Spatial Advertisement in Political Campaigns

Anja Prummer

Working Paper No. 805 December 2016 ISSN 1473-0278

School of Economics and Finance

Queen Mary
University of London

Spatial Advertisement in Political Campaigns*

Anja Prummer†

December 5, 2016

Abstract

This paper characterizes the optimal advertising strategy of candidates in an election campaign, where groups of heterogeneous voters are targeted through media outlets. We discuss its effects on the implemented policy and relate it to the well-documented increase in polarization. Additionally, we empirically establish that polarization displays electoral cycles. These cycles emerge in the model as candidates find it optimal to cater to different groups of voters and thus to adjust policies. Further, technologies that allow targeting voters more precisely tend to increase polarization. Our prediction is confirmed empirically as an increase in internet penetration leads to higher polarization.

Keywords: Targeting, Media, Networks, Voting

JEL Classification: D85, D72, D83

*I am deeply indebted to Francesco Nava for his support at each stage of this project. I am grateful to Jérôme Adda, Toke Aidt, Vasco Carvalho, Matt Elliott, Ingvil Gaarder, Elise Gourier, Sanjeev Goyal, Serafin Grundl, Matt Jackson, David Levine, Aniol Llorente-Saguer, Marco Manacorda, Massimo Morelli, Sujoy Mukerji, Sönje Reiche, Carly Urban and Fernando Vega Redondo for helpful comments and advice. I would also like to thank Roberta Dessi, Flavio Toxvaerd, Soheil Mahmoodzadeh as well as seminar participants at PET 2015, SAET 2015, CERGE-EI, University of Exeter, University of Sussex, University of Melbourne, Humboldt University, Queen Mary, University of Manchester, Bilkent University, Middlesex University, Aix-Marseille University, University of Birmigham, Montana State University, Academia Sinica, the 2nd EPEC Workshop, 2nd Annual Conference on Network Science and Economics at Stanford, Rotterdam Political Economy Workshop and NICEP in Nottingham for helpful comments.

†Queen Mary University of London, Mile End Road, London E1 4NS, UK, Telephone +44 (0)20 7882 5555, a.prummer@qmul.ac.uk

1 Introduction

Election campaigns increasingly invest in targeting certain subsets of voters via different media outlets. An article titled "Political Ads Take Targeting to the Next Level" in the Wall Street Journal highlights how this was done in the re-election campaign of Chris Christie, the governor of New Jersey, who wanted to reach Hispanic voters:[1]

> "[it was] *discovered that viewers of "Dama y Obrero", a Spanish-language telenovela about a woman torn between two men, would likely be more receptive to his message than people who watch "Porque el Amor Manda," a romantic comedy.*"

This emphasizes that candidates carefully consider which groups of voters to target, through which media outlets to advertise their platforms.[2] At the same time, Poole and Rosenthal (2000) document that Democrats and Republicans increasingly tend to favor very different policies, both in the House of Representatives as well as in Senate. This disparity in policy preferences is commonly referred to as *polarization*.

We seek to connect the increase in polarization to the advertising strategy of candidates, who tailor their policy platforms to the voters they target. We empirically establish that polarization increases and that it displays electoral cycles, with parties alternatively catering to swing voters or partisans. In order to explain these patterns, we develop a model of advertising with heterogeneous voters. Targeted voters learn the policy platform, whereas non-targeted voters cast their vote based on a prior belief. Candidates cater to voters who are a priori least likely to turn out on their behalf, that is voters whose ideal policy has the greatest distance to the prior.[3] If this prior belief is affected by past policies, it is optimal to target different sets of voters in each election, resulting in electoral cycles in polarization. Importantly, our framework cannot only help understand these cycles, but also the overall trend in polarization. Voters are targeted through media outlets and thus candidates are constrained in their advertising by the media network they face. In line with recent developments in the media landscape, we allow for advertising to become more precise. Candidates can target voters more specifically and tailor their policy to a more narrow subset of voters. This induces candidates to choose more extreme policies, leading to

[1]See the Wall Street Journal article on http://www.wsj.com/articles/political-ads-take-targeting-to-the-next-level-1405381606

[2] An important goal of advertising is to provide information and Freedman et al. (2004), among others, shows that campaigns are indeed crucial in informing voters.

[3]This follows immediately from each voter's concave utility function.

an increase in polarization. Further, we show that even if there are greater spillovers of the campaign among voters, polarization can increase. Finally, we empirically establish the effect of internet penetration as a measure of targeting precision in U.S. Congressional Districts on polarization. Our results confirm that higher internet adoption rates lead to an increase in polarization, in line with our theoretical predictions.

To establish the trend and cycles in polarization, we use the DW-NOMINATE data set, created by Poole and Rosenthal (2000). It provides a measure of parties' policy positions in the U.S. Senate and House of Representatives for each Congress. Based on these positions, we calculate the level of polarization, which is defined as the difference between the parties' policies and we establish, in line with the literature, that polarization has increased sharply since the 1970s. We show that, after removing the time trend, polarization exhibits electoral cycles. This implies that Democratic and Republican positions are closer together for one presidential election, only to be more divisive in the next election followed again by relatively more moderate policies. To the best of our knowledge, these electoral cycles have not been documented so far.

In order to explain both the trend and cycles in polarization, we develop a model of targeting in networks with heterogeneous agents. In our model, two candidates compete to win an election. Both candidates are purely office-motivated and maximize their chance of winning by simultaneously selecting a policy platform they commit to as well as an advertising, or targeting, strategy. Candidates target voters through media outlets. That is, they advertise their platform in certain media outlets and voters that follow these outlets receive information about a candidate's platform. A media network describes which voters can access a given outlet, where we allow for voters to belong to multiple media outlets, see Figure 1 as an example.

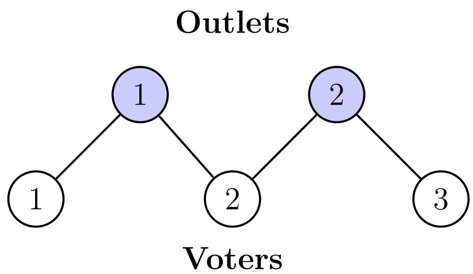


Figure 1: 3 Groups of Voters & 2 Outlets

Voters differ in what policy they prefer. That is, they have heterogeneous bliss points. A voter decides for a candidate according to a standard probabilistic voting model based on what he believes the candidates' platforms are. A voter is more likely to select the candidate whose platform he believes to be closest to his bliss point. Initially, every voter has a prior about the platform each candidate will implement. If a voter is connected to an outlet which is targeted by a candidate, then he learns what the true policy platform of this candidate is. Otherwise, the voter sticks to his prior.

As voters differ in what policies they prefer, audiences of media outlets can be heterogeneous as well. Additionally, we allow for media outlets to differ in how many voters they reach. We show that candidates are more likely to target an outlet if it is has a large audience. However, not only the size of the audience, but also its policy preference influences the choice of outlet. In particular, candidates are interested in targeting voters, whose bliss point is very different from the prior voters have about the candidate's choice. To see this, note that it is never optimal to cater to a group of voters that believes that the party implements their preferred policy. By targeting them, and thus selecting a policy that matches their position, the probability of this group to vote for the candidate remains the same as if the party did not disclose to them. However, selecting voters whose bliss point is very different from the prior they have yields an increase in the probability of voting for the candidate. By selecting these voters and a policy that coincides with their preferred policy, parties can increase the number of their supporters. The choice of such a policy does not alienate the voters whose bliss point is close to the prior as they do not learn about the true policy and therefore stick to the prior, based on which they cast the vote. Advertising thus allows candidates to increase the number of supporting voters, without loosing those already in favor of the party. We then take these considerations together: both the number of voters reached as well as their bliss points matter for the optimal target set.

Based on this we propose a new measure of centrality, which we refer to as *media centrality*. Media centrality captures the trade off between voters' bliss points and the number of voters targeted in a very simple way. This is, to the best of our knowledge, the first measure that explicitly highlights the trade off between an agent's characteristics, here their bliss points, and their position in the network.

We then extend our model to a dynamic setting, where we allow for the prior belief to change in each period. More precisely, we assume that the beliefs of voters regarding the candidate's platform adjust adaptively. If voters do not receive any information about a candidate's platform, they assume that it is going to be the same as last period's policy. We show that a candidate's platform changes with each election. A candidate caters to the voters whose preferred policy has the greatest distance from last period's policy (controlling for the size of the audience). This leads to policy cycles, a result that matches the electoral cycles we established in the data. We then calculate for a given cycle the difference in policy platforms, that is the level of polarization. Our model can thus generate polarization as well as electoral cycles. We also consider an alternative set up in which voters make an inference about the policy if they are not targeted, that is voters are rational, but face uncertainty about the network. We show that the optimal targeting strategy with rational voters and uncertainty yields polarization, but fails to generate electoral cycles. Therefore, our assumption of adaptive beliefs is not only a simplifying one, but also plausible as it matches the relevant stylized facts, unlike a model with rational voters.

Given our definition of polarization across an electoral cycle, we can now connect recent developments in media networks to the policy implemented. We first focus on the increase in the number of media outlets. This has made it easier for voters to find programs that match their interests. As demographic characteristics predict viewership as well as voting behavior, it is now feasible to target a certain set of voters specifically. We show that such a change in the media network generally leads to an increase in polarization. Additionally, it has become easier for voters to be connected to a higher number of media outlets. In particular the internet has created many spillovers. Voters, although not directly targeted, might be able to observe a campaign ad because a friend posts a link on Twitter or Facebook, or a blog reports about it. In this sense, voters are connected to a higher number of media outlets than they have been previously, that is they have a higher number of links to various outlets. We show that this development can also lead to an increase in polarization.

One development in the media landscape has been the increase in internet penetration. The internet has made it easier to target specific groups of voters and so we should see an effect on polarization. To test this hypothesis, we use data from the Federal Communication Commission on internet adoption and connect it to polarization in Congressional Districts. We indeed find that districts with higher levels of internet penetration elect congressmen that are relatively more polarized, confirming our theoretical predictions.

Related Literature Our paper contributes to various strands of literature which we discuss in turn.

Targeting & Advertising First, we contribute to the literature on targeting and advertising. There is a wide range of papers that look at targeting in networks. The key difference to most this literature is that we allow for heterogeneity among the agents in the network.[4] An exception is Groll and Prummer (2015) who also allow for networks with heterogeneous nodes. Their set up does not lend itself to the analysis of arbitrary networks, though. Galeotti and Mattozzi (2011) consider a model of word of mouth communication in which parties target voters in a network, but can only choose one of two policies. Political networks are at the heart of the analysis in Murphy and Shleifer (2004). Their setting, unlike ours, does not allow for overlap between different social communities, which drives opinion divergence among voters and leads to polarization. Work that deals explicitly with microtargeting, without networks, is Schipper and Woo (2014) and Hoffmann et al. (2013, 2014). In the literature on advertising in political campaigns, it is shown that rational voters often cannot infer parties' platforms (Schultz (2007), Coate (2004), Callander and Wilkie (2007)). This finding is in line with the ample empirical evidence that voters are misinformed and make systematic mistakes (Delli Carpini and Keeter (1997), Caplan (2011)).[5] We build on this literature and simply assume that voters cannot infer the platform, which allows us to focus on targeting and policy choices.

Media Our work also contributes to the literature on media and polarization. Different from much of this literature, our focus is not on the decision of the media to slant information.[6] Media is not an active player in our setting, rather, we take the media and its viewers as fixed.[7] Additionally, we assume that politicians do not have the means to influence the voters' decisions about the newspapers they read or the TV programs they watch. Instead, our paper assumes that voters differ in their access to media.

It is empirically well established that media exposure has an impact on voting behavior, which Bernhardt et al. (2008) aim to explain with a focus on the media's reporting decision.[8] They show how differences in information between liberals and conservatives impact their

[4]Early work (Galeotti and Goyal (2009), Dziubiński and Goyal (2013), Goyal and Vigier (2015, 2014)) analyzes the targeting strategy of a monopolist, whereas Goyal et al. (2014) show how two competitors target a network of homogeneous consumers. This is similar to Bimpikis et al. (2013) who consider competition over networks in which firms advertise their products. They highlight that there is a trade-off between an agent's attitude towards a firm and his centrality, but different to our work are unable to give a specific measure that combines centrality and the characteristics of the node.

[5]Based on this models with behavioral voters have been introduced by Esponda and Pouzo (2016), Spiegler (2013), Levy and Razin (2015) and Bisin et al. (2015)

[6]See for example Mullainathan and Shleifer (2005), Baron (2006), Gentzkow and Shapiro (2006), Duggan and Martinelli (2011) and Chan and Suen (2008).

[7]This is different from Garcia-Jimeno and Yildirim (2015) who discuss the strategic interaction between politicians and media outlets.

[8]The effect of media exposure on political outcomes is documented in DellaVigna and Kaplan (2007), Gerber et al. (2009), Gentzkow (2006), Strömberg (2004), Gentzkow and Shapiro (2004), Gentzkow et al. (2006), Gentzkow et al. (2011) and Campante and Hojman (2013).

voting decision, as rational voters are unable to recover full information. In line with our basic idea, Gul and Pesendorfer (2011) connect increased media competition to higher polarization. Different from their set up, in our model all candidates can provide information through each media outlet. Last, our work is related to Glaeser et al. (2005), who document that Republicans choose to divide on religious values in communities that are homogeneous in their ideology. We ask why it is beneficial to target communities with certain characteristics, that is what are the values parties should divide on.

Polarization Additionally, our work aims to explain the increase in polarization. It is well-established that there has been an increase in elite or party polarization, see McCarty et al. (2006). Whether there has also been an increase in polarization in the electorate has been disputed.[9] Even if there has been an increase in polarization among voters, this rise has been moderate compared to the polarization among politicians. Our model provides a novel explanation for this increasing gap between the voters' preferences and the politicians policy choices - the heightened fragmentation of media outlets.[10] Other explanations of polarization are policy-motivation, entry deterrence, incomplete information among voters or candidates, differential candidate valence, politicians catering to their own electorate and increased campaign spending.[11] Unlike our mechanism, these explanations fail to account for the fact that polarization has risen sharply in recent years (the exception being Herrera et al. (2008)).

The remainder of this paper is structured as follows. We first describe the policy patterns we aim to explain in Section 2 and present our model of targeting in Section 3. We show what the equilibrium policy and the optimal targeting strategy are in Section 4. We connect the media network to the optimal policy in Section 5 and analyze trends and cycles in polarization in Section 6. In Section 7 we show that the targeting strategy remains qualitatively the same if voters are rational, but face uncertainty regarding the network structure. Section 8 shows empirically that increased internet adoption leads to higher polarization. Section 9 concludes. All proofs are collected in Appendix A.

[9] Abramowitz and Saunders (2008), Abramowitz and Stone (2006) argue that there has been an increase and Fiorina and Abrams (2008), Fiorina et al. (2008), Fiorina et al. (2005), Fiorina and Levendusky (2006) state that polarization has remained the same.

[10] For an overview on polarization, its causes and consequences, see Layman et al. (2006).

[11] See Wittman (1983), Calvert (1985), Londregan and Romer (1993), Roemer (1994),Osborne and Slivinski (1996), Besley and Coate (1997), Martinelli (2001), Gul and Pesendorfer (2009), Palfrey (1984), Callander (2005), Castanheira (2003), Bernhardt et al. (2007), Callander (2008), Bernhardt and Ingerman (1985), Groseclose (2001), Soubeyran (2009), Krasa and Polborn (2010, 2012), Bierbrauer and Boyer (2013), Eyster and Kittsteiner (2007), Polborn and Snyder (2016), Herrera et al. (2008).

2 Trends and Cycles in Polarization

We first analyze how polarization has changed over time. To do so, we use the DW-NOMINATE data, collected by Poole and Rosenthal based on roll call voting in U.S. Congress.[12] Our analysis is based on the observation that candidates who run for president develop a position, which then becomes the party's new policy platform and is implemented in Congress (Budge and Hofferbert (1990)).

Data Description Poole and Rosenthal collected data on roll-call voting in the House of Representatives and Senate from 1879 to 2013.[13] They use multidimensional scaling techniques in order to project the roll-call voting data on a one- or two-dimensional space. They assume that politicians have symmetric and single-peaked utility functions that are centered around an ideal point. Further, politicians vote probabilistically and assign a higher probability to their preferred outcome. Poole and Rosenthal perform a maximum likelihood estimation, simultaneously estimating the points of the roll-call votes and the ideal points of the politicians on the one- or two-dimensional space.[14] They show that the one-dimensional policy space explains most of the votes and they argue that this dimension is the left-right, socio-economic dimension.[15]

The optimal parameters from the maximum likelihood estimation are then placed on a scale from -1 to 1, with -1 being liberal and 1 being conservative. Ideologically similar politicians are placed close to each other, ideologically different politicians are set further apart. Based on this, we can then calculate the average ideological position of each party over time and the difference in positions gives a measure of polarization. As the scale is bounded between -1 and 1, polarization must lie between 0 and 2. If parties' positions do not differ, polarization is zero.

Polarization We show that there has been an increase in polarization, a result first documented by Poole and Rosenthal (2000). The level of polarization is measured every second

[12]For a detailed description of the data see Keith Poole's website voteview.com, their findings are summarized in Poole and Rosenthal (2000).

[13]Roll-call votes show for each legislator whether he votes "yea" or "nay" on a given bill.

[14]The parameters that emerge from this estimation are those that make the real, observed roll-call votes as likely as possible.

[15]More precisely, Poole and Rosenthal show that the one-dimensional model was correct 83% of the time. This means that the legislator's bliss point was closer to the outcome he voted for than to the outcome he did not vote for (both as estimated by DW-NOMINATE). The two-dimensional model is correct 85% of the time, higher dimensional models do not lead to an improvement. The second dimension is a regional and social dimension.

Figure 2: Polarization

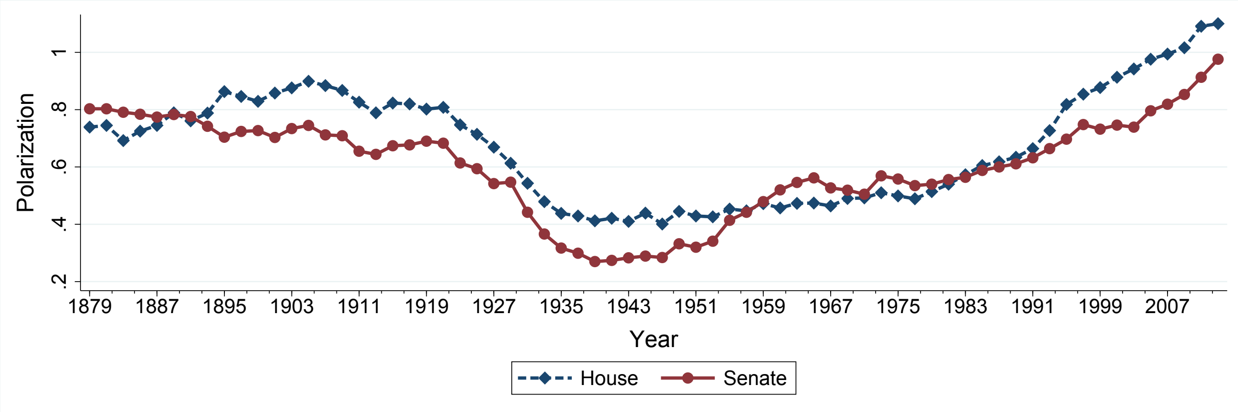
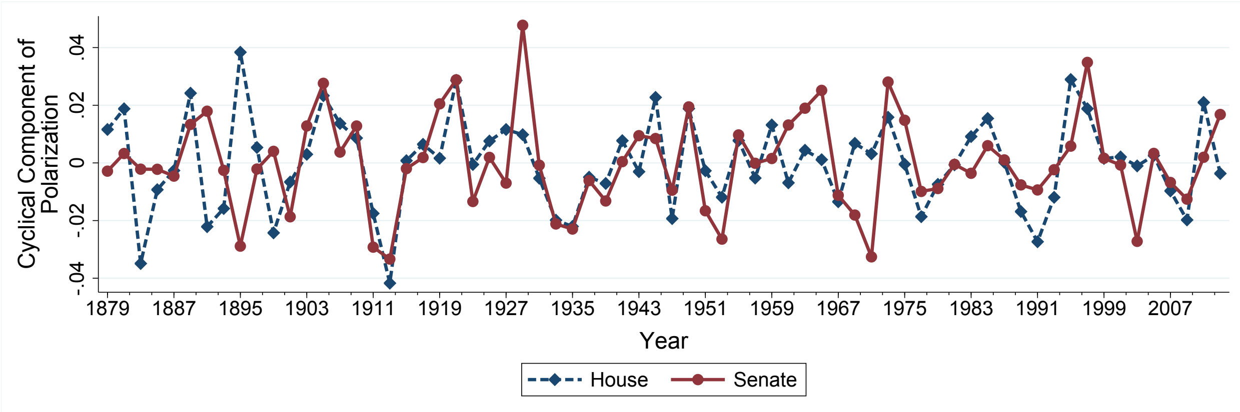


Figure 3: Cyclical Component of Polarization



Note: Cyclical component of polarization from Poole and Rosenthal's measure of polarization obtained by using an HP-filter with smoothing parameter 6.25.

year, for each Congress. Figure 2 highlights that there has been an increasing divergence in the parties' policy positions both in the House of Representatives as well as in the Senate. Polarization was at an all time low after World War II but has steadily increased since then. In particular in recent years the level of polarization has increased steeply.

