigma_{\mathcal{L}}(w)$ higher. □

Proposition 3.

Proof. As shown the proof of Lemma 1, the probability that a citizen of type w supports the correct party decreases with the variance of $w_{d_l} \cdot \theta + \epsilon$ which is equal to $\sum (w^k_{d_l})^2 \frac{1}{\rho^k} + \frac{1}{\eta}$. Increase in η decreases this value. Increases in depth of learning also can only lead to a decreases in $\sum (w^k_{d_l})^2 \frac{1}{\rho^k}$. □

Theorem 1.

Proof. As shown in the proof of Proposition 2, responsiveness to party platforms for any type w decreases when $\sigma_{\mathcal{L}}(w)$ increases. An increase in η leads to a decline in $\sigma_{\mathcal{L}}(w)$, while increased opportunities for specialization weakly increase $\sigma_{\mathcal{L}}(w)$ for any type. □