Going beyond the standard result in the literature regarding polarization, we are additionally interested in the the cyclical component of polarization. We therefore use an HP filter to de-trend the measure of polarization.[16] The cyclical components of polarization for the Senate and House are given in Figure 3. It can be seen that the difference in parties' platforms fluctuates around the trend. In order to analyze this pattern in more detail we

[16] We use an HP filter with a smoothing parameter of 6.25. The parameter of 6.25 was suggested for yearly data by Ravn and Uhlig (2002). To the best of our knowledge there does not exist a benchmark smoothing parameter for biannual data. We therefore report the results for a smoothing parameter of 6.25, but have used a variety of smoothing parameters, both larger and smaller that have not yielded qualitatively different results. One reason for sticking to 6.25 are also the problems with too small of a smoothing parameter as pointed out by Harvey and Trimbur (2008). They show that choosing the parameter too small results in standard errors that are too small and that trends can wrongly absorb cycles.

Table 1: Cyclical Components of Polarization

	Senate	House of Representatives
1st Lag	0.840***	1.414***
	(0.132)	(0.0927)
2nd Lag	-0.412***	-0.636***
	(0.116)	(0.104)
Observations	68	68

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in parentheses

Note: Cyclical Components in polarization from Poole and Rosenthal's DW-NOMINATE scores of first dimension obtained by using an HP-filter with smoothing parameter 6.25. Processes selected according to BIC and AIC were an ARMA (2,1) for the Senate and an ARMA (2,2) for the House of Representatives.

estimate various model specifications and select the best model based on the information criteria (AIC and BIC). The results are given in Table 1, which highlights that polarization is negatively correlated with the second lag. This implies that Democratic and Republican positions are closer together for one presidential election, only to be more divisive in the next election followed again by relatively more moderate policies.[17]

We have thus established that polarization has increased from a value of roughly .6 in the 1970s to a level of above 1, but that it has not done so smoothly, rather there are fluctuations around the trend. These fluctuations in a given year can explain up to 5% of the deviation from the trend.[18] Based on these observations, we aim to develop a model that can help to understand both the trend and cycles in polarization.

3 Model

We introduce a model of probabilistic voting on a network. Two candidates, A and B, aim to win an election while catering to their party. The candidates choose a policy platform and decide how to advertise it in a media network. The targeted audience will observe the policy platform chosen by the politician, but others will decide how to vote based on their prior beliefs about the platform that the candidates intend to implement.

[17]We show that these difference in polarization are indeed caused by parties selecting different platforms in Appendix B. Both Democrats and Republicans in the Senate and the House choose policies that exhibit electoral cycles.

[18]We take both the average of polarization in the House and Senate for the years later than 1990 and show that both for the House and the Senate the fluctuations can on average explain about 1.5 % of the trend. If we take into account the entire time series, then the cycles explain about 2% of the trend.

Voters Policy platforms are contained in the unit interval $[0, 1]$. A voter's utility only depends on the policy implemented and on the identity of the politician implementing it. In particular, the utility of a voter with bliss point x, who believes that candidate c will implement policy y amounts to

$$U(c, y|x) = \begin{cases} u(y|x) & \text{if} \quad c = A \\ u(y|x) + \theta & \text{if} \quad c = B \end{cases}.$$

Voters have an idiosyncratic bias θ in favor of politician B. To simplify later discussions, we assume, as is standard, a quadratic loss function for the baseline policy preferences of voters – so that

$$u(y|x) = 1 - (y - x)^2.$$

The quadratic loss function implies that $u(y|x)$ is single-peaked, concave and symmetric in y.

We further assume, as is common in the literature, that the idiosyncratic biases are uniformly, independently and identically distributed across voters on $[-1, 1]$ – so that $\theta \sim U[-1, 1]$. Thus, a voter with bliss point x would prefer voting for politician A when facing platforms y_A and y_B if and only if

$$u(y_A|x) - u(y_B|x) > \theta.$$

This specification makes the problem symmetric for party B.[19] The probability that a voter with bliss point x votes for A when facing platforms y_A and y_B then simply amounts to

$$P_A(y_A, y_B|x) = \frac{1}{2}\left[1 + u(y_A|x) - u(y_B|x)\right], \tag{1}$$

while the probability of voting for B is given by $P_B(y_A, y_B|x) = 1 - P_A(y_A, y_B|x)$.[20]

Communities, Outlets and the Media Network Voters are partitioned into k *communities*. Let K denote the set of communities, and suppose that a measure 1 of voters belongs to every community. In community $i \in K$, the bliss points of voters are distributed according to a cumulative distribution $G_i(x)$ which admits a density function $g_i(x)$ with support on

[19] Our results hold for any quadratic loss function. If we consider a utility function with $u(y|x) = \alpha_0 - \alpha_1(y - x)^2$ and adjust the uniform distribution appropriately to $\theta \sim U[-\alpha_1, \alpha_1]$, then the candidates' optimization problem is unchanged. Put differently, we only need to adjust the uniform distribution and ensure the difference between the utility of voting for A and B has full support.

[20] This specification implies that $P_A(y, y|x) = P_B(y, y|x) = 1/2$, for any platform $y \in [0, 1]$.

$[0,1]$. Denote its expectation by $E_i(X)$.

There are m possible media outlets where politicians can advertise the platform they select. Let M denote the set of media outlets. The *media network* $\{K, M, N\}$ is a bipartite network that describes which communities have access to a given outlet. Formally, $N \subseteq K \times M$ and community i observes outlet j if and only if $ij \in N$. Denote the neighborhood of media outlet j by $K(j) = \{i \in K \mid ij \in N\}$ Thus, $K(j)$ consists of all communities that can observe a given outlet j. Without loss of generality, assume that every community observes at least 1 outlet – or formally: $K = \cup_{j \in M} K(j)$.[21]

Parties We assume that each candidate needs to cater to the interests of their party in order to gain their support. Parties are instrumental in providing funds for the election campaign, volunteers are needed for canvassing. Party members can choose whether to support their candidate. If a party member does not support his candidate then his utility is given by $\overline{u}$. Thus, a party member of A faces the following utility:

$$U_A(c_A, y_A|x) = \begin{cases} u(y_A|x) & \text{if support} \\ \overline{u} + \theta & \text{if no support} \end{cases} .$$

We impose the same assumptions on θ as previously and we obtain a probability of supporting candidate c by

$$P_c(y_c|x) = \frac{1}{2}\left[1 + u(y_c|x) - \overline{u}\right] \tag{2}$$

Once the party members decide to support their candidate they contribute to the campaign with D.[22] We assume that this holds for both parties. Further, the bliss points of the parties have a cumulative distribution $G_c(x)$, which admits a density function $g_c(x)$ with support on $[0,1]$.[23]

We let the parties of candidate A and B have bliss points $E_A(X)$ and $E_B(X)$, respectively, with $E_A(X) < E_B(X)$. We further assume that every community i is more moderate, $E_A(X) \leq E_i(X) \leq E_B(X)$. This implies, as is common in the literature that party members are more extreme than voters that are not affiliated with any party.

[21] As the model allows for an arbitrary number of communities, any measure of voters could observe any subset of outlets and so there is no loss of generality in assuming that all communities have measure one.

[22] D is a fixed amount and serves as a scaling parameter.

[23] This assumes that the party members are of measure 1. This assumption is without loss of generality as we allow for D to vary.

Candidates: Targeting and Policy Setting Candidates simultaneously select the platform they are going to implement and the media outlets through which to advertise it, so as to maximize their expected vote share as well as the support from their party.[24] Candidates can target an arbitrary number of outlets.[25]

A strategy for a candidate $c \in \{A, B\}$ consist of a pair $\{x_c, T_c\}$. We denote the *policy* of candidate c by $x_c \in [0, 1]$ and the *target* subset of media outlets in which the candidate chooses to advertise by $T_c \subseteq M$. If an outlet is targeted by candidate c, c commits to a platform x_c.[26]

If a community is connected to any outlet targeted by c, every voter belonging to that community knows that the candidate will set platform x_c. However, if a community is not targeted, voters stick to their *prior belief* about the candidates' policy which is denoted by π_c and consists of a single policy in $[0, 1]$.[27]

For any subset of media outlets $T \subseteq M$, define its *coverage* $K(T)$ as the set of communities that are linked to at least one outlet in T. Formally, this set is defined as $K(T) = \bigcup_{j \in T} K(j)$. Denote its cardinality by $k(T)$, and define voters' *posterior beliefs* of community i as

$$y_c^i(T) = \begin{cases} x_c & \text{if} \quad i \in K(T) \\ \pi_c & \text{if} \quad i \notin K(T) \end{cases}$$

Unlike voters, party members always know about the platforms parties commit to. We assume that the candidate's objective function is increasing in both the expected vote share as well as in support D and that it is linear in both elements. The problem faced by candidate $c \in \{A, B\}$ is then given by

$$\max_{x_c, T_c} \left(\textstyle\sum_{i \in K} \frac{1}{k} \int_0^1 P_c(y_A^i(T_A), y_B^i(T_B)|x) dG_i(x) + D \int_0^1 P_c(x_c|x) dG_c(x) \right), \tag{3}$$

where the first part denotes the expected vote share and the second part gives party support.[28]

[24] Maximizing expected vote share, expected plurality (the vote share relative to that of the other party) and maximizing the probability of winning are equivalent in our setting. For a general discussion of the relation between these concepts, see for example Banks and Duggan (2005).

[25] In Appendix C we show an example that highlights some of the implications arising if the number of outlets that can be targeted is restricted.

[26] Note that our results are unchanged if we allow for sequential choice of target set and platform.

[27] Alternatively, we can interpret the policy space as that of a policy dimension. As we have a probabilistic voting model, it is straightforward to extend our framework to a multi-dimensional policy space, where the analysis would then be carried out for each policy dimension separately.

[28] Instead of a utility function that depends on vote share and party support, we could consider a model in which candidates maximize vote share subject to party support being sufficiently large, for example above $\overline{D}$. Then, we could solve the problem using a Lagrangian. Our results remain unchanged, with the exception

An *equilibrium* is thus defined by a strategy profile $\{x_c, T_c\}_{c\in\{A,B\}}$ which solves (3) for every candidate $c \in \{A, B\}$.

4 Equilibrium Policy and Targeting

A candidate faces the problem of which outlets to include in the target set and what policy to set for the targeted communities as well as the party members. In order to characterize the optimal strategy set by the politician, we first simplify expression (3). Given the voting probabilities in expression (1), the maximization problem of candidate c will have the same solution as the following problem which abstracts from the competitor's strategy

$$\max_{x_c,T_c} \int_0^1 u(x_c|x)\textstyle\sum_{i\in K(T_c)} g_i(x)dx + \int_0^1 u(\pi_c|x)\textstyle\sum_{i\in K\setminus K(T_c)} g_i(x)dx + D\int_0^1 u(x_c|x)g_c(x)dx. \tag{4}$$

The first part of expression (4) maximizes the utility of the communities that are contained in the target set, whereas the second part refers to the communities not contained in the target set, which we denote by $K\setminus K(T_c)$. The problem that candidates face is therefore to determine which outlets to target depending on their audience, as well as to choose the optimal policy for the outlets that are targeted, while also taking into account that the party members can always observe the policy. For ease of exposition we initially set $D = 0$. This allows us to focus on the characteristics of the targeted voters and to provide several benchmark results.

We first show that the optimal policy coincides with the average of the bliss points of the communities in the targeted outlets. We denote this expected bliss point by $E(X|T_c)$. We then establish that the optimal targeting strategy prescribes selection of the target set with the highest *media centrality* $W_c(T)$. Formally, media centrality is defined as

$$W_c(T) = k(T)\left[E(X|T) - \pi_c\right]^2, \tag{5}$$

where $k(T)$ denotes the number of communities in the target set. This highlights that the most valuable targets are media outlets with a large coverage, which have viewers whose ideal policies are not aligned with the prior beliefs about the politician's platform. This is summarized in Proposition 1.

of λD replacing D, where λ is the Lagrange multiplier.

Proposition 1. *In every equilibrium $\{x_c, T_c\}$, for every candidate $c \in \{A, B\}$:*

(1) The policy set amounts to the average bliss point in the target's coverage, $x_c = E(X|T_c)$.

(2) The outlets targeted, T_c, maximize media centrality $W_c(T)$.

Candidates select a policy platform that caters to a subset of voters whose bliss point has the greatest distance from the prior. Unless all communities have similar expected bliss points, candidates never choose to target all outlets and all voters, even though there is no cost to targeting. Candidates can essentially segment voters according to their bliss points (subject to the constraint of the media network). They choose a policy that matches the preferences of the voters they target. This induces these voters to cast the vote in their favor. At the same time, the voters who are not targeted do not learn about the platform and vote based on their belief. As the non-targeted voters have a preferred policy close to their prior, this still leads them to vote for the candidate with a high probability. Therefore, candidates can improve their expected vote share, above what they would obtain were they to disclose in all outlets with a policy that is the average of all voters. But it is not straightforward to show how many voters will be targeted. On the one hand adding voters to the target set allows candidates to take their policy preferences into account. This increases their probability of voting on the candidate's behalf. But voters with a bliss point that is further from the prior than that of the newly added voters are now less likely to vote for the candidate as the policy has moved away from their bliss point. The optimal target set balances this trade off and the way this is done is highlighted in the following paragraph.

Characterization of Target Set This discussion aims to further characterize the optimal target set, which will be of use when we connect the implemented policy to the structure of the media network. We write in what follows $T' - T$ if we refer to the *communities* contained in target set T', but not in T. Thus, we denote the expected bliss points of these communities by $E(X|T' - T)$ and their cardinality by $k(T' - T)$.[29] In order to gain some intuition, fix any two target sets satisfying $T' \supset T$ and let the expected bliss points of these target sets lie below the prior, $\max\{E(X|T), E(X|T')\} < \pi_c$.[30] First we establish that target T' has a lower media centrality than a given target set T whenever $E(X|T' - T)$ and $E(X|T)$ lie to different sides of the prior, $E(X|T' - T) \geq \pi_c > E(X|T')$. If this is the case, the candidate prefers

[29]Formally, the distribution of ideal points in $K(T)$ that do not belong to the coverage $K(S)$ of media outlets $S \subseteq M$ is defined as $G(x|T - S) = \frac{1}{|K(T)\setminus K(S)|} \sum_{i \in K(T)\setminus K(S)} G_i(x)$. We denote the expectation of $G(x|T - S)$ by $E(X|T - S)$. For convenience, we also let $k(T - S)$ denote $|K(T)\setminus K(S)|$.

[30]This is equivalent to both bliss points being to the right of the prior.

target set T to T', that is to not include additional voters. Voters in communities contained in the target set T', but not in T are more likely to vote for the candidate if they are not targeted. Their prior is closer to their preferred policy than the actual policy if they were targeted and thus targeting reduces the probability with which they vote for this candidate. Additionally, the voters in target set T are also less likely to vote for the candidate if they observe the policy associated with T' as the new policy is further away from their preferred one. This is formalized in Lemma 1.

Lemma 1. *Let* $\max\{E(X|T), E(X|T')\} < \pi_c$ *and* $E(X|T' - T) \geq \pi_c$. *Then,* $W_c(T) > W_c(T')$.

Outlets may be added to a given target only if $E(X|T')$ and $E(X|T' - T)$ lie to the same side of the prior π_c. The next result presents sufficient conditions for media centrality to be higher in T' than in the original target set T.

Lemma 2. *Let* $E(X|T) < \pi_c$. *If either of the following conditions hold,* $W_c(T') > W_c(T)$:
(1) $E(X|T' - T) \leq [\pi_c + E(X|T)]/2$;
(2) $W_c(T) > W_c(T'')$ *for some* $T'' \subset T$ *with* $E(X|T - T'') \in (E(X|T' - T), \pi_c)$.

The first part establishes that including communities whose bliss point is closer to $E(X|T)$ than to the prior unambiguously increases media centrality. The second part of the result shows that if adding more moderate communities increases the support for a politician, then adding communities with more extreme bliss points also increases support. Suppose there is a target set T''. Then, if it is optimal to add outlets with bliss points $E(X|T - T'')$ such that target set T emerges, then it is also optimal to add any other outlet with $E(X|T' - T) < E(X|T - T'')$. Thus, if media centrality increases by including an outlet, then any remaining outlet which has an expected bliss point below the one included should also be targeted.

When considering a target outlet T' such that $E(X|T' - T) \in ((\pi_c + E(X|T))/2, \pi_c)$, no further simplification is possible. Two forces drive the politician in opposite direction – on the one hand more outlets guarantee more coverage, on the other hand they dilute the effect of disclosure as the new policy is closer to the prior.

Overall Lemma 1 and 2 make it straightforward to determine which outlets will be targeted and which ones are not. They allow us to determine how target sets are affected if the underlying media networks change.

5 Media & Policy Platforms

We aim to capture recent developments in information technologies that have broadened the supply of entertainment and made it easier to access various programs and analyze their effect on the policy set. In particular the internet has created many spillovers. Voters, although not directly targeted, might be able to observe a campaign ad because a friend posts a link on Twitter or Facebook, or a blog reports about it. In this sense, voters are connected to a higher number of media outlets than they have been previously. Put differently, voters have a higher number of links to various outlets. But not only the number of outlets voters are connected to has changed, but also the way in which voters are connected to media outlets. Given the increase in the number of media outlets it has become easier for everyone to find programs that match their interests. As demographic characteristics predict viewership as well as voting behavior, this has made it possible to target a certain set of voters specifically. In order to analyze the effect of this development, we focus on partitions and the impact they have on the implemented policy.

Partitions We allow for the number of media outlets to increase, such that voters can be targeted more specifically and analyze how this affects the implemented policy. We keep the number of communities fixed and assign communities to new outlets without increasing the number of links. This is captured by a partition.

Definition 1 (Partitioned Media Network)**.** *Consider an outlet $j \in M$. A collection $K^P(j)$ is a partition of $K(j)$ if*
(1) $\bigcup_{K_\imath \in K^P(j)} K_\imath = K(j)$
(2) $\emptyset \notin K^P(j)$
(3) $K_\imath \cap K_{\imath'} = \emptyset$ *for any* $K_\imath, K_{\imath'} \in K^P(j)$

A partition splits up the communities belonging to an outlet and assigns them to new outlets. This leads to the creation of a new network and we denote by $K'(\cdot)$, $E'(\cdot)$ and $W'(\cdot)$ the new operators. The created outlets are denoted by $j_\imath \in M^P$, which leads to a new set of outlets $M' = M \cup M^P$. All the partitioned communities $K_\imath \in K^P(j)$ belong to the neighborhood of some $j_\imath$ and so we can write $K_\imath = K'(j_\imath)$.

We first consider the case where an outlet contained in the target set is partitioned. Note that it can always be the case that the expected bliss point of the optimal target set after the partition lies on a different side of the prior, that is it can always be the case that

$E(X|T) < \pi < E'(X|T')$, where T' denotes the target set after the partition. An example of this is given in Appendix C. In what follows we rule out that the target set switches to a different side of the prior. Additionally, we focus on the case where the communities connected to any outlet $j_\imath$ only belong to this outlet, that is there is no overlap with other outlets. Formally, $K'(j_\imath) \cap K'(j) = \emptyset \; \forall j_\imath \in M^P,\, j \in M'$. This is essentially a restriction on how we partition. We can then show that the target set after the partition contains weakly fewer communities.

Lemma 3. *Let $K'(j_\imath) \cap K'(j) = \emptyset \; \forall j_\imath \in M^P,\, j \in M'$. If $\max\{E(X|T), E'(X|T')\} < \pi$, $j \in T$, and $K(j)$ is partitioned, then $K'(T') \subseteq K(T)$ and $E'(X|T') \leq E(X|T)$.*

A partition of an outlet contained in the target set leads to fewer targeted communities. Partitioning allows the candidates to target communities more specifically and therefore communities with a high bliss point will be omitted from the target set. This leads to the expected bliss point of the target set to become more extreme. To see this recall Lemmas 1 and 2, where we have shown that if communities are omitted from the target set, their bliss point lies above $(\pi + E(X|T))/2$ and below π, which results in the bliss point decreasing, that is $E'(X|T') \leq E(X|T)$.

If the outlet that is partitioned does not lie in the target set, then there can be fewer or more communities in the target set, the policy can move closer to the prior or further away. This depends on the specific network characteristics.

Adding Links We turn now to what happens when a community forms an additional link to an outlet. We first focus on the case where a community i contained in target set T forms an additional link to outlet l. As before, we write $K'(\cdot)$, $E'(\cdot)$ etc. to denote the operators of the new network. For our further analysis, we denote by $j \in T$ the outlets to which i belongs in the original network, that is $i \in K(j)$. Again, it can be the case that the new target set has a bliss point to the other side of the prior, that is $E(X|T) < \pi < E'(X|T')$.[31]

If the bliss points of the target set does not switch then the new target set depends on whether the link is formed to an outlet l already contained in the target set (that is, $l \in T$) or whether the link is formed to a previously non-targeted outlet (that is, $l \notin T$). We establish that if $l \in T$ the target set either remains the same or outlets are omitted from the target set. In this case it can never be the case that an outlet is added to the target set (that is,

[31] We provide an example of this in Appendix C.

the target set cannot increase). If $l \notin T$, then it can still be the case that the target set remains the same (it never shrinks), but it can also be that j is omitted from the target set and replaced by outlet l.

Additionally, the change in the target set has an impact on the expected bliss point of the target set and thus also on the policy selected by the candidate.

Lemma 4. *Let* $\max\{E(X|T), E'(X|T')\} < \pi$ *and* $i \in K(T)$

*(**1**)If* $l \in T$*, then* $T' \subseteq T$ *and* $E(X|T) \geq E'(X|T')$*.*

*(**2**)If* $l \notin T$ *and* $j \in T'$*, then* $T = T'$*. If* $j \notin T'$*,* $l \in T'$*.*

If a community, that is already connected to a targeted outlet, forms a link to another outlet in the target set, then it is intuitively plausible that the target set remains unchanged.[32] It is more surprising that the addition of such a link can in fact lead to fewer outlets in the target set. Note that not only outlet j might be dropped from the target set, but other outlets might be dropped as well. To see this consider the example in Figure 4 with $\pi = 1/2$, $E_L(X) = 1/4$, $E_R(X) = 1/2$.

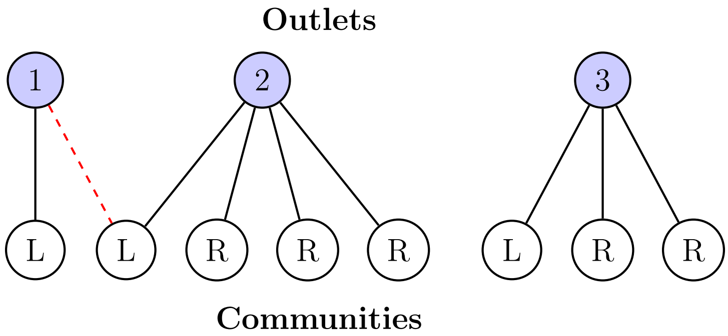


Figure 4: 3 Communities in 2 outlets

Initially, the target set contains all outlets. But after adding the red, dashed link, not only outlet 2 is no longer targeted, but even outlet 3 is dropped from the target set. Again, by Lemma 2, it must be the case that the bliss point decreases, if outlets are dropped from the target set. Similar to the case of the partition, if an already targeted community joins another targeted media outlet, then the candidates can potentially tailor their policy to a more narrow subset of voters. This allows candidates to cater to the most extreme voters (relative to the prior), without losing the vote of the communities that are closer to the prior, as these are no longer informed about the true platform.

[32] An example of this is the case of outlets that have homogeneous communities as their audience. Forming a link has no influence on the expected bliss point of a given neighborhood of the media outlet and therefore the target set will remain the same. If the communities are heterogenous, but the heterogeneity is sufficiently small, the same logic applies.

If a community forms a link to an outlet that does not belong to the target set, then by the same logic as before, it can be that $T = T'$ when the additional link has no impact on the target set. But it can also be that outlet j is replaced by outlet l, an example of this is given in Appendix C.

Last, if a community that is not contained in the target set forms a link, then the target set can increase or decrease, the policy can move closer to the prior or further away. This is similar to the case of the partition, with the added complication that communities can form a link to an outlet contained in the target set or an outlet outside the target set.

This section highlights that it is far from obvious how a change in the media network affects the target set and with it the policy set. In particular, if there is a change in the media network that affects communities and outlets that are not contained in the target set, then the target set and with it the policy can change in an arbitrary way. The policies can increase or decrease relative to the prior and they can lie to the other side of the prior. We can however show how changes within a target set affect the implemented policies.

6 Polarization

We now analyze the differences in the platforms that candidates A and B set. We refer to this discrepancy as *polarization*. Suppose first that party supporters do not matter for the election campaign, $D = 0$. Then, the policies of both candidates can coincide.

Homogeneous Candidates We analyze the benchmark case of *homogeneous* candidates, that is the two candidates are ex-ante identical with, $\pi_c = \pi$ for all $c \in \{A, B\}$. We first establish that in this special case, without parties ($D = 0$), candidates select the same policy and same target set. The expression *symmetric equilibrium* as usual refers to an equilibrium in which both candidates play the same strategy.

Proposition 2. *When candidates are homogeneous:*

(1) *Symmetric equilibria always exist.*

(2) *Both candidates win with probability* $1/2$ *in any equilibrium.*

(3) *In any equilibrium, candidates generically target the same communities* $K(T_A) = K(T_B)$ *and set the same policy* $x_A = x_B$*.*

The generic uniqueness result relies on the observation that no two media coverages can generate the same support for a politician generically. The observation follows, as for any two distinct media coverages $K(T) \neq K(S)$, any small perturbation to bliss points implies that $E(X|T) \neq E(X|S)$ with probability 1.

One implication of this result is that platforms can differ if voters have different beliefs about candidates.[33] But platforms generally differ, even if voters have the same prior about a candidate, if candidates take their own party's preferences into account.

Without party members, the policy set corresponds to the expected bliss point in the target set, $x_c = E(X|T_c)$. This is no longer the case if party members matter for the campaign. Then, the policy weights the expected bliss point in the target set as well as the preferences of the party and is given by

$$x_c = \frac{k(T_c)E(X|T_c) + DE_c(X)}{k(T_c) + D}. \tag{6}$$

We generally do not know where the platforms of parties A and B lie relative to each other, but we can show that if party support is sufficiently important with D high enough, then the platform selected by candidate B lies to the right of A's policy. Party A becomes left-wing, party B right-wing. We define two target sets T_{max} and T_{min}, where $T_{max} \in \arg\max_T k(T)E(X|T)$ and $T_{min} \in \arg\min_T k(T)E(X|T)$.

Proposition 3. *The policy platform of candidate A always lies strictly below that of candidate B, $x_A \leq x_B$, for any prior π_A, π_B, if $D > \overline{D}$ where*

$$\overline{D} = \sqrt{k(T_{max})k(T_{min})}\sqrt{\frac{E(T_{max}) - E(T_{min})}{E_B(X) - E_A(X)}}. \tag{7}$$

Equation (7) highlights that if party support is sufficiently important, namely with D above a threshold value $\overline{D}$, then party A's policies are always to the left of B's policies, *independently of the priors*. The threshold is decreasing in the difference in preferences of the parties. Additionally, a higher number of communities in the target sets lead to a higher threshold as they diminish the influence of the party members. Last, if the difference in preferences of the communities in the target sets are small, then the threshold is lower.

[33] In principle one could analyze polarization by assuming that the priors differ regarding parties. We refrain from doing so as this implies that a party that is perceived to be right-wing will choose a left-wing policy and vice versa. This is not the case if we take party members that support candidates and their preferences into account.

Dynamic Framework Given that polarization is an inherently dynamic phenomenon, we analyze how polarization changes over time.[34] In particular, we allow for the prior beliefs to change over time. If the prior were the same in each election, then parties would select the same policy platform. However, we know that the policies change in each presidential election, as we have established empirically the existence of electoral cycles. Our model can generate these cycles if we allow for beliefs to adjust with each election. In particular, cycles emerge if voters change their beliefs adaptively, that is today's prior equals last period's policy. We denote the prior about candidate c in period t by π_c^t and we index the policy in each period with t. We impose the following assumption.

Assumption 1. *Adaptive Learning: $\pi_c^t = x_c^{t-1}$ for $t > 1$.*

This assumption is in line with a well-established literature on retrospective voting that shows that voters base their decision on who to vote for on past policies, see the seminal book by Fiorina (1981) and the work building on it. Note that despite adaptive expectations, in aggregate, there will be no mistakes in voting. This implies that voters have no incentive to depart from their very simplistic voting strategy.[35]

We further set the discount factor for the candidates to zero. This captures the fact that a candidate cares first and foremost about winning a given election. Put differently, long-term concerns of how the party positions itself are of limited importance. This implies that candidates maximize expected vote share and party support for one period.

We restrict attention to the case where the number of media outlets is at least two, that is $m \geq 2$ and we ask how policy platforms change over time.[36] We first show that it is never optimal to set the same policy in each period.

Proposition 4. *It is never optimal for a party to select the same policy in each election period.*

Selecting exactly the same policy as in the last period prevents parties from segmenting voters. Voters believe that the policy today will be the same as last period's policy. By selecting a

[34] We provide in Appendix C an analysis of how polarization is affected by a change in the media network in a static framework. The previous section has shown that a change in the media network can affect the policies in an arbitrary way. This carries over to polarization and so polarization can increase or decrease. The dynamic framework on the other hand allows us to make sharper predictions.

[35] This holds under the same assumptions that are needed to establish that being able to target voters more precisely tends to increase polarization.

[36] In case of one media outlet, both candidates always target this outlet, but candidates' platforms differ due to their parties.

policy that coincides with the prior, candidates do not single out a subset of voters where they increase the probability that these communities vote for them. This cannot be optimal as we have already established that it is never optimal to target all voters. Note that it does not imply that a given *outlet* cannot be targeted in two consecutive periods. Rather, the policy and with it the target set cannot remain unchanged over time.

This result does not depend on the assumption of adaptive expectations, but only requires that voters adjust their belief sufficiently over time. Thus, our model provides one possible motive of why the electoral cycles we document empirically emerge.

We can further show that these platforms cycle between exactly two policies *in the long run*. We establish that for each $t > \bar{t}$, the optimal platform can take one of two possible values and parties fluctuate between two possible target sets, which we denote by T_L and T_R. The policies are then given by $x_c^t \in \{x_{cL}, x_{cR}\}$ with $x_{cL} < x_{cR}$. However, the electoral cycles are not necessarily unique. There can be two cycles with distinct target sets and *distinct policies*. We denote the policies in the first cycle as x_{cL} and x_{cR}, those in the second one by x'_{cL} and x'_{cR}. There is cycling between x_{cL} and x_{cR} if one of the two platforms is reached and cycling between x'_{cL} and x'_{cR} if either x'_{cL} and x'_{cR} is set at some point.

Lemma 5. *A candidate's electoral cycle has the following properties*

(1) *In the long run, a candidate's platform cycles between exactly two policies.*

(2) *For any two distinct cycles, it must hold that $x_{cL} < x'_{cL} < x_{cR} < x'_{cR}$.*

The target sets are characterized by the number of communities they contain as well as the expected bliss point. However, the size of the target set and the expected bliss points are not correlated.[37]

Therefore, our proof is solely based on how the optimal policies change over time. For a given initial policy, x^1, we distinguish between two possibilities of how the target sets change in the following periods, namely (i) $x^1 < x^2$ and $x^2 > x^3$ or (ii) $x^1 < x^2 < x^3$. We focus on the latter case here and refer the reader for the other case to the proof in the Appendix. We can show that T^4 either equals T^2, in which case we have established that there is cycling between exactly two platforms, or $x^4 \in (x^1, x^2)$. We then turn to the optimal target set

[37]To see this more clearly consider an example with two outlets. We write T_1 (T_2) if only the first (second) outlet is targeted and T_{12} contains both media outlets. There are two cases, namely (i) $x_1 < x_2 < x_{12}$ and (ii) $x_1 < x_{12} < x_2$. An example of case (i) is depicted in Figure 5 where for some parameter values it can be shown to be the case that $x_1 < x_2 < x_{12}$. This holds for example for candidate B if $E_L(X) = 1/4$, $E_R(X) = 5/6$, $E_B(X) = 1$, $D = 2$. Additionally, it holds that $k(T_{12}) > k(T_1)$ and $k(T_{12}) > k(T_2)$ and so the a target set with a more extreme bliss point can be larger or smaller than a target set with a less extreme bliss point.

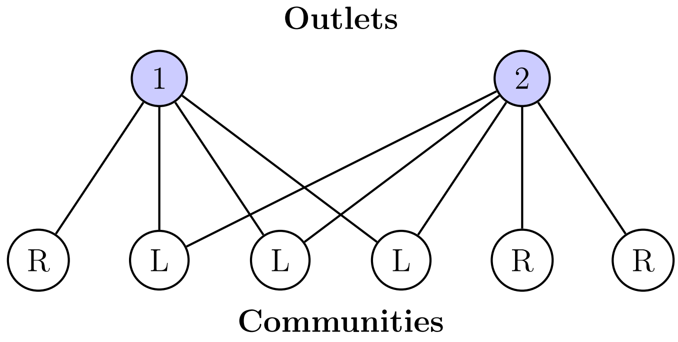


Figure 5: $E(X|T_1) < E(X|T_{12}) < E(X|T_2)$

in period 5. If this equals T_3, we have again established that there is cycling between two policies. Otherwise, $x^5 \in (x^2, x^3)$. Therefore, it must be the case that x_{cL} lies in a bounded subset of the policy space, namely $x_{cL} \in (x^1, x^4)$. Similarly, $x_{cR} \in (x^2, x^5)$. This holds true more generally, that is x_{cL} lies in a bounded subset or the policy space and so does x_{cR}. It further holds that for all $t > t'$ either $x_c^t \leq x_c^{t+2}$ and $x_c^{t+1} \leq x_c^{t+3}$ or $x_c^t \geq x_c^{t+2}$ and $x_c^{t+1} \geq x_c^{t+3}$. This implies that from a certain time period onwards, each policy in the cycle either moves to the left or the right, that is each policy chosen lies weakly to the left or the right of the one two periods before. As the policy space is bounded, this implies that in the long run, there is cycling between exactly two alternatives. However, this cycle does not necessarily have to be unique for each party. It can be the case that there are two cycles which are shifts of each other, that is one cycle contains policies that lie to the left of the policies of the other cycle.

As we have established that in the long run, there are exactly two target sets for each candidate, we can now define polarization over a cycle.

Definition 2 (Polarization)**.** *Polarization is the average distance between parties' platforms for an electoral cycle,* $\Delta_P = \frac{1}{2}(x_{BL} + x_{BR} - x_{AL} - x_{AR})$.

It is straightforward to see that polarization is increasing in D and $E_B(X)$ and decreasing in $E_A(X)$. The level of polarization also depends on the network structure. We have shown that a partition as well as adding links affect the target set as well as the policy set. Therefore, these changes also have an implication on the difference between policies and therefore polarization.

Symmetric Media Outlets In order to analyze the impact of a partition or of adding links, we restrict attention to symmetric media outlets. This allows us to connect the features of the media network to polarization.

Assumption 2 (Symmetry)**.** *Media outlets as well as party preferences are symmetric.*

A.1 *For any outlets $j, j' \in M$, there exist outlets $j'', j''' \in M$ such that*

$$k(j - j') = k(j'' - j''') \quad \textit{and} \quad 1 - E(X|j - j') = E(X|j'' - j''')$$

A.2 *Parties have symmetric preferences, $E_B(X) = 1 - E_A(X)$.*

Assumption **A.1** implies that for each media outlet that lies to the left of $1/2$, there exists a symmetric media outlet to the right of one half. It presumes that the distribution of voters among media outlets is the same on the left and on the right of $1/2$. We assume this symmetry also holds for party members (Assumption **A.2**).

Based on these assumptions we can then establish that there always exists an equilibrium, in which the policies are symmetric.

Lemma 6 (Symmetric Policy Platforms)**.** *If media outlets and party members are symmetric, there exists an equilibrium in which policies are symmetric, that is $x_{AL} = 1 - x_{BR}$ and $x_{AR} = 1 - x_{BL}$ for $D \geq 0$.*

Lemma 6 shows that there exists an equilibrium in which policies are symmetric. However, it can still be the case that there are asymmetric equilibria. As an electoral cycle is not necessarily unique, see Lemma 5, it can be the case that the two parties end up with different policy cycles. This results in asymmetric equilibria, where $x_{AL} \neq 1 - x_{BR}$ and $x_{AR} \neq 1 - x_{BL}$. But even if policies are symmetric, it might not necessarily be the case that the target sets are the same. What target sets emerge matters as a change in the media network, either through a partition or by adding links affects the target sets which in turn impacts the chosen policies which then leads to higher or lower levels of polarization. We show that there are two possibilities, namely (i) target sets are the same for both parties or (ii) target sets differ, but are still symmetric.[38]

To see this consider Figure 6, which gives the symmetric bliss points associated with 4 different target sets, *without* taking party preferences into account. We rank the bliss points with $E(X|T_{-2}) < E(X|T_{-1}) < E(X|T_1) < E(X|T_2)$. Once we also take party preferences into account, we obtain the actual policies which are as given in Figure 7. Note that it is not necessarily the case that the ordering of the bliss points carries over to the policies. Rather, it can be the case that a target set associated with a higher bliss point will lead

[38] In any case, the policies are symmetric around $1/2$.

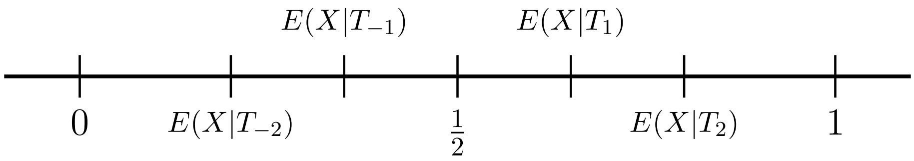


Figure 6: Bliss Points with Symmetric Media Outlets

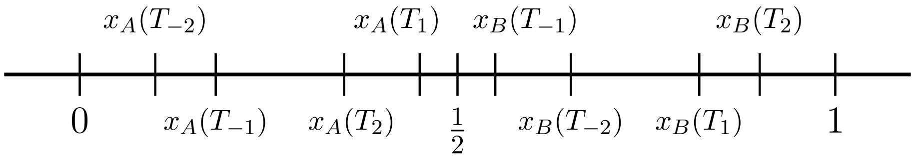


Figure 7: Policies with Symmetric Media Outlets

to a lower policy than another target set with a lower bliss point. To see this recall that policies are given by $x_c(T) = \frac{\sum_{i \in T} E_i(X) + DE_c(X)}{k(T)+D}$. Therefore, not only the bliss points in the target set matters, but also the number of communities that belong to the target set. If a target set contains only few communities, then the impact of the party members is larger. They have a weight of D and each community contains a measure one of voters. If there is only one community in the target set, then the candidates set policies that cater greatly to their own party. If the number of the communities is high, then the influence of the party is limited and the implemented policy places a higher weight on voter's preferences. From this it follows that the policy associated with T_2 might lie below the one selected when the target set is T_1 despite the communities in target set T_2 having a higher bliss point than those in T_1. This is possible if the number of communities in T_2 is sufficiently low compared to those in T_1, resulting in the interest group gaining more influence if T_2 is targeted than if T_1 is chosen. Formally, if $k(T_1) > k(T_2)$ and $E(X|T_1) < E(X|T_2)$, then it can be the case that $x_A(T_2) < x_A(T_1)$. By symmetry, for candidate B the policies associated with target sets T_{-1} and T_{-2} have a reversed ranking compared to the bliss points as well.

This then allows us to discuss what target sets candidates select. First, parties might select the same target sets. Suppose, for example, that party A selects T_{-1} and T_1. Then, by symmetry, there exists an equilibrium in which T_{-1} and T_1 are also optimal for party B. But it can also be the case that both candidates select different target sets. Suppose, for example, that party A cycles between T_{-2} and T_1. Then, by symmetry, T_{-1} and T_2 are also optimal for party B. This still implies that the policy platforms are symmetric, as target set T_{-1} contains the same number of communities as target set T_1 and the bliss points are symmetric around 1/2, that is $E(X|T_1) = 1 - E(X|T_{-1})$. The same holds true for T_{-2} and T_2.

Based on this we can simplify the level of polarization. We denote by T_{BL}, the left target set of candidate B, by T_{BR} his right one. Polarization is then given by

$$\begin{aligned}\Delta_P =& \frac{1}{2}\Bigg(\frac{2\left(\sum_{i\in K(T_{BL})} E_i(X)\right) - k(T_{BL})}{k(T_{BL}) + D} + \frac{2\left(\sum_{i\in K(T_{BR})} E_i(X)\right) - k(T_{BR})}{k(T_{BR}) + D}\\ &+ D\left(E_B(X) - E_A(X)\right)\left(\frac{1}{k(T_{BL}) + D} + \frac{1}{k(T_{BR}) + D}\right)\Bigg), \end{aligned} \tag{8}$$

where we omit the target sets of candidate A due to symmetry. This allows us to characterize how different symmetric equilibria affect the level of polarization.

Lemma 7. *In equilibria with symmetric policies, higher levels of polarization are associated with an increase in $\sum_{i\in K(T_{BL})} E_i(X)$ and $\sum_{i\in K(T_{BR})} E_i(X)$ and with a decrease in $k(T_{BL})$ and $k(T_{BR})$.*

Lemma 7 highlights which features of an equilibrium impact the level of polarization.[39] In particular, if a symmetric equilibrium arises in which the bliss points in the target sets candidate B selects are higher, then polarization increases. Higher bliss points in the target sets for B imply lower bliss points in the target sets selected by candidate A and thus the gap between policies chosen grows, which results in higher polarization. At the same time, polarization is more moderate, in an equilibrium in which target sets contain a higher number of voters. Recall that polarization arises due to party preferences. A higher number of voters in the target set diminishes the influence of the party on the policy set and therefore moderates the policies selected. This ultimately leads to a decrease in polarization. However not only the overall number of communities contained in the target sets matters, but also how many communities are contained in the left and right target set, respectively.

To simplify our further analysis, we assume that each outlet is contained in at least one of the two target sets of each party. That is, a given outlet must either be included in the left or right target set.

Assumption 3 (Exclusion of Non-Targeting)**.** *For any two target sets T_L and T_R, with $E(X|T_L) < E(X|T_R)$, there does not exist an outlet $j \in M$ such that $E(X|j - T_L) > \frac{x_{cL}+x_{cR}}{2}$ and $E(X|j - T_R) < \frac{x_{cL}+x_{cR}}{2}$.*

[39] This is not a comparative static result, rather we illustrate features of the equilibrium.

Recall that Lemma 2 established that any outlet with communities that have a bliss point closer to the policy than the prior should be added to the target set. Here, the prior is last period's policy. Therefore, if the bliss points of the communities in a given outlet are closer to the policy associated with the left target set, then the outlet should be added to this target set. Otherwise, the outlet should be added to the right target set. However, it can be the case that there are outlets that are never added to a target set as they have a sufficiently large overlap with other outlets. An example of this is given in Figure 8. The majority of left communities already belongs to T_L and the same holds true for the right communities that are also connected to outlets already contained in T_R. These type of outlets would not be contained in any target set and we rule that this type of outlets exist. We essentially focus on media networks in which each outlet has a sufficiently large voter base that is only connected to it.[40]

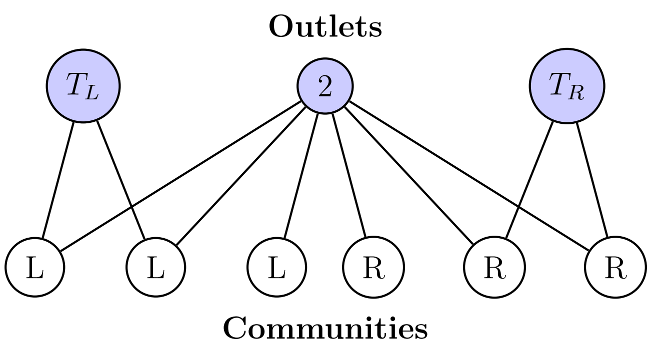


Figure 8: Exclusion Non-Targeting

One implication of Assumptions 2 and 3 is that the target sets associated with the more moderate policies contain weakly more voters than those with the more extreme policies. This is summarized in Lemma 8.

Lemma 8. *Let $D > \overline{D}$. Then, the number of voters in B's left target set is weakly larger than the number of voters in his right target set, $k(T_{BL}) \geq k(T_{BR})$.*

By symmetry it then also must be the case that there are more voters in A's right target set than in his left target set. The result follows from the fact that B's left target set lies to

[40]Note that imposing this assumption on the long run policies is less restrictive than doing so for two arbitrary target sets. More precisely, this assumption will not hold for arbitrary target sets. To see this consider three policies that are chosen over time with $x^1 < x^2 < x^3$, with each policy becoming next period's prior. Then the target set associated with x^2 does not contain outlets with communities that have a bliss point below x^1. Similarly, the target set associated with x^3 does not contain outlets with communities that have a bliss point below x^2. This implies that there are outlets not contained in the target set associated with x^2 or x^3.

the right of $\frac{1}{2}$. By Lemma 2, it must be that any outlet with a bliss point to the left of this target set must be contained in T_{BL}. As half of the communities lie to the left of 1/2 due to the symmetry of the media outlets, the number of voters in T_{BL} must be weakly larger than the number of voters in T_{BR}.

For equilibria with symmetric policies and networks which satisfy Assumption 3, we can then analyze the effect of a change in the media network on polarization. We know that each outlet is contained in a target set and therefore, we can use the results developed in Section 5. There we have shown that as long as the changes in the media network occur within a target set, we can analyze their impact on policy. This carries over to the effect on polarization.

Partitions & Polarization We now discuss the impact of a partition on polarization if candidates select the same target sets before turning to the case of different, but symmetric target sets. We consider two outlets, j_- and j_+ that are symmetric, that is $k(j_-) = k(j_+)$ and $E(X|j_-) = 1 - E(X|j_+)$ and let the communities connected to these outlets be *partitioned symmetrically.*

Definition 3 (Symmetric Partitions). *Two symmetric outlets j_- and j_+ are partitioned symmetrically if for every outlet $j_{-\imath}$, there exists an outlet $j_{+\imath}$, such that $k(j_{-\imath}) = k(j_{+\imath})$ and $E(X|j_{-\imath}) = 1 - E(X|j_{+\imath})$.*

By focussing on symmetric partitions, we preserve the symmetry of the underlying media outlets. If partitioning has no effect on the target set, then the policies set are unchanged and so is polarization. However, it can be the case that an outlet $j_\imath$ is omitted from one of the target sets and then polarization is affected.[41]

This allows us to turn to our main result. We distinguish between a setting where candidates of both parties target the same voters and when they target different voters.[42] In U.S. presidential elections only few states are contested and so candidates tend to target the same states. However, within these states it can be either more important for a party to capture ideologically moderate swing voters or to cater to partisans, with empirical evidence

[41]It can be the case that several outlets are omitted from the target set. This does not change the analysis as long as for any omitted outlet j', $K'(j') \cap K(j) = \emptyset \ \forall j', j \in M'$. By partitioning appropriately, we can always ensure that this condition holds and thus we can focus on outlet $j_\imath$.

[42]The fact that candidates target either the same or different voters is an intermediate results and depends on the specific network structure.

supporting both types of targeting strategies.[43] Our model can generate distinct targeting strategies and thus reconciles the different empirical findings. If candidates target the same voters, then partitioning weakly increases polarization. If, on the other hand, candidates target different voters, then the effect on polarization depends on how the media network is partitioned. To see this, note that there are two cases, namely (i) both partitioned outlets are contained in T_{BL} and (ii) one partitioned outlet is contained in T_{BL}, the other in T_{BR}. We can focus on these two cases as it is never optimal to include two symmetric outlets in B's right target set, see Lemma 8. We show that if both partitioned outlets are contained in $B's$ left target, then polarization can increase or decrease. If outlets in both target sets are partitioned, then polarization unambiguously increases. Put differently, polarization can decrease if only moderate voters select new outlets. If however, due to the partition, also more extreme voters can be targeted specifically, then polarization increases. We believe the latter case to be more reasonable, due to the "Theory of the Long Tail". This idea has been put forward by Anderson (2006). He argues that the number of differentiated goods that can be offered is increasing. In particular, it has become easier to offer niche products as cost constraints disappear. This implies that for new entrants it is often more attractive to produce products that satisfy the preferences of a minority rather than to offer another mainstream product. In line with this theory, there has been an emergence of new media channels with highly specific programmes in the US. As an example, Fox news caters mainly to conservative viewers. The same pattern holds true for TV programs, for example televangelists that tend to cater to a conservative minority. Moreover, individuals' preferences and their political attitudes are correlated, as is well-documented by the initial quote about Chris Christie's re-election campaign. We therefore find it plausible that it has not only become easier to identify and reach moderate voters, but also more extreme ones. Then, polarization increases as the media network becomes more fragmented.

Based on this we now turn to the formal statement. We fix the preferred policies of parties A and B at zero and one, respectively if candidates choose different target sets. We further impose Assumptions 2 and 3, which then allows us to connect media fragmentation and polarization in Proposition 5.

Proposition 5. *Let $K'(j_i) \cap K'(j) = \emptyset \ \forall j_i \in M^P$, $j \in M'$ and let partitions be symmetric.*
(1) *If both candidates select the same target sets, partitioning weakly increases polarization.*

[43] Cox (2009) provides an overview of the empirical evidence of targeting swing voters versus partisans and highlights that there is fairly equal evidence for both targeting strategies.

(2) Let candidates select different target sets and suppose $E_B(X) = 1$.

(a) Let $j_-, j_+ \in T_{BL}$. If $E(X|j_\imath)$ is sufficiently high, partitioning increases polarization, otherwise polarization weakly decreases.

(b) If $j_- \in T_{BL}$ and $j_+ \in T_{BR}$, polarization weakly increases.

Independently of whether candidates select the same target sets or different ones, it can always be the case that the target sets are unaffected by the partition. Then, polarization is unchanged as well. If partitioning does affect the target sets, then the induced change and its effect on polarization depends on the type of equilibrium.

Suppose first that both candidates select the same target sets. Then it can either be the case that communities switch target sets, that is communities originally in the left target set move to the right one and vice verse. Alternatively, it can be the case that communities move from B's right target set to his left one. It cannot occur that communities initially contained in the left target set move to the right target set, again due to Lemma 8. Additionally, it cannot be the case that the partitioned outlets belong to both target sets.

We now connect the changes in the target sets to polarization. Suppose first that outlets switched. Then, polarization remains unchanged. Target sets still remain symmetric and so polarization simplifies to

$$\Delta_P = \frac{D}{k(T) + D} \left(E_B(X) - E_A(X) \right). \tag{9}$$

Equation (9) shows that only the number of communities in the target sets matters for the level of polarization and as this number is unchanged, polarization is unaffected by the change in the target sets.[44] Intuitively, the resorting of communities in target sets does not have an effect on polarization as the increase in B's right policy is set off by the same increase in $A's$ right policy. The same holds for both candidate's left policies and so polarization remains unchanged.

We then turn to the case where the partitioned outlet moves from the right target set to the left. This leads to an increase in the number of communities contained in the left target set and a decrease in the number of communities in the right target set, again from the perspective of candidate B. For candidate A the change is reversed with the communities in the right target set increasing and those in the left one decreasing. This implies that candidates no longer select the same target sets. We can show that the policy associated

[44] We write here $k(T)$ and drop the subscript as both target sets are equal sized.

with the right targeted sets becomes more extreme whereas the policy associated with the left target set can be more or less extreme. Overall, polarization increases as B's right policy increases more than B's left policy decreases. There are two different forces at play. First, as the number of communities associated with B's right target set has shrunk, this leads to a greater influence of the party. As the party members have more extreme preferences, this leads the policy associated with B's right target set to become more extreme, moving closer to one. By symmetry, the policy of candidate A associated with its left target set moves towards zero. This leads to an increase in polarization. Second, as the number of communities in B's right target set has decreased, only the communities whose preferences are more extreme remain, see Lemmas 1 and 2. The candidates therefore cater to more extreme voters, choosing more extreme policies leading again to higher polarization. Thus we have established the first part of Proposition 5, namely that partitioning weakly increases polarization if candidates select the same target sets.

We then turn to the effect of a partition on polarization if candidates do not select the same target sets. If one partitioned outlet is contained in T_{BL}, and the other in T_{BR}, then the logic is exactly the same as if candidates select the same target sets. We therefore focus on what happens if both partitioned outlets are contained in B's left target set, and by symmetry, $A's$ right target set. If both outlets are contained in T_{BL}, then an outlet might be omitted from the left target set and added to the right target set. If this happens, polarization increases if the bliss point of the communities added to the right target set is sufficiently high. Otherwise polarization decreases. There are two opposing forces at play. On the one hand, increasing the number of communities in the right target set decreases polarization, as the influence of the party is reduced. At the same time adding communities that have an extreme bliss point to the right target set induces a more extreme policy. If the second effect outweighs the first one, polarization increases.

Adding Links & Polarization We again restrict attention to the case of of symmetric media outlets and symmetric equilibria. We focus on symmetric outlets j_- and j_+ where $E(X|j_+) = 1 - E(X|j_-)$ and $k(j_+) = k(j_-)$. The same conditions hold for l_- and l_+, the outlets communities connect to. Similarly, communities i_- and i_+ are symmetric if $E_{i_+}(X) = 1 - E_{i_-}(X)$. We aim to preserve the symmetry of the media network and therefore add links *symmetrically.*

Definition 4 (Symmetrically Added Links)**.** *If a community $i_+ \in K(j_+)$ forms an additional* *link to outlet l_+, the community $i_- \in K(j_-)$ forms a link to outlet l_-.*

We then consider how polarization is affected by symmetrically adding links.

Proposition 6. *Suppose communities i_-, i_+ form an additional link. Then, polarization can increase or decrease.*

The key departure in the analysis of adding links relative to the case of partitions is that more communities can now belong to both target sets. This feature per se decreases polarization. Also, eventually it must be the case that polarization decreases as in a complete network, where all voters observe every outlet, polarization is minimized. Nevertheless, it can be the case that polarization increases. To see this we focus on the case where candidates select the same target set. If communities form links across target sets, then it can be that (i) target sets remain unchanged, (ii) outlets switch between target sets symmetrically and (iii) candidates no longer choose the same target sets. If target sets are unchanged, then polarization *decreases* as the overall number of targeted communities, $k(T_{BL}) + k(T_{BR})$ has increased. The same holds true if outlets switch between target sets symmetrically, again as $k(T_{BL}) + k(T_{BR})$ increases and both target sets are of equal size. This is no longer true if candidates choose different target sets after the links are added. In this case it can then occur that polarization increases. In particular, it can be that communities switch from the right target set to the left target set, which also in the case of adding links can lead to an increase in polarization.

Proposition 6 highlights that there is a non-monotonicity in how adding links affects the level of polarization. This implies that even if new technologies have made it possible for voters to better observe politicians campaigns, this does not necessarily lead to policies that cater to a greater set of voters. Instead, candidates adjust their campaign strategies, which ultimately can result in a higher level of polarization.

7 Rational Voters & Uncertainty

We assume in the baseline model that voters do not make an inference about the policy if they are not targeted. In what follows, we allow for voters to be rational and to make this inference while simultaneously introducing uncertainty about the environment. Unlike voters, candidates know the realization of the state, that is they know the exact structure and bliss points of communities in the network.

Voters face a set of states, which describe the structure of the media networks *as well as* the bliss point of a given community. This implies that we are allowing voters not only to be uncertain about the overall structure of the media outlet, they also do not know the bliss point of their own community.[45] This set of states is denoted by $S = \{1, 2, \ldots, n\}$. A realization of the state is given by s. The probability that a state s has been realized is p_s.

In each state s a candidate selects an optimal policy x_s and a target set T_s, taking into account the beliefs of the voters.[46] If a group of voters is targeted, they learn about the policy that is being implemented. If a community i is not targeted, the voters use Bayesian updating to assign a posterior probability to each possible policy, which we denote by q_{is}. We provide further details in Appendix C. We can show that without parties ($D = 0$) the optimal target set *minimizes*

$$\left(k(T_{\tilde{s}}) + \sum_{i \in K \setminus K(T_{\tilde{s}})} q_{i\tilde{s}}\right) \left(E^w(X^2|T_{\tilde{s}}) - E^w(X|T_{\tilde{s}})^2\right) + \int_0^1 \sum_{s \neq \tilde{s}, s \in S} (x_s - x)^2 \sum_{i \in K \setminus K(T_{\tilde{s}})} q_{is} g_i(x) dx, \tag{10}$$

where $E^w(X|T_{\tilde{s}})$ is the expected bliss point of the target set also taking into account the preferences of the non-targeted communities (weighted by their posterior).[47] $E^w(X^2|T_{\tilde{s}})$ the expectation of these bliss points squared. We denote the realized state of the world by $\tilde{s}$. The two parts of equation (10) are two measures of variance. The first part of the equation measures the loss from informing voters, the second part entails the loss of votes from not disclosing. To see the difference to the model where voters are not fully rational, note that in this case the optimal target set minimizes

$$k(T) \left(E(X^2|T) - E(X|T)^2\right) + \int_0^1 (\pi_c - x)^2 \sum\nolimits_{i \in K \setminus K(T)} g_i(x) dx \tag{11}$$

If voters are rational, then candidates take into account that not-disclosing reveals some information and therefore adjust their policy to also cater to some extent to non-targeted voters. However, as long as q_{is} lies below one, that is voters cannot perfectly infer the policy, candidates still place a greater weight on the preferences of the voters they target and the basic targeting strategy that is optimal with voters who do not make an inference if not

[45] This seems realistic as voters only have limited information about the audience of a given TV show.

[46] We assume that there are no two states with the same policy and/or same target set. It is always feasible to find conditions on the states and the network structure that occurs that guarantees this.

[47] For an explicit definition of $E^w(X|T_{\tilde{s}})$ and $E^w(X^2|T_{\tilde{s}})$, see Appendix A, Proof of Proposition 7.

targeted is still valid for Bayesian agents. To see this more clearly, we discuss an example in more detail in Appendix C.

This result implies that the policy patterns with voters who do not make an inference if they are not targeted are exactly the same as in the case of rational voters. The only difference is the extent of the polarization which is greater in our baseline model. Thus, at one extreme voters can perfectly infer the policy which is the case in standard policy setting models. This leads to the mean voter theorem with probabilistic voting. At the other extreme, voters cannot infer the policy if they are not targeted as is the case in the model presented here. This implies more extreme policies in which only the preferences of a subset of the voters are reflected. Last, a setting with uncertainty and Bayesian voters leads to a policy in which the preferences of some voters are taken into account more than that of others. Again, the policy set departs from the mean voter theorem, but the deviation is not as extreme in our example as with the unsophisticated voters in the baseline model.

One key assumption we have made so far is that there is no state in which voters are not targeted. Suppose instead that such a state exists with probability $\overline{q}$ and voters think the implemented policy equals π in this state. Equation (10) can then be written as

$$\left(k(T_{\tilde{s}}) + \sum_{i\in K\setminus K(T_{\tilde{s}})} q_{i\tilde{s}}\right)\left(E(X^2|T_{\tilde{s}}) - E(X|T_{\tilde{s}})^2\right) + \int_0^1 \sum_{i\in K\setminus K(T_{\tilde{s}})} \left(\sum_{s\neq\tilde{s}, s\in S} (x_s - x)^2 q_{is} + (\pi - x)^2\overline{q}\right) g_i(x)dx.$$

We can show that the media centrality we develop is the limit case of the setting with rational voters if $\overline{q} \to 1$.

Proposition 7. *Let $\overline{q} \to 1$. Then, the optimal targeting strategies in the setting with myopic voters and the setting with rational voters and uncertainty coincides with equations (10) and (11) being equivalent.*

This highlights, in line with existing papers (Schultz (2007), Callander and Wilkie (2007), Bernhardt et al. (2008)), that even when voters are rational, they are unable to infer the optimal targeting strategy as well as the policy, which results in polarization. Note however that the model with rational voters fails to account for the electoral cycles. While different policies will emerge if states change, we would not expect this to lead to the cycles we observe in the data. Therefore, our model with behavioral voters performs better at matching the stylized facts derived from the data.

8 Internet & Polarization

The key prediction of the paper is that media fragmentation leads to a higher level of polarization. One measure of how easy it is to target specific groups of voters is internet penetration, and so we ask whether higher internet penetration increases polarization. We test this hypothesis using data on internet penetration collected by the Federal Communications Commission (FCC) for each county in the U.S. In order to match the counties to U.S. congressional districts, we use Census data. To measure polarization we use again the DW-NOMINATE scores, which have already been described in Section 2.

Data Description The FCC collects data for each county on three measures of internet penetration.[48] The data comprises of information about how many households in a given county have a fixed high-speed connection over 200 kbps and how many households have a speed of 3 Mbps downstream and 768 per kbps upstream (broadband). Additionally, the data provides information on the overall number of internet providers in a given county. In previous studies the number of internet providers has been the best measure of internet penetration.[49] However, the available data from the FCC allows a more direct measure of actual internet connection and speed and therefore improves on previous studies. All three measures of internet penetration are collected bi-annually since December 2008. We use the data until 2013.[50] We then need to match the counties to congressional districts. We do

Table 2: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.
Dist. Mean. Pol.	0.527	0.165	0.016	1.208
Mean Polarization	0.172	0.062	0.084	0.216
# Internet Providers	26.759	11.549	7.166	62.5
Internet (200MB)	0.787	0.105	0.421	1
Broadband	0.611	0.139	0.295	1
Percent White	72.445	14.844	20.8	96.867
Median Age	37.357	2.955	27.179	50.776
Owner Occupied	65.163	8.897	18.5	81.146
Observations		1304		

[48] For detailed information, see https://www.fcc.gov/general/form-477-county-data-internet-access-services

[49] For a discussion of this, see Larcinese and Miner (2012) and Kolko (2010).

[50] Only for these years we have all three measures of internet penetration.

so using 2010 Census data.[51] The data set contains information about the congressional districts. As we only have information about internet penetration for a limited number of years, we restrict attention to the 111th to the 113th Congress, which comprises of years 2009-2015.[52] For these years of congress, we use the DW-NOMINATE score for each member of the House of Representative. The descriptive statics of internet penetration, polarization and controls are depicted in Table 2, where owner occupied denotes the houses that are occupied by the owner.

Whether households in a county have internet access, low speed or broadband, is measured on a scale from 1 to 5. A county has a value of one if less than 200 out of 1000 households have internet access, two if less than 400 out of 1000 households have internet access etc. To simplify the interpretation of the coefficients, we redefine the ordinal scales to percentages, with one being replaced by 20%, two by 40% etc. As we are averaging across time over internet penetration rates, we get an average percentage of internet penetration, which is no longer ordinal. Additionally, we include some controls. We consider the share of voters that are white. Further, we include median age as well as the percentage of houses that are owner occupied to have a measure of wealth in a congressional district.[53]

We calculate the mean of polarization for each congress and consider the deviation from this mean. We document what this looks like for the 113th Congress, see Figure 9. Figure 9a shows the individual levels of polarization in the House of Representative. It can be seen that the Democrats lie clearly to the left of zero and the Republicans are to the right. Further, Republicans are more extreme compared to Democrats. The black line depicts the mean of polarization. Subtracting the individual levels of polarization from the mean and taking absolute values, gives Figure 9b. This is our dependent variable in the regression.

Identification Strategy We have information about internet penetration and polarization for three consecutive congresses. To exploit the panel structure of the available data, we estimate a fixed effects model with fixed effects at congressional district level. We further impose clustered standard errors, which are clustered at the congressional district level. Due to our fixed effect estimation, we only use the variation within each congressional district

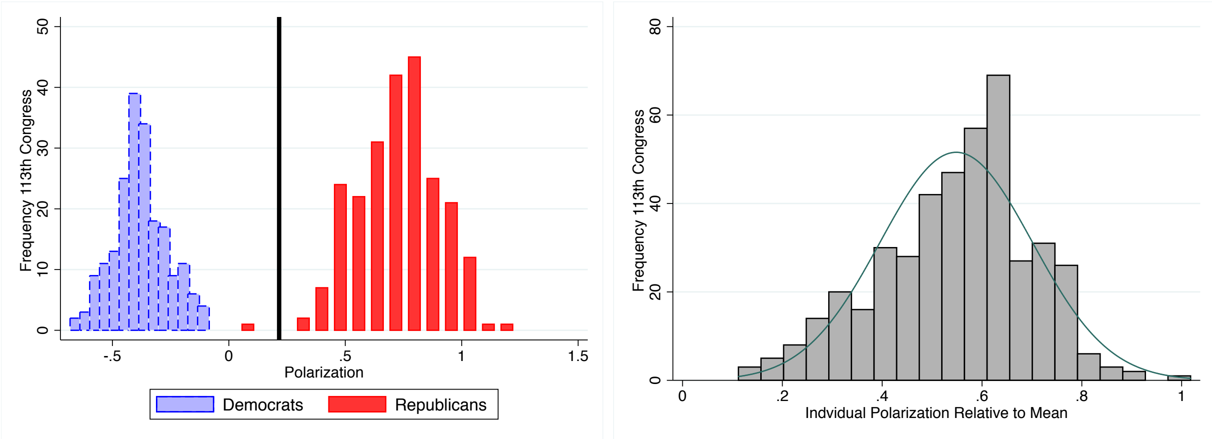
[51] The data is available at http://mcdc.missouri.edu/websas/geocorr12.html.

[52] We exclude the 110th Congress as for this congress, we only have one observation for internet penetration, December 2008, whereas for the 111th-113th Congress we can access several measures over time. The regression results when including the 110th Congress are qualitatively the same and are available upon request.

[53] We run various specifications with additional controls, which are available upon request. Throughout all of our specifications our qualitative outcomes remain unchanged.

Figure 9: Polarization and Deviation of Polarization from Mean, 113th Congress

(a) Polarization (b) Relative Polarization



Note: Figure 9a shows the polarization in the House of Representatives for the 113th Congress. We distinguish between Democrats and Republicans. The measure used is the first dimension of the DW NOMINATE score. The black line represents the mean of polarization, which is positive. In Figure 9b we depict the absolute value of the difference between the individual level of polarization and its mean.

over time. Additionally, we include time dummies to make sure our results are not driven by a time trend. Although, congressional districts remain fixed over time, that is, there is no redistricting over time, the allocation of some counties to congressional district changes.[54] This can have an impact on the voter composition in the congressional districts. In order to control for this, we add variables such as race, what percentage of housing is owner occupied etc. We further control for the party affiliation of the Congressman, that is whether he is a Democrat or Republican.

Results The results of our estimation are given in Table 3. Columns (1) through (3) show the effect of an increase of internet penetration on polarization. We add controls in columns (4)-(6) and focus on these regressions. A one unit increase in households with low speed internet penetration leads to an increase in polarization by .219. Note that the standard deviation of the overall mean of polarization across the 111th to 113th Congress is .165, see Table 2. This implies that an increase by one standard deviation in internet penetration leads to an increase in polarization of .139 standard deviations.[55] The effect of broadband is slightly smaller, but highly significant. An increase in the number of internet providers also has a significantly positive impact on polarization.

So summarize, we find a positive effect of internet penetration on polarization. This effect

[54]Counties can be split among congressional districts. As an example, suppose 70% of a county is included in congressional district one and 30% is in congressional district two for the 111th district. Then it can be the case that 50% are included in district one and 50% in district two for the 112th congress.

[55]This comes about as a unit increase is a switch from zero to one.

Table 3: Internet Penetration and Polarization

	Distance between Individual Polarization and Mean Polarization					
	(1)	(2)	(3)	(4)	(5)	(6)
Internet (200MB) %	0.112			0.219**		
	(0.0976)			(0.0909)		
Broadband %		0.119*			0.182***	
		(0.0677)			(0.0658)	
# Internet Providers			0.00367***			0.00272**
			(0.00131)			(0.00127)
112th Congress	0.0416***	0.0600***	0.0489***	0.0226**	0.0532***	0.0359***
	(0.00963)	(0.00987)	(0.00754)	(0.00879)	(0.00942)	(0.00710)
113th Congress	0.0446***	0.0682***	0.0517***	0.0248**	0.0654***	0.0434***
	(0.0135)	(0.0107)	(0.00921)	(0.0125)	(0.0101)	(0.00878)
Republican				0.0933***	0.0937***	0.0909***
				(0.0197)	(0.0199)	(0.0201)
% White				-0.00554***	-0.00528***	-0.00506**
				(0.00189)	(0.00190)	(0.00198)
Median Age				0.0141***	0.0145***	0.0153***
				(0.00443)	(0.00442)	(0.00441)
Owner Occupied				-0.00657***	-0.00683***	-0.00536**
				(0.00218)	(0.00220)	(0.00229)
Fixed Effects						
N	1304	1304	1304	1304	1304	1304

Standard errors in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Dependent variable is always the absolute value of the difference between mean of polarization and individual level of polarization for each congress. All regressions include fixed effects as well as clustered standard errors. Fixed effects are at congressional district level, standard errors are clustered at congressional district level as well.

is significant at either the one or five percent level. Our result provides evidence that indeed the fragmentation of the media network as captured by internet adoption levels yields higher levels of polarization, in line with our theoretical predictions. Our paper thus provides a novel explanation for why polarization occurs, in line with the empirical evidence. In particular, it does not seem to be the case that new information technologies lead to better information about politicians which could according to our model result in a decrease in polarization.

9 Conclusion

The paper discusses the increase in polarization and establishes the presence of electoral cycles. In order to explain these policy patterns, we develop a framework to analyze targeting in a network with heterogeneous agents and provide a new measure of network centrality that explicitly takes the characteristics of the agents into account. Based on this we show how candidates' advertising strategies and ultimately, the policies they set are influenced by the features of the network. We extend the model to a dynamic setting and show that due to advertising, policies exhibit electoral cycles, that is with each election they fluctuate. We then ask how the features of the media network influence the level of polarization. Changes in the media landscape have made it possible to target messages very narrowly to a certain subset of voters, which generally induces an increase in polarization. Additionally, we show that if communities are connected to a higher number of media outlets then, counterintuitively, in certain environments polarization increases. We then take our predictions to the data. We show that media fragmentation as measured by internet penetration rates leads to higher levels of polarization.

Our model provides a theoretical framework for targeting in networks taking into account not only differences in the network but also allowing for heterogeneity among agents. This comes at the cost of voters being behavioral and not making an inference about the policy if they are not targeted. We show that our qualitative results carry over to a setting with rational voters and uncertainty about the state. Further, in the symmetric, dynamic setting with cycles, voters make no mistakes in aggregate and therefore have no incentive to adjust their adaptive updating strategy. Most importantly, our model outperforms a setting with rational agents as it is able to generate electoral cycles.

References

Abramowitz, A. I. and K. L. Saunders (2008). Is Polarization a Myth? *The Journal of Politics 70*(02), 542–555.

Abramowitz, A. I. and W. J. Stone (2006). The Bush Effect: Polarization, Turnout, and Activism in the 2004 Presidential Election. *Presidential Studies Quarterly 36*(2), 141–154.

Anderson, C. (2006). *The Long Tail: Why the Future of Business is Selling More for Less.* Hyperion.

Banks, J. S. and J. Duggan (2005). Probabilistic Voting in the Spatial Model of Elections: The Theory of Office-Motivated Candidates. In *Social Choice and Strategic Decisions*, pp. 15–56. Springer.

Baron, D. P. (2006). Persistent Media Bias. *Journal of Public Economics 90*(1), 1–36.

Bernhardt, D., J. Duggan, and F. Squintani (2007). Electoral Competition with Privately-Informed Candidates. *Games and Economic Behavior 58*(1), 1–29.

Bernhardt, D., S. Krasa, and M. Polborn (2008). Political Polarization and the Electoral Effects of Media Bias. *Journal of Public Economics 92*(5), 1092–1104.

Bernhardt, M. D. and D. E. Ingerman (1985). Candidate Reputations and the 'Incumbency Effect'. *Journal of Public Economics 27*(1), 47–67.

Besley, T. and S. Coate (1997). An Economic Model of Representative Democracy. *The Quarterly Journal of Economics 112*(1), 85–114.

Bierbrauer, F. J. and P. C. Boyer (2013). Political Competition and Mirrleesian Income Taxation: A First Pass. *Journal of Public Economics 103*, 1–14.

Bimpikis, K., A. Ozdaglar, and E. Yildiz (2013). Competing over networks. Working Paper.

Bisin, A., A. Lizzeri, and L. Yariv (2015). Government Policy with Time Inconsistent Voters. *The American Economic Review 105*(6), 1711–1737.

Budge, I. and R. I. Hofferbert (1990). Mandates and Policy Outputs: US Party Platforms and Federal Expenditures. *American Political Science Review 84*(01), 111–131.

Callander, S. (2005). Electoral Competition in Heterogeneous Districts. *Journal of Political Economy 113*(5), 1116–1145.

Callander, S. (2008). Political Motivations. *The Review of Economic Studies 75*(3), 671–697.

Callander, S. and S. Wilkie (2007). Lies, Damned Lies, and Political Campaigns. *Games and Economic Behavior 60*(2), 262–286.

Calvert, R. L. (1985). Robustness of the Multidimensional Voting Model: Candidate Motivations, Uncertainty, and Convergence. *American Journal of Political Science 29*(1), 69–95.

Campante, F. R. and D. A. Hojman (2013). Media and Polarization: Evidence from the Introduction of Broadcast TV in the United States. *Journal of Public Economics 100*, 79–92.

Caplan, B. (2011). *The Myth of the Rational Voter: Why Democracies Choose Bad Policies.* Princeton University Press.

Castanheira, M. (2003). Why vote for losers? *Journal of the European Economic Association 1*(5), 1207–1238.

Chan, J. and W. Suen (2008). A Spatial Theory of News Consumption and Electoral Competition. *The Review of Economic Studies 75*(3), 699–728.

Coate, S. (2004). Political Competition with Campaign Contributions and Informative Advertising. *Journal of the European Economic Association 2*(5), 772–804.

Cox, G. W. (2009). *Swing Voters, Core Voters, and Distributive Politics*, pp. 342. Cambridge University Press.

DellaVigna, S. and E. D. Kaplan (2007). The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics 122*(3), 1187–1234.

Delli Carpini, M. X. and S. Keeter (1997). *What Americans Know About Politics and Why It Matters.* Yale University Press.

Duggan, J. and C. Martinelli (2011). A Spatial Theory of Media Slant and Voter Choice. *The Review of Economic Studies 78*(2), 640–666.

Dziubiński, M. and S. Goyal (2013). Network Design and Defence. *Games and Economic Behavior 79*, 30–43.

Esponda, I. and D. Pouzo (2016). Conditional Retrospective Voting in Large Elections. *American Economic Journal: Microeconomics*.

Eyster, E. and T. Kittsteiner (2007). Party Platforms in Electoral Competition with Heterogeneous Constituencies. *Theoretical Economics 2*, 41–70.

Fiorina, M. P. (1981). *Retrospective Voting in American National Elections.* Yale University Press.

Fiorina, M. P., S. A. Abrams, and J. C. Pope (2008). Polarization in the American Public: Misconceptions and Misreadings. *The Journal of Politics 70*(02), 556–560.

Fiorina, M. P. and S. J. Abrams (2008). Political Polarization in the American Public. *Annu. Rev. Polit. Sci. 11*, 563–588.

Fiorina, M. P., S. J. Abrams, and J. C. Pope. (2005). *Culture War? The Myth of a Polarized America.* New York: Pearson Longman.

Fiorina, M. P. and M. S. Levendusky (2006). *Red and Blue Nation?: Characteristics and Causes of America's Polarized Politics*, Volume 1, Chapter Disconnected: The Political Class versus the People, pp. 49–71. Brookings Institution Press.

Freedman, P., M. Franz, and K. Goldstein (2004). Campaign Advertising and Democratic Citizenship. *American Journal of Political Science 48*(4), 723–741.

Galeotti, A. and S. Goyal (2009). Influencing the Influencers: a Theory of Strategic Diffusion. *The RAND Journal of Economics 40*(3), 509–532.

Galeotti, A. and A. Mattozzi (2011). "Personal Influence": Social Context and Political Competition. *American Economic Journal: Microeconomics 3*(1), 307–327.

Garcia-Jimeno, C. and P. Yildirim (2015). Matching Pennies on the Campaign Trail: An Empirical Study of Senate Elections and Media Coverage. Working Paper.

Gentzkow, M. (2006). Television and Voter Turnout. *The Quarterly Journal of Economics 121*(3), 931–972.

Gentzkow, M., E. L. Glaeser, and C. Goldin (2006). The Rise of the Fourth Estate. How Newspapers Became Informative and Why it Mattered. In *Corruption and Reform: Lessons from America's Economic History*, pp. 187–230. University of Chicago Press.

Gentzkow, M. and J. M. Shapiro (2006). Media Bias and Reputation. *Journal of Political Economy 114*(2), 280–316.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2011). The Effect of Newspaper Entry and Exit on Electoral Politics. *The American Economic Review 101*(7), 2980.

Gentzkow, M. A. and J. M. Shapiro (2004). Media, Education and Anti-Americanism in the Muslim World. *The Journal of Economic Perspectives 18*(3), 117–133.

Gerber, A. S., D. S. Karlan, and D. Bergan (2009). Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions. *American Economic Journal: Applied Economics 1*(2), 35–52.

Glaeser, E. L., G. A. Ponzetto, and J. M. Shapiro (2005). Strategic Extremism: Why Republicans and Democrats Divide on Religious Values. *The Quarterly Journal of Economics 120*(4), 1283–1330.

Goyal, S., H. Heidari, and M. Kearns (2014). Competitive Contagion in Networks. *Games and Economic Behavior*.

Goyal, S. and A. Vigier (2014). Attack, Defense and Contagion in Networks. *The Review of Economic Studies 81*(4), 1518–1542.

Goyal, S. and A. Vigier (2015). Interaction, Protection and Epidemics. *Journal of Public Economics 125*, 64–69.

Groll, T. and A. Prummer (2015). Whom to Lobby? Targeting in Political Networks. Working Paper.

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. and W. Pesendorfer (2009). Partisan Politics and Election Failure with Ignorant Voters. *Journal of Economic Theory 144*(1), 146–174.

Gul, F. and W. Pesendorfer (2011). Media and Policy. Working Paper.

Harvey, A. and T. Trimbur (2008). Trend Estimation and the Hodrick-Prescott filter. *Journal of the Japan Statistical Society 38*(1), 41–49.

Herrera, H., D. K. Levine, and C. Martinelli (2008). Policy Platforms, Campaign Spending and Voter Participation. *Journal of Public Economics 92*(3), 501–513.

Hoffmann, F., R. Inderst, and M. Ottaviani (2013). Hypertargeting, limited attention, and privacy: Implications for marketing and campaigning. Working Paper.

Hoffmann, F., R. Inderst, and M. Ottaviani (2014). Persuasion through Selective Disclosure: Implications for Marketing, Campaigning, and Privacy Regulation. JW Goethe University Frankfurt, UCL, and Bocconi.

Kolko, J. (2010). A New Measure of US Residential Broadband Availability. *Telecommunications Policy 34*(3), 132–143.

Krasa, S. and M. Polborn (2010). Competition between Specialized Candidates. *American Political Science Review 104*(04), 745–765.

Krasa, S. and M. K. Polborn (2012). Political Competition between Differentiated Candidates. *Games and Economic Behavior 76*(1), 249–271.

Larcinese, V. and L. Miner (2012). The Political Impact of the Internet on US Presidential Elections. Working Paper.

Layman, G. C., T. M. Carsey, and J. M. Horowitz (2006). Party Polarization in American Politics: Characteristics, Causes, and Consequences. *Annu. Rev. Polit. Sci. 9*, 83–110.

Levy, G. and R. Razin (2015). Correlation Neglect, Voting Behavior, and Information Aggregation. *The American Economic Review 105*(4), 1634–1645.

Londregan, J. and T. Romer (1993). *Polarization, Incumbency, and the Personal Vote*, pp. 355–377. Cambridge Univ. Press Cambridge.

Martinelli, C. (2001). Elections with Privately Informed Parties and Voters. *Public Choice 108*(1-2), 147–167.

McCarty, N., K. T. Poole, and H. Rosenthal (2006). *Polarized America: The Dance of Ideology and Unequal Riches*, Volume 5. MIT Press.

Mullainathan, S. and A. Shleifer (2005). The Market for News. *American Economic Review 95*(4), 1031–1053.

Murphy, K. M. and A. Shleifer (2004). Persuasion in Politics. *American Economic Review 94*(2), 435–439.

Osborne, M. J. and A. Slivinski (1996). A Model of Political Competition with Citizen-Candidates. *The Quarterly Journal of Economics 111*(1), 65–96.

Palfrey, T. R. (1984). Spatial Equilibrium with Entry. *The Review of Economic Studies 51*(1), 139–156.

Polborn, M. K. and J. M. Snyder (2016). Party Polarization in Legislatures with Office-Motivated Candidates. Working Paper.

Poole, K. T. and H. Rosenthal (2000). *Congress: A Political-Economic History of Roll Call Voting.* Oxford University Press.

Ravn, M. O. and H. Uhlig (2002). On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations. *Review of Economics and Statistics 84*(2), 371–376.

Roemer, J. E. (1994). A Theory of Policy Differentiation in Single Issue Electoral Politics. *Social Choice and Welfare 11*(4), 355–380.

Schipper, B. C. and H. Y. Woo (2014). Political Awareness, Microtargeting of Voters, and Negative Electoral Campaigning. Working Paper.

Schultz, C. (2007). Strategic Campaigns and Redistributive Politics. *The Economic Journal 117*(522), 936–963.

Soubeyran, R. (2009). Does a Disadvantaged Candidate Choose an Extremist Position? *Annals of Economics and Statistics/Annales d'Économie et de Statistique April/June*(93/94), 327–348.

Spiegler, R. (2013). Placebo Reforms. *The American Economic Review 103*(4), 1490–1506.

Strömberg, D. (2004). Radio's Impact on Public Spending. *The Quarterly Journal of Economics 119*(1), 189–221.

Wittman, D. (1983). Candidate Motivation: A synthesis of Alternative Theories. *American Political Science Review 77*(01), 142–157.

Appendix A

Proof Proposition 1: We first establish that for any target set, both parties disclose the average bliss point in the target's coverage. Fix T_c the target set of politician c, and recall that his policy decision needs to solve

$$\max_{x_c} \frac{1}{2}\left[\int_0^1 u(x_c|x)\textstyle\sum_{i\in K(T_c)} g_i(x)dx + \int_0^1 u(\pi_c|x)\textstyle\sum_{i\in K\setminus K(T_c)} g_i(x)dx\right]$$

Our functional form assumption and requirement on voting behaviour implicit in the model imply that the problem is equivalent to solving

$$\max_{x_c} \frac{1}{2}\int_0^1 (1-(x_c-x)^2)\textstyle\sum_{i\in K(T_c)} g_i(x)dx$$

Necessary conditions for an interior optimum then require that

$$\int_0^1 (x_c - x)\textstyle\sum_{i\in K(T_c)} g_i(x)dx = 0$$

By our functional form assumptions, the condition is also sufficient for an optimum, and fully characterizes the equilibrium policy for any given target decision. Simple manipulations then establish that

$$\begin{aligned} x_c &= \left[\int_0^1 x\textstyle\sum_{i\in K(T_c)} g_i(x)dx\right] / \left[\int_0^1 \textstyle\sum_{i\in K(T_c)} g_i(x)dx\right] = \\ &= \frac{1}{k(T_c)}\textstyle\sum_{i\in K(T_c)} \int_0^1 x g_i(x)dx = E(X|T_c) \end{aligned}$$

which establishes the first part of the proposition.

Begin by calculating the payoff of any possible targeting strategy. Consider a target set T, and observe that, at the optimal policy, the preference of candidate c are a positive affine transformation of

$$\int_0^1 u(E(X|T)|x)\textstyle\sum_{i\in K(T)} g_i(x)dx + \int_0^1 u(\pi_c|x)\textstyle\sum_{i\in K\setminus K(T)} g_i(x)dx =$$
$$\int_0^1 \left[1-(E(X|T)-x)^2\right]\textstyle\sum_{i\in K(T)} g_i(x)dx + \int_0^1 \left[1-(\pi_c - x)^2\right]\textstyle\sum_{i\in K\setminus K(T)} g_i(x)dx$$

If so target set T is better than target set S if and only if

$$\int_0^1 (E(X|T)-x)^2 \textstyle\sum_{i\in K(T)} g_i(x)dx + \int_0^1 (\pi_c - x)^2 \textstyle\sum_{i\in K\setminus K(T)} g_i(x)dx <$$
$$\int_0^1 (E(X|S)-x)^2 \textstyle\sum_{i\in K(S)} g_i(x)dx + \int_0^1 (\pi_c - x)^2 \textstyle\sum_{i\in K\setminus K(S)} g_i(x)dx$$

However, the definition of variance and simple manipulations establish that

$$\begin{aligned}\int_0^1 (E(X|T)-x)^2 \textstyle\sum_{i\in K(T)} g_i(x)dx &= k(T)\int_0^1 (E(X|T)-x)^2 dG(x|T)\\ &= k(T)\left(E(X^2|T)-E(X|T)^2\right)\end{aligned}$$

Similarly, it also follows that

$$\begin{aligned}\int_0^1 (\pi_c - x)^2 \textstyle\sum_{i\in K\setminus K(T)} g_i(x)dx &= \int_0^1 (\pi_c^2 - 2\pi_c x + x^2) \textstyle\sum_{i\in K\setminus K(T)} g_i(x)dx\\ &= k(M-T)\left(\pi_c^2 - 2\pi_c E(X|M-T) + E(X^2|M-T)\right)\end{aligned}$$

Using the three previous observations then implies that

$$\begin{aligned}&k(T)\left(E(X^2|T)-E(X|T)^2\right) + k(M-T)\left(\pi_c^2 - 2\pi_c E(X|M-T) + E(X^2|M-T)\right)\\ &< k(S)\left(E(X^2|S)-E(X|S)^2\right) + k(M-S)\left(\pi_c^2 - 2\pi_c E(X|M-S) + E(X^2|M-S)\right) \quad (12)\end{aligned}$$

Now, observe that for any T, we have that

$$k(T)E(X^2|T) + k(M-T)E(X^2|M-T) = kE(X^2|M),$$

as we are averaging over all communities in K. Further,

$$\begin{aligned}k(M-S)E(X|M-S) - k(M-T)E(X|M-T) &= \textstyle\sum_{i\in K\setminus K(S)} E_i(X) - \sum_{i\in K\setminus K(T)} E_i(X)\\ &= \textstyle\sum_{i\in K(T)} E_i(X) - \sum_{i\in K(S)} E_i(X)\\ &= k(T)E(X|T) - k(S)E(X|S),\end{aligned}$$

where the second equality folds by adding and subtracting $\sum_{i\in K} E_i(X)$. Based on the last two simplifications, inequality (12) becomes

$$k(M-T)\left(\pi_c^2-2\pi_c E(X|M-T)\right)-k(T)E(X|T)^2<k(M-S)\left(\pi_c^2-2\pi_c E(X|M-S)\right)-k(S)E(X|S)^2$$
$$k(S)\left(E(X|S)^2+\pi_c^2\right)-k(T)\left(E(X|T)^2+\pi_c^2\right)<2\pi_c\left(k(M-T)E(X|M-T)-k(M-S)E(X|M-S)\right)$$
$$k(S)\left(E(X|S)^2+\pi_c^2\right)-k(T)\left(E(X|T)^2+\pi_c^2\right)<-2\pi_c\left(k(T)E(X|T)-k(S)E(X|S)\right)$$
$$k(S)\left(E(X|S)-\pi_c\right)^2<k(T)\left(E(X|T)-\pi_c\right)^2$$

As the same logic applies to every target set it follows that, it is optimal to choose that target set that maximizes $k(T)\left(E(X|T)-\pi_c\right)^2$. We can use the concept of media centrality to compare target sets. The following lemma will be useful for this.

Lemma 9. *Consider any two target sets $T, T' \subseteq M$ such that $T \subseteq T'$. Then $W_c(T) < W_c(T')$ if and only if*

$$k(T)\left[E(X|T'-T)-E(X|T)\right]^2<k(T')\left[E(X|T'-T)-\pi_c\right]^2. \tag{13}$$

Proof Lemma 9: Observe first that, since $T \subseteq T'$,

$$W(T')=k(T')\left[E(X|T')-\pi_c\right]^2=k(T')\left[\frac{k(T)E(X|T)+k(T'-T)E(X|T'-T)}{k(T')}-\pi_c\right]^2.$$

If so, $W(T) < W(T')$ is equivalent to

$$k(T)\left[E(X|T)-\pi_c\right]^2<k(T')\left[\frac{k(T)E(X|T)+k(T'-T)E(X|T'-T)}{k(T')}-\pi_c\right]^2.$$

Simple manipulations then show that this is equivalent to

$$k(T)\left[E(X|T'-T)-E(X|T)\right]^2<k(T')\left[E(X|T'-T)-\pi_c\right]^2.$$

Proof Lemma 1: We assume without loss of generality that $E(X|T') < \pi_c$. A symmetric logic applies to the converse scenario. Observe that $E(X|T'-T) \geq \pi_c$ implies that $E(X|T) < \pi_c$, as

$$E(X|T')=\frac{k(T)E(X|T)+k(T'-T)E(X|T'-T)}{k(T)+k(T'-T)}<\pi_c.$$

Moreover, this implies that

$$\pi_c - E(X|T) > \pi_c - E(X|T') > 0. \tag{14}$$

For a candidate **not** to be willing to add outlets $T'\backslash T$ to the target set it must be that

$$k(T')\left[\pi_c - E(X|T')\right]^2 < k(T)\left[\pi_c - E(X|T)\right]^2.$$

Moreover by (14) the latter condition is equivalent to

$$k(T')^{1/2}\left[\pi_c - E(X|T')\right] < k(T)^{1/2}\left[\pi_c - E(X|T)\right].$$

Writing explicitly the expected target set yields the following

$$k(T')^{-1/2}\textstyle\sum_{i\in K(T')}\left[\pi_c - E_i(X)\right] < k(T)^{-1/2}\textstyle\sum_{i\in K(T)}\left[\pi_c - E_i(X)\right].$$

Thus the politician does not target outlets $T'\backslash T$ if

$$\textstyle\sum_{i\in K(T')}\left[\pi_c - E_i(X)\right] < \left(\dfrac{k(T')}{k(T)}\right)^{1/2}\textstyle\sum_{i\in K(T)}\left[\pi_c - E_i(X)\right]. \tag{15}$$

But, as $\pi_c - E(X|T'-T) \leq 0$ is equivalent to $\sum_{i\in K(T')\backslash K(T)}\left[\pi_c - E_i(X)\right] \leq 0$, we have that

$$\textstyle\sum_{i\in K(T')}\left[\pi_c - E_i(X)\right] \leq \textstyle\sum_{i\in K(T)}\left[\pi_c - E_i(X)\right].$$

If so, condition (15) must hold as well since $k(T') > k(T)$.

Proof Lemma 2: (1) First consider the case in which $E(X|T'-T) \leq E(X|T)$. If so,

$$E(X|T') = \frac{k(T)E(X|T) + k(T'-T)E(X|T'-T)}{k(T')} \leq E(X|T),$$

and therefore $W_c(T) \leq W_c(T')$, as it is possible to both increase the number of the communities targeted – as $k(T') \geq k(T)$ – without decreasing the difference between the expected bliss point and the prior – as $\pi_c - E(X|T') \geq \pi_c - E(X|T)$.

Next consider the case in which $E(X|T'-T) > E(X|T)$. It is better to disclose also in outlet j if

$$k(T')\left[\frac{k(T)E(X|T)+k(T'-T)E(X|T'-T)}{k(T')}-\pi_c\right]^2 \geq k(T)\left[E(X|T)-\pi_c\right]^2.$$

this inequality by Lemma 9 and some algebra is equivalent to

$$k(T'-T)\left(E(X|T'-T)-\pi\right)^2 \geq k(T)\left(E(X|T)-\pi_c\right)\left(E(X|T)+\pi_c-2E(X|T'-T)\right). \quad (16)$$

However by assumption we have that $E(X|T)+\pi_c > 2E(X|T'-T)$ and $E(X|T) < \pi_c$. Thus the RHS of (16) is negative, whereas the LHS of (16) is positive; and thus inequality of (16) must hold.

(2) By part (1), including outlets $T'\backslash T$ to the target set increases the candidate's payoff if $E(X|T'-T) < E(X|T) < \pi_c$. Thus, suppose that $E(X|T'-T) > E(X|T)$. By assumption T has a higher media centrality than T'' and thus by Lemma 9 we have that

$$k(T)\left[\pi_c - E(X|T-T'')\right]^2 > k(T'')\left[E(X|T'')-E(X|T-T'')\right]^2.$$

We want to establish that

$$k(T')\left[\pi_c - E(X|T'-T)\right]^2 > k(T)\left[E(X|T)-E(X|T'-T)\right]^2. \quad (17)$$

Observe that, as $E(X|T-T'') \in (E(X|T'-T),\pi_c)$, we have that

$$\left[\pi_c - E(X|T'-T)\right]^2 > \left[\pi_c - E(X|T-T'')\right]^2,$$

Similarly, as $E(X|T'-T) > E(X|T)$, we have that

$$\left[E(X|T)-E(X|T-T'')\right]^2 > \left[E(X|T)-E(X|T'-T)\right]^2.$$

But, from the previous inequalities we have that

$$\begin{aligned}
k(T')\left[\pi_c - E(X|T'-T)\right]^2 &> k(T')\left[\pi_c - E(X|T-T'')\right]^2 \\
&> \frac{k(T')k(T'')}{k(T)}\left[E(X|T'') - E(X|T-T'')\right]^2 \\
&= \frac{k(T')k(T)}{k(T'')}\left[E(X|T) - E(X|T-T'')\right]^2 \\
&> k(T)\left[E(X|T) - E(X|T-T'')\right]^2 \\
&> k(T)\left[E(X|T) - E(X|T'-T)\right]^2.
\end{aligned}$$

where the equality follows as $k(T'')E(X|T'') = k(T)E(X|T) - k(T-T'')E(X|T-T'')$. This completes the proof by establishing (17).

We add an additional lemma, which turns out helpful to see what target sets are optimal if priors change. The proof is immediate and therefore omitted.

Lemma 10. *Let one of the following conditions hold.*

C.1 $0 < E(X|T) < E(X|T') < \pi' < \pi < 1$

C.2 $0 < E(X|T') < E(X|T) < \pi < \pi' < 1$

C.3 $0 < E(X|T) < \pi < \pi' < E(X|T') < 1$

Then, if T is preferred to T' given prior π, it is also preferred given prior π'.

Proof of Lemma 3 A partition of the communities of outlet j can either lead to (i) no change at all (ii) an outlet dropped from the target set.

Suppose first that the communities in the target set remain the same. Then, $k'(T') = k(T)$, $E'(X|T') = E(X|T)$, which implies $W'(T') = W(T)$. By Lemma 9, it then follows that $o \notin T$ implies $o \notin T'$ as $k'(o-T') = k(o-T)$ and $E'(X|o-T') = E(X|o-T)$.

If an outlet is dropped from the target set such that fewer communities remain in the target set, $K'(T') \subset K(T)$, then by the definition of media centrality, it must hold that $E'(X|T') < E(X|T)$ as $k'(T') < k(T)$. We then show that it cannot be optimal for an outlet $o \notin T$ to be added, that is it must hold that $o \notin T'$. We denote the outlets that are dropped by $j_\imath$. Given that is has not been optimal to add outlet o to target set T it must hold that

$$k(T)\left(E(X|o-T) - E(X|T)\right)^2 > \left(k(T) + k(o-T)\right)\left(E(X|o-T) - \pi\right)^2 \tag{18}$$

It is not optimal to add outlet o to target set T', if

$$k'(T')\left(E'(X|o-T')-E'(X|T')\right)^2 > \left(k'(T')+k'(o-T')\right)\left(E'(X|o-T')-\pi\right)^2 \tag{19}$$

Note first that given $K'(j_\imath)\cap K'(o)=\emptyset$, $E(X|o-T)=E'(X|o-T')$ and $k(o-T)=k'(o-T')$. This implies that RHS of inequality (18) is larger than the right hand side of inequality (19). Additionally, it has to hold that

$$k'(T')\left(E'(T')-\pi\right)^2 > k(T)\left(E(T)-\pi\right)^2 \tag{20}$$

If $E'(X|o-T')>\pi$, then by Lemma 1, $o\notin T'$. We therefore focus on $E'(X|o-T')<\pi$. By Lemma 2 for $o\notin T$, it must be that $E(X|T)<E(X|o-T)$. This implies that

$$E(X|T') < E(X|T) < E(X|o-T) < \pi$$

Lemma 10 **C.1** together with inequality 20 implies that the LHS of inequality (19) is greater than the LHS of inequality (18) which establishes that inequality (19) is fulfilled.

Proof Lemma 4: We first consider case 1, that is a link if formed within the target set. Recall that $j\in T$ denotes an outlet such that $i\in K(j)$ – which must exist by assumption. We first show that for $j\in T'$, it must be that $T=T'$.

Lemma 11. *If $j\in T\cap T'$, then $T=T'$.*

Proof: If $j\in T'$, it must be that $W'(T')=W(T')$. Assume by contradiction that $T'\neq T$. If so, it must be that $W(T')<W(T)$. But as $j\in T$ it must be that $W(T)=W'(T)$. But this implies that $W'(T')<W'(T)$, which contradicts the optimality of T'.

Next we show that if $j\in T\backslash T'$, it there does not exists an outlet $o\notin T$, but $o\in T'$.

Lemma 12. *Let $j\in T\backslash T'$ and $l\in T$ then $T'\subseteq T$.*

Proof: To see this consider three outlets, $\{j,l,o\}$. Assume for now that their coverages do not overlap and that $j=T\backslash T'$. We then proceed by contradiction, that is it is not optimal to include j in the new target set, but it is optimal to include o.

Denote by $k(l)$ the number of groups in outlet l, by $k'(j-l)$ the number of groups in outlet j that only belong to outlet j and by $k(o)$ the number of groups in outlet o. The bliss points are given along the same lines that is $E(X|l)$ is the expected bliss point in outlet l,

$E(X_i)$ the bliss point of group i that forms a new connection, $E'(X|j-l)$ is the bliss point of all other groups in outlet j, $E(X|o)$ the bliss point in outlet o. If $T = \{l, j\}$, but $o \in T'$, then it has to be the case that

$$(k(l)+1)\left(E'(X|j-l) - \frac{k(l)E(X|l)+E(X_i)}{k(l)+1}\right)^2 \\ > (k(l)+k'(j-l)+1)\left(E'(X|j-l)-\pi\right)^2 \tag{21}$$

$$(k(l)+1+k'(j-l))\left(E(X|o) - \frac{k(l)E(X|l)+E(X_i)+k'(j-l)E'(X|j-l)}{k(l)+1+k'(j-l)}\right)^2 \\ > (k(l)+1+k'(j-l)+k(o))\left(E(X|o)-\pi\right)^2 \tag{22}$$

$$(k(l)+1+k(o))\left(E(X|o)-\pi\right)^2 \\ > (k(l)+1)\left(E(X|o) - \frac{k(l)E(X|l)+E(X_i)}{k(l)+1}\right)^2 \tag{23}$$

The first equation requires $W(l,j) > W(l)$; for the second $W'(l,j) < W'(l)$; the third requires $W(l,j,o) < W(l,j)$; and for the last $W'(l,o) > W'(l)$. The last condition holds as $T' \neq \{l,j,o\}$ since $W'(l,j,o) = W(l,j,o) < W(l,j) = W'(l,j)$. Note that

$$\begin{aligned} k'(l) &= 1 + k(l), \\ k(l+j) &= k(l) + 1 + k'(j-l), \\ k'(l+o) &= k(l) + 1 + k(o), \\ K &= k(l) + 1 + k'(j-l) + k(o). \end{aligned}$$

Equations (21) can be rewritten as

$$k'(j-l)E'(X|j-l) > k(l+j)\pi - k(l)E(X|l) - E(X_i) \\ - \sqrt{\frac{k(l+j)}{k'(l)}}\left((1+k(l))\pi - k(l)E(X|l) - E(X_i)\right). \tag{24}$$

Similarly, equations (22) and (23) become

$$E(X|o) > \frac{1}{k(o)}\Bigg(K\pi - k(l)E(X|l) - E(X_i) - k'(j-l)E'(X|j-l) - \sqrt{\frac{K}{k(l+j)}}\left(k(l+j)\pi - k(l)E(X|l) - E(X_i) - k'(j-l)E'(X|j-l)\right)\Bigg).$$

$$E(X|o) < \frac{1}{k(o)}\left(k'(l+o)\pi - k(l)E(X|l) - E(X_i) - \sqrt{\frac{k'(l+o)}{k'(l)}}(k'(l)\pi - k(l)E(X|l) - E(X_i))\right).$$

For the last two equations to hold, it must be that

$$\left(k'(l+o)\pi - \sqrt{\frac{k'(l+o)}{k'(l)}}(k'(l)\pi - k(l)E(X|l) - E(X_i))\right) >$$
$$\left(K\pi - k'(j-l)E'(X|j-l) - \sqrt{\frac{K}{k(l+j)}}\left(k(l+j)\pi - k(l)E(X|l) - E(X_i) - k'(j-l)E'(X|j-l)\right)\right)$$

which is equivalent to

$$\frac{1}{\sqrt{\frac{K}{k(l+j)}}-1}\Bigg(\left(k'(l+o) - K + \sqrt{k(l+j)K} - \sqrt{k'(l)k'(l+o)}\right)\pi + \left(\sqrt{\frac{k'(l+o)}{k'(l)}} - \sqrt{\frac{K}{k(l+j)}}\right)(k(l)E(X|l) + E(X_i))\Bigg) > k'(j-l)E'(X|j-l)$$

As we also have a lower bound on $k'(j-l)E'(X|j-l)$ given by equation (24) is must be that

$$\frac{1}{\sqrt{\frac{K}{k(l+j)}}-1}\Bigg(\left(k'(l+o) - K + \sqrt{k(l+j)K} - \sqrt{k'(l)k'(l+o)}\right)\pi \tag{25}$$
$$+ \left(\sqrt{\frac{k'(l+o)}{k'(l)}} - \sqrt{\frac{K}{k(l+j)}}\right)(k(l)E(X|l) + E(X_i))\Bigg) \tag{26}$$
$$> \left(k(l+j)\pi - k(l)E(X|l) - E(X_i) - \sqrt{\frac{k(l+j)}{k'(l)}}(k'(l)\pi - k(l)E(X|l) - E(X_i))\right),$$

which never holds and which gives the contradiction we were looking for. The first part of the argument relies on two assumptions, namely: (1) coverages have no overlap, (2) $j = T\backslash T'$. The proof without these two assumptions follows along similar lines and can be found in Appendix D. This establishes that if the target set remains on the same side of the prior and an outlet is omitted, that it then cannot be optimal for an outlet that was originally not included in the target set to be included in the new target set.

If a community establishes an additional link to a an outlet contained in the target set, then $T' \subseteq T$. It is clear that if the target sets are identical with $T' = T$, $E(X|T) = E'(X|T')$. Consider next the case that $T' \subset T$. Then $k(T) \geq k'(T')$. Given optimality of T', it must be that $W(T) = W'(T) < W'(T')$, where $W(T) = k(T)\,(E(X|T) - \pi)^2$ and $W'(T') = k'(T')\,(E'(X|T') - \pi)^2$. For $W(T) < W'(T)$ and $k(T) \geq k'(T')$ it then must be the case that $(E(X|T) - \pi)^2 < (E'(X|T') - \pi)^2$. As we have assumed that the bliss points of the target sets do not switch $E(X|T), E'(X|T') < \pi$, $\pi - E'(X|T') > \pi - E(X|T)$, which is equivalent to $E(X|T) > E'(X|T')$.

We then turn to the case where a group that belongs to a targeted outlet forms a link to an outlet that does not belong to the target set. There we establish that next lemma.

Lemma 13. *Let $l \notin T$ and $i \in K(T)$. If $j \in T'$, $T' = T$. If $j \notin T'$, $l \in T'$.*

Proof: We have already established in the proof to Lemma 11, that if $j \in T'$, it must be that $T = T'$. If $j \notin T'$, then $l \in T'$. Suppose by contradiction that $j \notin T'$ and $l \notin T'$. It must be the case that $W(T \backslash j) < W(T)$, otherwise it would not have been optimal to add outlet j. Note that $W'(T \backslash j) = W(T \backslash j)$ and $W'(T) = W(T)$ and so $W'(T \backslash j) < W'(T)$. This establishes that it cannot be optimal to only omit outlet j from the target set, without making any other changes to it. There are two possibilities given our assumption that $l \notin T'$. Either another outlet $o \neq l$, $o \in T$ is omitted from the target set or an outlet $o \notin T$ is added to the target set. Consider first the case where an outlet is omitted from the target set. The media centrality of such a target set is $W(T \backslash (j \cup o)) = W'(T \backslash (j \cup o))$, as the change in the linking structure has not affected outlet o. It must be the case that $W(T \backslash (j \cup o)) < W(T)$. But this implies that $W'(T') = W'(T \backslash (j \cup o)) < W'(T)$, contradicting the optimality of T'. Consider next the case where $o \notin T$ is added to the target set. The media centrality of such a target set is $W((T \cup o) \backslash j) = W'((T \cup o) \backslash j)$, as again the added link does not affect o. It holds that $W((T \cup o) \backslash j) < W(T)$ and thus $W'((T \cup o) \backslash j) = W'(T') < W'(T)$, contradicting the optimality of T'. Therefore, it must be the case that if $j \notin T'$, then $l \in T'$.

Proof Proposition 2: (2) Obviously, the probability of either of the two candidates winning the election must necessarily equal 1 – that is, $P_A(x_A, x_B|x) + P_B(x_A, x_B|x) = 1$. However, if candidate c enjoys a probability of winning that strictly exceeds 1/2, then the other candidate must be losing with a probability that exceeds 1/2. But if so, the strategy of the latter candidate cannot be optimal as, by mimicking candidate c, he could increase his odds of winning to $1/2$, when candidates are homogeneous. Thus, if $\pi_A = \pi_B$, both candidates must win with probability $1/2$ in any equilibrium. Moreover, (1) must hold as candidates' best responses are independent of the competitor's strategy and identical.

(3) By Proposition 1, both candidates set the same policy when targeting the same media outlets. Thus, to prove the last part of claim it suffices to show that candidates generically select the same media coverage. Clearly, symmetric equilibria always exist as either candidate chooses the optimal target by solving

$$\max_{T \subseteq M} k(T) \left[E(X|T) - \pi\right]^2 . \tag{27}$$

If (27) the admits a single solution, any equilibrium is necessarily symmetric. To conclude that this is generically the case, denote by E_i the average bliss point $E_i(X)$ in every community $i \in K$. Suppose that every E_i is drawn from a continuous distribution with no atoms and with support $[l, h] \subseteq [0, 1]$.[56] If so, for any $T, S \subseteq M$ such that $K(T) \neq K(S)$, we have that

$$\Pr(E(X|T) = E(X|S)) = \Pr\left(k(S) \textstyle\sum_{i \in K(T)} E_i = k(T) \textstyle\sum_{i \in K(S)} E_i\right) = 0.$$

The second equality holds as $k(S) \sum_{i \in K(T)} E_i = k(T) \sum_{i \in K(S)} E_i$ defines a hyperplane in $[l, h]^k$ when $K(T) \neq K(S)$, and because any such hyperplane has measure zero in $[l, h]^k$ when all the E_i are drawn independently from a common distribution with no atoms.

Proof of Proposition 3 The policy platform of candidate c is given by

$$x_c = \frac{k(T_c) E(X|T_c) + D E_c(X)}{k(T_c) + D}$$

We want to ensure that

$$\frac{k(T_{max}) E(X|T_{max}) + D E_A(X)}{k(T_{max}) + D} < \frac{k(T_{min}) E(X|T_{min}) + D E_B(X)}{k(T_{min}) + D}$$

[56] Implicitly, we are assuming that G_i is determined once $E_i(X)$ has been set so to satisfy

$$E_i(X) = \int_0^1 x dG_i(x).$$

This can be rewritten as

$$k(T_{max})k(T_{min})\left(E(X|T_{max})-E(X|T_{min})\right)+D\left(k(T_{max})E(X|T_{max})-k(T_{min})E(X|T_{min})\right)$$
$$<D^2\left(E_B(X)-E_A(X)\right)+D\left(k(T_{max})E_B(X)-k(T_{min})E_A(X)\right)$$

It is always the case that

$$D\left(k(T_{max})E(X|T_{max})-k(T_{min})E(X|T_{min})\right)\leq D\left(k(T_{max})E_B(X)-k(T_{min})E_A(X)\right)$$

as $E_A(X)\leq E_i(X)\leq E_B(X)$. Therefore, if

$$k(T_{max})k(T_{min})\left(E(X|T_{max})-E(X|T_{min})\right)\leq D^2\left(E_B(X)-E_A(X)\right)$$

it must be that $x_A\leq x_B$.

Proof of Proposition 4 Let there be two outlets $j,j'\in M$ with $k(j-j')>0$. Consider an arbitrary target set T^t. Then the media centrality in $t+1$ is given by

$$(k(T^{t+1})+D)\left(\frac{k(T^{t+1})E(X|T^{t+1})+DE_c(X)}{k(T^{t+1})+D}-\frac{k(T^t)E(X|T^t)+DE_c(X)}{k(T^t)+D}\right)^2$$

By setting $T^{t+1}=T^t$, media centrality is zero. Choosing any other target set leads to an increase in media centrality.

Proof of Lemma 5 (1)

We aim to show that *in the long run* a politician's platform $x_c^t\in\{E(X|T_L),E(X|T_R)\}$ for $t>\bar{t}$, with $E(X|T_L)<E(X|T_R)$ and switches between these two policies that is for some t, $x_c^t=E(X|T_L)$, $x_c^{t+1}=E(X|T_R)$ and $x_c^{t+2}=E(X|T_L)$.

Note first that for a given π, there exists a generically unique target set T that maximizes $W_c(T)$. To see this denote by $\mathcal{P}(M)$ the power set of outlets M. The power set contains any possible target set. Every set has a maximum. This maximum is generically unique, by the same logic as used in the proof of Proposition 2. Denote by T^t the target set in period t. Such a target set leads to policy $E(X|T^t)$.

Suppose first that $E(X|T^1)<E(X|T^2)<E(X|T^3)$. We are then interested in where $E(X|T^4)$ lies.

We first establish that

$$E(X|T^1) < E(X|T^4) < E(X|T^3) \tag{28}$$

To see this note that it must have been optimal to select $E(X|T^2)$ over $E(X|T^3)$ as the platform given prior $E(X|T^1)$, that is

$$k(T^2)\left(E(X|T^2) - E(X|T^1)\right)^2 > k(T^3)\left(E(X|T^3) - E(X|T^1)\right)^2$$

Further, it must be the case that it is better to select platform $E(X|T^3)$ over $E(X|T^4)$ given prior $E(X|T^2)$,

$$k(T^3)\left(E(X|T^3) - E(X|T^2)\right)^2 > k(T^4)\left(E(X|T^4) - E(X|T^2)\right)^2$$

Last, it must be the case that $E(X|T^4)$ is preferred to $E(X|T^2)$ for prior $E(X|T^3)$,

$$k(T^4)\left(E(X|T^4) - E(X|T^3)\right)^2 > k(T^2)\left(E(X|T^2) - E(X|T^3)\right)^2$$

Taking these three equations together, it is straightforward to show that they hold if and only if $E(X|T^1) < E(X|T^4) < E(X|T^3)$. If $E(X|T^4) = E(X|T^2)$, that is the target sets and policies in period 2 and 4 are identical, then we have established that there is indeed cycling between two alternatives.

If, on the other hand, $E(X|T^4) \neq E(X|T^2)$, it has to be the case that

$$k(T^2)\left(E(X|T^2) - E(X|T^1)\right)^2 > k(T^4)\left(E(X|T^4) - E(X|T^1)\right)^2.$$

It has to be better to choose $E(X|T^2)$ rather than $E(X|T^4)$ given prior $E(X|T^1)$. This can only hold for

$$E(X|T^1) < E(X|T^4) < E(X|T^2). \tag{29}$$

We then have a sequence of policy platforms, where $E(X|T^1) < E(X|T^4) < E(X|T^2) < E(X|T^3)$.

Based on this we turn to $E(X|T^5)$. If $E(X|T^5) = E(X|T^3)$, then we have again established that there is cycling between two alternatives. If $E(X|T^5) \neq E(X|T^3)$, we again need to characterize where $E(X|T^5)$ can lie. The details of the proof can be found in Appendix E. We can show that eventually for any target set, $E(X|T^{t+1}) < E(X|T^t)$. As the policy space is bounded, this implies the at some point $E(X|T^t) = E(X|T^{t+2})$.

Proof of Lemma 5: (2) We aim to show that $E(X|T_L) < E(X|T'_L) < E(X|T_R) < E(X|T'_R)$. First, by Lemma 10 it cannot be the case that

$$E(X|T'_L) < E(X|T_L) < E(X|T_R) < E(X|T'_R)$$

Further, it cannot be the case that

$$E(X|T_L) < E(X|T_R) < E(X|T'_L) < E(X|T'_R)$$

For these cycles to be optimal, the following equations have to hold

$$\begin{aligned} k(T_R)\left(E(X|T_R) - E(X|T_L)\right)^2 &> k(T'_L)\left(E(X|T'_L) - E(X|T_L)\right)^2, \\ k(T'_L)\left(E(X|T'_L) - E(X|T'_R)\right)^2 &> k(T_R)\left(E(X|T_R) - E(X|T'_R)\right)^2, \end{aligned}$$

which cannot occur.

Proof of Lemma 6 For $D = 0$, there always exists an identical policy cycle, if the initial prior is the same. We therefore focus on $D > 0$. For a given target set T_i there always exists a target set T_{-i} by symmetry, with $E(X|T_i) = 1 - E(X|T_{-i})$ and $k(T_i) = k(T_{-i})$. Then, for any given target set the following holds

$$\frac{k(T_i)E(X|T_i) + DE_B(X)}{k(T_i) + D} = 1 - \frac{k(T_{-i})E(X|T_{-i}) + DE_A(X)}{k(T_{-i}) + D} \tag{30}$$

This can be simplified as follows

$$\begin{aligned} \frac{k(T_i)E(X|T_i) + k(T_i)(1 - E(X|T_i)) + DE_B(X) + DE_A(X)}{k(T_i) + D} &= 1 \\ \frac{k(T_i) + D}{k(T_i) + D} &= 1 \end{aligned}$$

and establishes that equation (30) holds. This implies that for each policy that is available for candidate A there exists a symmetric policy for candidate B. Suppose that T_{AL} and T_{AR} are the target sets for candidate A. Then, there exists target sets T_{BL} and T_{BR} such that

$x_{AL} = 1 - x_{BR}$ and $x_{AR} = 1 - x_{BL}$. Additionally, if T_{AL} and T_{AR} are optimal, then it also has to be the case that T_{BL} and T_{BR} are optimal due to the symmetry. This completes the proof.

Proof of Lemma 7 We take derivatives of equation (8) with respect to $\sum_{i \in K(T_L)} E_i(X)$, $\sum_{i \in K(T_R)} E_i(X)$ and $k(T_L)$ and $k(T_R)$.

1. Change in $\sum_{i \in K(T_L)} E_i(X)$

$$\frac{\partial \Delta}{\partial \sum_{i \in K(T_L)} E_i(X)} = \frac{2}{k(T_L) + D} > 0$$

2. Change in $\sum_{i \in K(T_R)} E_i(X)$

$$\frac{\partial \Delta}{\partial \sum_{i \in K(T_R)} E_i(X)} = \frac{2}{k(T_R) + D} > 0$$

3. Change in $k(T_L)$

$$\frac{\partial \Delta}{\partial k(T_L)} = \frac{-D - 2\left(\sum_{i \in K(T_L)} E_i(X)\right) - D\left(E_B(X) - E_A(X)\right)}{\left(k(T_L) + D\right)^2} < 0$$

4. Change in $k(T_R)$

$$\frac{\partial \Delta}{\partial k(T_R)} = \frac{-D - 2\left(\sum_{i \in K(T_R)} E_i(X)\right) - D\left(E_B(X) - E_A(X)\right)}{\left(k(T_R) + D\right)^2} < 0$$

Proof of Lemma 8 We want to show that $k(T_L) \geq k(T_R)$. If both candidates select the same target sets, it is clear that the target sets are of equal size. We therefore focus on the case where candidates select different target sets that yield symmetric policies. Due to symmetry, it must hold that the number of communities to the left of 1/2 equals the number of communities to the right of 1/2. As $D > \overline{D}$, we know that for any possible target set T_L, $x_{BL} > \frac{1}{2}$. To see this note that for $x_{BL} < \frac{1}{2}$, by symmetry it must hold that $\frac{1}{2} < x_{AR}$. But then $x_{AR} > x_{BL}$, which yields a contradiction as $D > \overline{D}$. Suppose now by contradiction that the number of outlets in T_R is higher than the number of outlets in T_L. This can only be the case if for some outlet j_+, the symmetric outlet j_- is also contained in T_R. Otherwise, the number of communities in each target set would be half of all communities. By symmetry, it must be that $E(X|j_- \cup j_+) = \frac{1}{2}$. Lemma 1 establishes that it cannot be optimal to include outlets with $E(X|j_- \cup j_+) < x_{BL}$, which yields a contradiction and completes the proof.

Proof of Proposition 5 It can always be the case that the target sets are not affected by the partition in which case polarization remains unchanged. In what follows we focus on what happens if target sets are changed by the partition. Consider first the case of candidates selecting the same target sets. Then it must be the case that $k(T_L) + k(T_R)$ is unchanged as $K'(j_\imath) \cap K(j) = \emptyset \ \forall j_\imath, j \in M'$. We denote the bliss point of a subset of the partitioned communities by $E(X|j_\imath)$. Denote the bliss point of the target set with $j_\imath$ omitted by $E(X|T_1 - j_\imath)$ and the bliss point of the target set which is unchanged by $E(X|T_2)$. Then by Lemma 9 for $j_\imath$ to be contained in both target sets it must be that both

$$\begin{aligned}
(k(T_2) + D)\left(E(X|j_\imath) - \frac{k(T_2)E(X|T_2) + DE_c(X)}{k(T_2) + D}\right) \\
< (k(T_2) + k(j_\imath) + D)\left(E(X|j_\imath) - \frac{k(T_1)E(X|T_1) + DE_c(X)}{k(T_1) + D}\right) \\
(k(T_1 - j_\imath) + D)\left(E(X|j_\imath) - \frac{k(T_1 - j_\imath)E(X|T_1 - j_\imath) + DE_c(X)}{k(T_1 - j_\imath) + D}\right) \\
< (k(T_1) + D)\left(E(X|j_\imath) - \frac{k(T_2)E(X|T_2) + k(j_\imath)E(X|j_\imath) + DE_c(X)}{k(T_2) + D + k(j_\imath)}\right)
\end{aligned}$$

We can show that these equations never hold simultaneously taking into account that by Lemmas 1 and 2, it must be that

$$\frac{k(T_2)E(X|T_2) + k(j_\imath)E(X|j_\imath) + DE_c(X)}{k(T_2) + D + k(j_\imath)} > E(X|j_\imath) > \frac{k(T_1)E(X|T_1) + DE_c(X)}{k(T_1) + D}$$

Thus, it must be the case that $j_\imath$ either in T_L or T_R. Due to symmetry, $j_{-\imath} \in T_L$, $j_{+\imath} \in T_R$. Then, there are three cases to consider.

1. $j_{-\imath}, j_{+\imath} \in T'_L$
2. $j_{-\imath}, j_{+\imath} \in T'_R$
3. $j_{-\imath} \in T'_R$, $j_{+\imath} \in T'_L$

Case 1: $j_{-\imath}, j_{+\imath} \in T'_L$ This implies that communities originally contained in T_R now belong to T_L. The new level of polarization is given by

$$\frac{2\left(\sum_{i\in K(T_L)} E_i(X) + \sum_{i\in K'(j_{+\imath})} E_i(X)\right) - k(T_L) - k(j_{+\imath})}{k(T_L) + k(j_{+\imath}) + D}$$
$$+ \frac{2\left(\sum_{i\in K(T_R)} E_i(X) - \sum_{i\in K'(j_{+\imath})} E_i(X)\right) - k(T_R) + k(j_{+\imath})}{k(T_R) + k(j_{+\imath}) + D}$$
$$+ D\left(E_B(X) - E_A(X)\right)\left(\frac{1}{k(T_L) + D + k(j_{+\imath})} + \frac{1}{k(T_R) + D - k(j_{+\imath})}\right)$$

It is straightforward to show that this term is larger than $\frac{2D}{k(T_L)+D}\left(E_B(X) - E_A(X)\right)$, the old level of polarization and thus polarization increases.

Case 2: $j_{-\imath}, j_{+\imath} \in T'_R$ This can never occur, see Lemma 8.

Case 3: $j_{-\imath} \in T'_R$, $j_{+\imath} \in T'_L$ Target sets change symmetrically, as $k(j_{-\imath}) = k(j_{+\imath})$ and $E(X|j_{-\imath}) = 1 - E(X|j_{+\imath})$. This implies that the level of polarization is unchanged, that is polarization still equals $\frac{2D}{k(T_L)+D}\left(E_B(X) - E_A(X)\right)$

We then turn to symmetric equilibria in which candidates select different target sets. Again, we can show that j_i is only contained in one target set, but not in both, which follows from the same equations as before. We assume first that $j_{-\imath} \in T_L$, $j_{+\imath} \in T_R$. Then, we distinguish again between the following two cases

1. $j_{-\imath}, j_{+\imath} \in T'_L$
2. $j_{-\imath} \in T'_R$, $j_{+\imath} \in T'_L$

Case 1:$j_{-\imath}, j_{+\imath} \in T'_L$ Here again outlets are omitted from T_R and added to T_L.

$$\frac{2\left(\sum_{i\in K(T_L)} E_i(X) + \sum_{i\in K'(j_{+\imath})} E_i(X)\right) - k(T_L) - k(j_{+\imath})}{k(T_L) + k(j_{+\imath}) + D}$$
$$+ \frac{2\left(\sum_{i\in K(T_R)} E_i(X) - \sum_{i\in K'(j_{+\imath})} E_i(X)\right) - k(T_R) + k(j_{+\imath})}{k(T_R) - k(j_{+\imath}) + D}$$
$$+ D\left(E_B(X) - E_A(X)\right)\left(\frac{1}{k(T_L) + D + k(j_{+\imath})} + \frac{1}{k(T_R) + D - k(j_{+\imath})}\right)$$

Before partitioning the level of polarization was given by

$$\frac{2\left(\sum_{i\in K(T_L)} E_i(X)\right) - k(T_L)}{k(T_L)+D} + \frac{2\left(\sum_{i\in K(T_R)} E_i(X)\right) - k(T_R)}{k(T_R)+D} + D\left(E_B(X) - E_A(X)\right)\left(\frac{1}{k(T_L)+D} + \frac{1}{k(T_R)+D}\right)$$

and we can again establish that polarization has increased due to partitioning.

Case 2: $j_{-\imath} \in T'_R$, $j_{+\imath} \in T'_L$ Again, outlets switch between target sets. This implies that the number of communities in each set remains unchanged. It therefore suffices to show that

$$\frac{2\left(\sum_{i\in K(T_L)} E_i(X) + \sum_{i\in K'(j_{+\imath})} E_i(X) - \sum_{i\in K'(j_{-\imath})} E_i(X)\right) - k(T_L)}{k(T_L)+D} + \frac{2\left(\sum_{i\in K(T_R)} E_i(X) - \sum_{i\in K'(j_{+\imath})} E_i(X) + \sum_{i\in K'(j_{-\imath})} E_i(X)\right) - k(T_R)}{k(T_R)+D} > \frac{2\left(\sum_{i\in K(T_L)} E_i(X)\right) - k(T_L)}{k(T_L)+D} + \frac{2\left(\sum_{i\in K(T_R)} E_i(X)\right) - k(T_R)}{k(T_R)+D}$$

which always holds due to the symmetry and $\frac{\sum_{i\in K'(j_{-\imath})} E_i(X)}{k(j_{-\imath})} > \frac{1}{2}$. Otherwise it can never be optimal to add $j_{-\imath}$ to T_R, see Lemma 1.

Next we consider the case where $j_{-\imath}, j_{+\imath} \in T_L$. Then, outlets can be omitted from T_L and added to T_R and polarization can increase or decrease. Polarization increases if and only if

$$\frac{2\left(\sum_{i\in K(T_L)} E_i(X) - \sum_{i\in K'(j_\imath)} E_i(X)\right) - k(T_L) + k(j_\imath)}{k(T_L) - k(j_\imath) + D} + \frac{2\left(\sum_{i\in K(T_R)} E_i(X) + \sum_{i\in K'(j_\imath)} E_i(X)\right) - k(T_R) - k(j_\imath)}{k(T_R) + k(j_\imath) + D} + D\left(E_B(X) - E_A(X)\right)\left(\frac{1}{k(T_L)+D-k(j_{+\imath})} + \frac{1}{k(T_R)+D+k(j_{+\imath})}\right) > \frac{2\left(\sum_{i\in K(T_L)} E_i(X)\right) - k(T_L)}{k(T_L)+D} + \frac{2\left(\sum_{i\in K(T_R)} E_i(X)\right) - k(T_R)}{k(T_R)+D} + D\left(E_B(X) - E_A(X)\right)\left(\frac{1}{k(T_L)+D} + \frac{1}{k(T_R)+D}\right)$$

Equating the left and right hand side of the inequality and solving for $\sum_{i\in K'(j_i)} E_i(X)$ shows that polarization is increasing for $\sum_{i\in K'(j_i)} E_i(X)$ sufficiently large.

Proof of Proposition 6

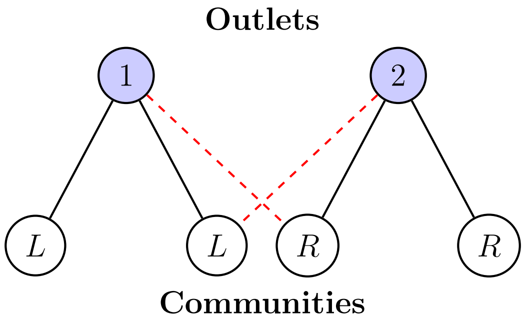


Figure 10: Polarization Decreases

For an example when polarization decreases, consider Figure 10. Let $E_X(B) = 1$, $E_R(X) = 3/4$, by symmetry it follows that $E_L(X) = 3/4$. Let $D = 2$. The level of polarization is given by 1. Once links are added symmetrically, it is still optimal to set $T'_L = \{1\}$ and $T'_R = \{2\}$. Polarization is now given by 4/5 and has thus decreased. Consider next

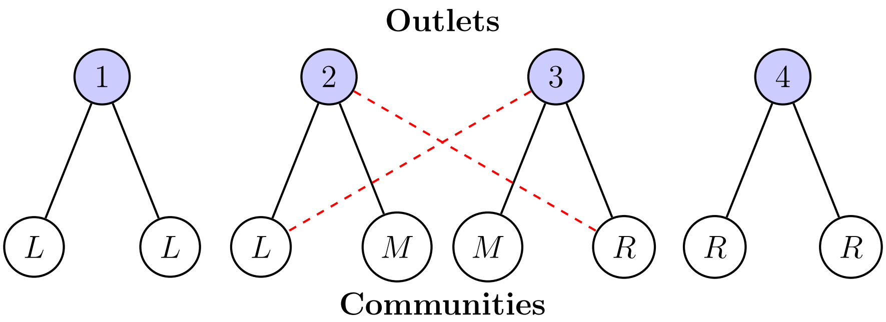


Figure 11: Polarization Increases

an example where polarization increases. Let the network be as depicted in Figure 11. Let $E_M(X) = 1/2$. The other bliss points are as in the previous example. Then, $T_L = \{1, 2\}$, $T_R = \{3, 4\}$ and $T'_L = \{1, 3\}$, $T'_R = \{4\}$. Polarization is initially given by 1/3, after links are added polarization is at 7/16.

Proof of Proposition 7 The optimal policy for a given target set in state s is

$$x_s = \frac{\sum_{i\in K(T_s)} E_i(X) + \sum_{i\in K\setminus K(T_s)} q_{is} E_i(X)}{k(T_s) + \sum_{i\in K\setminus K(T_s)} q_{is}}.$$

We define

$$E^w(X|T_s) = \frac{\sum_{i\in K(T_s)} E_i(X) + \sum_{i\in K\setminus K(T_s)} q_{is} E_i(X)}{k(T_s) + \sum_{i\in K\setminus K(T_s)} q_{is}}$$
$$E^w(X^2|T_s) = \frac{\sum_{i\in K(T_s)} E_i(X^2) + \sum_{i\in K\setminus K(T_s)} q_{is} E_i(X^2)}{k(T_s) + \sum_{i\in K\setminus K(T_s)} q_{is}}$$

As $\overline{q} \to 1$, $q_s \to 0$. Then,

$$x_s = \frac{\sum_{i\in K(T_s)} E_i(X))}{k(T_s)} = E(X|T_s)$$
$$E^w(X|T_s) = \frac{\sum_{i\in K(T_s)} E_i(X)}{k(T_s)} = E(X|T_s)$$
$$E^w(X^2|T_s) = \frac{\sum_{i\in K(T_s)} E_i(X^2)}{k(T_s)} = E(X^2|T_s)$$

This implies that equation (10) simplifies to

$$k(T_{\tilde{s}})\left(E(X^2|T_{\tilde{s}}) - E(X|T_{\tilde{s}})^2\right) + \int_0^1 \sum_{i\in K\setminus K(T_{\tilde{s}})} \left(\sum_{s\neq \tilde{s}, s\in S} (x_s - x)^2 q_{is} + (\pi - x)^2 \overline{q}\right) g_i(x)dx$$

It is immediate that

$$\lim_{\overline{q}\to 1} \int_0^1 \sum_{i\in K\setminus K(T_{\tilde{s}})} \left(\sum_{s\neq \tilde{s}, s\in S} (x_s - x)^2 q_{is} + (\pi - x)^2 \overline{q}\right) g_i(x)dx = \int_0^1 \sum_{i\in K\setminus K(T_{\tilde{s}})} (\pi - x)^2 g_i(x)dx,$$

which completes the proof.

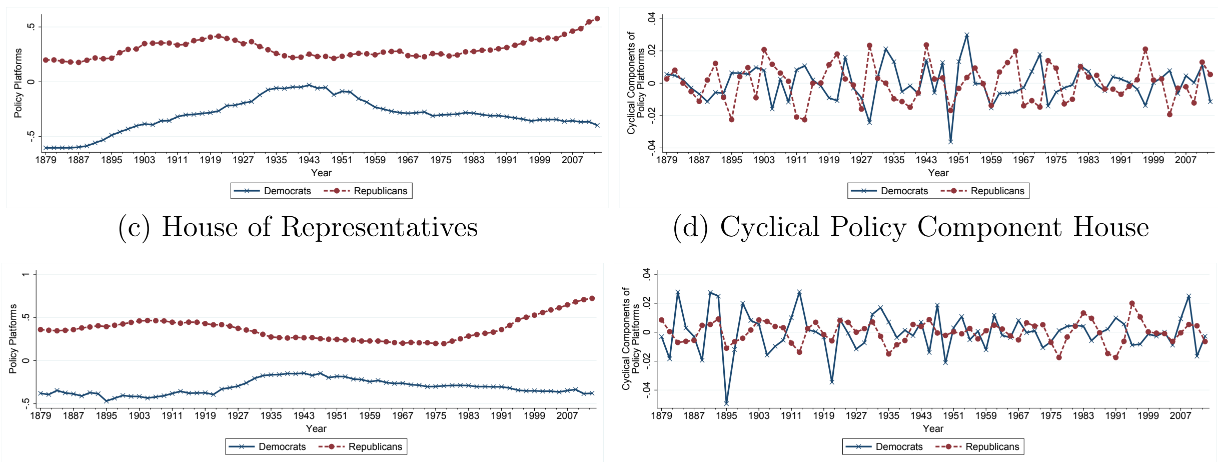
Supplementary Material

Appendix B: Trends and Cycles in Polarization

Figure 12: Cycling of Policies in Senate and House of Representatives

(a) Senate (b) Cyclical Policy Component Senate

(c) House of Representatives (d) Cyclical Policy Component House



Note: The figures on the left hand side show the party means on the [-1,1] scale over time for the House and Senate. On the right side we show the deviations from the trend. We apply an HP filter with smoothing parameter 6.25 to obtain the cyclical components.

We show that there are not only cycles in polarization, but parties' positions fluctuate themselves as well. In order to see this we consider the parties' average position both in the Senate and House, see Figures 12a and 12c. These figures highlight that in both chambers of Congress, Democrats choose a more left-wing position than Republicans. As there is a trend in the platforms of Democrats and Republicans, we again apply a HP filter, which identifies the cyclical component of the party platforms. The cycles in the platforms are depicted in Figures 12b and 12d. This highlights that in different years parties platforms are more central, in others they are more extreme. We can show that these cycles explain about 2% of the trend on average, when taking into account years later than 1990.

But from the graphs it does not become clear how exactly the party positions fluctuate around the trend. In order to investigate this, we estimate various ARMA specifications and select the best model based on the information criteria (AIC and BIC).[57] The results, given in Table 4, show that there is a clear negative correlation with the second lag. That is, for a given presidential election year the position of the party in the next election year is

[57]The best model for the House democrats is an ARMA (2,1), the best fit for all other platform changes is an ARMA(2,2).

Table 4: Autoregressive Processes

	House of Representatives		Senate	
	Dem. Position	Rep. Position	Dem. Position	Rep. Position
1st Lag	0.465***	1.024***	1.090***	1.412***
	(0.102)	(0.143)	(0.304)	(0.0916)
2nd Lag	-0.287*	-0.683***	-0.501**	-0.688***
	(0.123)	(0.107)	(0.190)	(0.0907)
Observations	68	68	68	68

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, standard errors in parentheses

Note: The estimation results given here are the processes with the best AIC and BIC. For columns (1)-(4): ARMA(2,1), ARMA(2,2), ARMA (2,2), ARMA(2,2), respectively.

significantly more conservative (liberal) and becomes in the next election cycle more liberal (conservative) again and thus, we observe electoral cycles.

Appendix C: Further Examples and Extensions

Restricted Targeting To see some of the complications that arise when candidates can only target a limited number of outlets, consider the case where all communities have the same bliss point. Consider the media network depicted in Figure (13). Outlet 2 has the highest degree in that it caters to 4 distinct communities. Clearly, candidates will target such an outlet if they are constrained to not disclose their policy in more than one outlet.

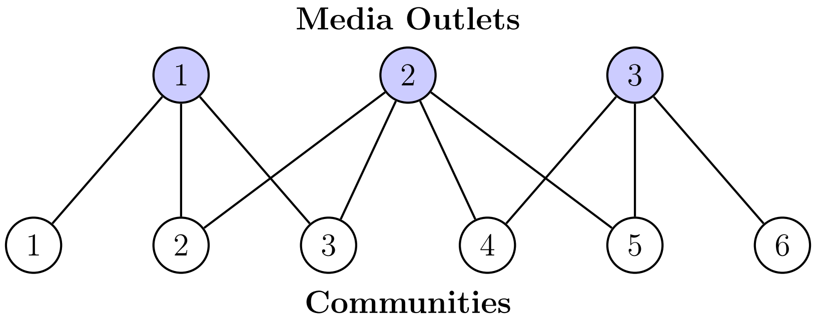


Figure 13: Maximal Target Set

However, if politicians can target two outlets, they are better off disclosing in outlets 1 and 3. Targeting outlets 1 and 2 covers 5 communities, whereas targeting outlets 1 and 3 covers 6. The example highlights the trade-offs implicit at the targeting stage when the communities are homogenous and there are limits to full disclosure.

Due to the complications arising when only a limited number of outlets can be targeted, we allow for an arbitrary number of outlets being targeted. Note that this does not imply that all outlets will be targeted if communities are heterogeneous. If we restrict attention to networks without overlap, that is there are no communities that belong to more than one outlet, where media outlets have equal sized audiences, then targeting with and without costs leads to similar results. Candidates target voters whose preferences have the greatest distance from the prior until they cannot afford to target an additional outlet.

Partition: Target Set Switches It can always be the case that the expected bliss point of the optimal target set after the partition lies on a different side of the prior, that is it can always be the case that $E(X|T) < \pi < E'(X|T')$, where T' denotes the target set after the partition. An example of this is given in Figure 14. We consider three types of nodes, left nodes with $E_L(X) = 1/4$, right nodes with bliss point $E_R(X) = 3/4$ and medium points with $E_M(X) = 2/3$. The prior is given by $\pi = 1/2$. If only two outlets, outlet 1 and 2 exist, it is optimal to target outlet 1. Once outlet 3 is created, which only caters to a right community previously connected to outlet 1, it is instead optimal to target outlets 2 and 3.

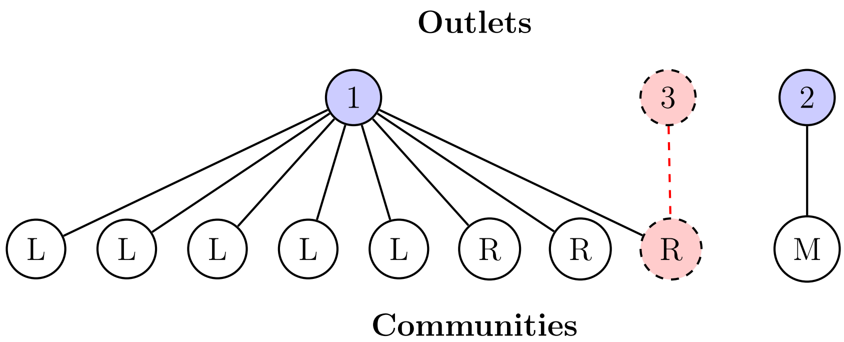


Figure 14: Partition: Target Set Switches

Adding Links: Target Set Switches We assume that $E(X|T) < \pi$. Then, it can always be the case that $E'(X|T') > \pi$ – that is, the expected bliss point of the new optimal target set might lie on the other side of the prior no matter how the additional link is formed. We provide an example in Figure 15 of such a switch in target sets if a link is added between a targeted outlet and a targeted community. The original network N is given by the black links the new link is the dashed red link. Suppose that $\pi = 1/2$, and consider two types of communities – namely left communities with $E_L(X) = 1/4$ and right communities with $E_R(X) = 2/3$. It is straightforward to verify that under linking structure N it is optimal to disclose in outlets 1 and 2, whereas once the additional link is formed, it is optimal to disclose in outlets 2 and 3.

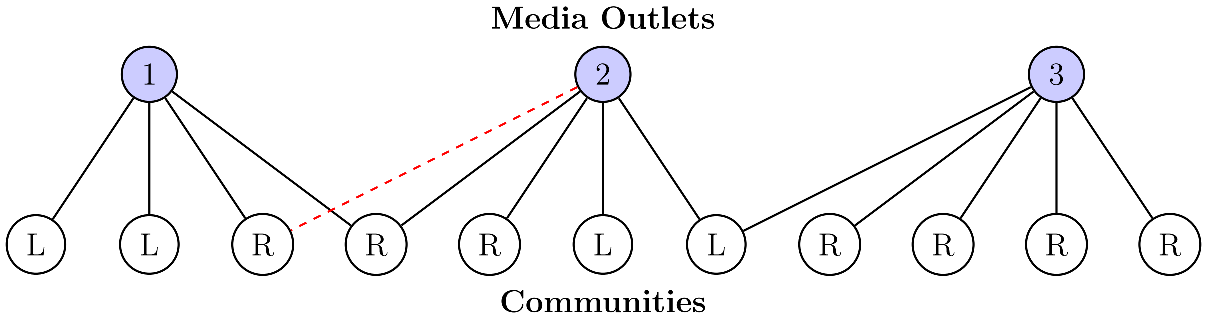


Figure 15: Target Set Switches

Adding Links: $l \notin T$ We consider a community that is contained in the target set and forms a link to an outlet l not contained in the target set and show that if j is omitted, then outlet l can be added to the target set. This result is highlighted in Figure 16. The expected bliss points of the various groups are shown in the nodes.

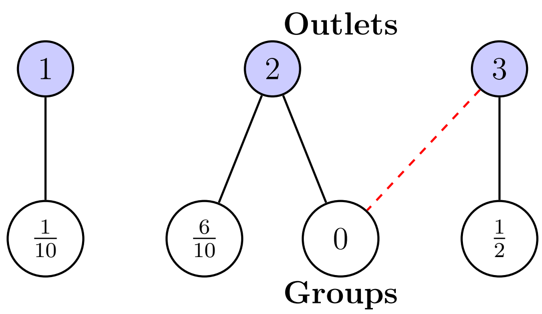


Figure 16: Target Set Changes

We set $\pi = \frac{1}{2}$. Then, the original target set, T, in absence of the red, dashed link contains outlets 1 and 2. If the group with expected bliss point of zero forms a link to outlet 3, then it is no longer optimal to include outlet 2, but instead it is better to include outlet 3, that is $T' = \{1, 3\}$.

Comparative Statics: Polarization is given by

$$\Delta = \frac{1}{2}\Bigg(\frac{\sum_{i \in T_{1B}} E_i(X) + DE_B(X)}{k(T_{1B}) + D} + \frac{\sum_{i \in T_{2B}} E_i(X) + DE_B(X)}{k(T_{2B}) + D} - \frac{\sum_{i \in T_{1A}} E_i(X) + DE_A(X)}{k(T_{1A}) + D} - \frac{\sum_{i \in T_{2A}} E_i(X) + DE_A(X)}{k(T_{2A}) + D} \Bigg)$$

This term is increasing in D and $E_B(X)$ and decreasing $E_A(X)$.

1. Change in D

$$\frac{\partial \Delta}{\partial D} = \frac{k(T_{1B})E_B(X) - \sum_{i \in T_{1B}} E_i(X)}{(k(T_{1B}) + D)^2} + \frac{k(T_{2B})E_B(X) - \sum_{i \in T_{2B}} E_i(X)}{(k(T_{2B}) + D)^2}$$
$$- \frac{k(T_{1A})E_A(X) - \sum_{i \in T_{1A}} E_i(X)}{(k(T_{1A}) + D)^2} - \frac{k(T_{2A})E_A(X) - \sum_{i \in T_{2A}} E_i(X)}{(k(T_{2A}) + D)^2} > 0$$

as $E_B(X) > E_i(X) > E_A(X)$ for any $E_i(X)$.

2. Change in $E_A(X)$

$$\frac{\partial \Delta}{\partial E_A(X)} = -\frac{D}{k(T_{1A}) + D} - \frac{D}{k(T_{2A}) + D} < 0$$

3. Change in $E_B(X)$

$$\frac{\partial \Delta}{\partial E_B(X)} = \frac{D}{k(T_{1B}) + D} + \frac{D}{k(T_{2B}) + D} > 0$$

Media Networks & Polarization: Constant Prior Using assumptions A.1 and A.2, we show under what circumstances polarization increases or decreases if outlets are partitioned. This depends on whether the policy the candidate chooses lies to the left or the right of the prior. We consider the problem from candidate B's perspective. If candidate B chooses a policy that lies to the right of his prior, then polarization is higher than if he chooses a policy to the left of his prior. Partitioning allows candidates to target the voters they aim for more precisely. If candidate B chooses a target set to the right of the prior, then a partition allows to target more ideologically extreme voters and this leads to an increase in polarization. If on the other hand, the candidate selects a target set to his left, then a partition results in a policy that is more centrist and leads to a decrease in polarization. This is summarized in Proposition 8. We denote by T_{BL} (T_{BR}) the target set of candidate B if the policy lies to the left (right) of the prior. We further refer to the policy associated with the left target set as x_{BL}, that associated with the right target set as x_{BR}. We index the target sets and policies after the partition by prime. We distinguish between two cases, namely (i) both partitioned outlets are contained in target set and (ii) at least one outlet is not contained in the target set. We denote by $E(X|j_\iota)$ the bliss point of the communities *added to the target set.*

Proposition 8. *Let $K'(j_\iota) \cap K(j) = \emptyset \ \forall j_\iota, j \in M'$ and let partitions be symmetric. Consider a symmetric equilibrium.*

(1) Let π_B be sufficiently low, such that $\max\{x_{BL}, x'_{BL}\} < \pi_B$.

(a)If $j_-, j_+ \in T_{BL}$, polarization weakly decreases.

(b)If at least $j_+ \notin T_{BL}$ and $E(X|j_\imath) > x_{BL}$, polarization increases. Otherwise, polarization weakly decreases.

(2) If $x_{BL} < \pi_B < x'_{BL}$, polarization increases.

(3) Let π_B be sufficiently high, such that $\max\{x_{BR}, x'_{BR}\} > \pi_B$. *If $E(X|j_\imath) < x_{BR}$, polarization decreases. Otherwise, polarization weakly increases.*

(4) If $x_{BR} > \pi_B > x'_{BR}$, polarization decreases.

Proposition 8 shows that partitioning can increase or decrease polarization depending on whether the candidate selects a moderate policy or an extreme one. Generally a partition allows for a more precise targeting of voters and a partition decreases polarization if candidates select a moderate policy, whereas it induces a rise in polarization if candidates select a more extreme policy. However, it can also be the case that a partition increases polarization in case of a moderate policy. This occurs if at least one of the partitioned outlets is not contained in the target set. Then, it can be the case that communities are added to the target set, which have a bliss point that lies above the previously implemented policy, leading to an increase in polarization. By the same logic it can also be the case that a partition decreases polarization if more ideologically extreme voters are targeted as at least one the partitioned outlets is not contained in the target set. Note again that it cannot be the case that both partitioned outlets are contained in B's right target set. Further, it can always be the case that the target sets switch, that is before the partition it was better to choose a policy to the left of the prior whereas it is now better to choose a policy to the right of the prior.

Proposition 9. *Suppose communities i_-, i_+ form an additional link. Then, polarization can increase or decrease.*

An additional complication arising in the added links case, compared to the partitions, is that it not only matters whether the communities that form a link are contained in the target set or not, but also whether they form link to an outlet in the target set or out of the target set. This changes the expected bliss point of the target set and can lead to outlets being added to or omitted from the target set and ultimately to an increase or decrease in the target set. If both communities that form an additional link are contained in the target set and the policy caters to more ideologically extreme voters, then polarization weakly increases, see Lemma 4.

If on the other hand the target set is not affected by the addition of links, then polarization can increase or decrease, depending on the effects of the target set.

The proofs of both propositions are omitted as they follow exactly the proofs of Propositions 5 and 6.

Rational Voters & Uncertainty The posterior probability is given by

$$q_{is} = P(x_s|i \text{ not targeted}) = \frac{P(i \text{ not targeted}|x_s)P(x_s)}{P(i \text{ not targeted})}.$$

Note that $P(i \text{ not targeted}|x_s) = 1$ as for each state voters know who is targeted and who is not. But communities can differ in the number of states in which they are targeted. Taking this into account, we can write

$$q_{is} = \frac{p_s}{\sum_{s=1}^{n} \sum_{i \in K \setminus K(T_s)} p_s}.$$

The maximization problem of candidate c if $D = 0$ is given by

$$\max_{x_c, T_c} \sum_{i \in K} \frac{1}{k} \int_0^1 P_c(y_A^i(T_A), y_B^i(T_B)|x) dG_i(x), \tag{31}$$

which is equivalent to

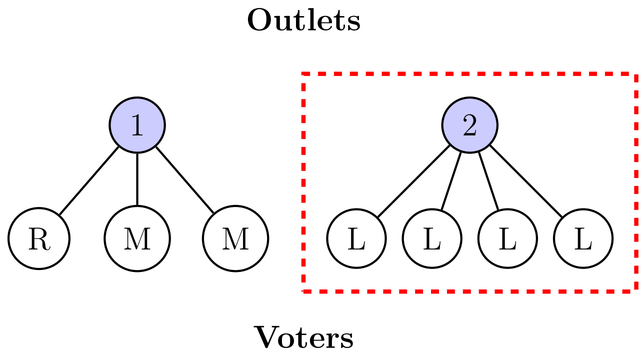
$$\max_{x_c, T_c} \sum_{i \in K(T_c)} \int_0^1 u(x_c|x) g_i(x) dx + \sum_{i \in K \setminus K(T_c)} \sum_{s=1}^{n} q_{is} \int_0^1 u(x_s|x) g_i(x) dx$$

Denote the state that is realized by $\tilde{s}$. Now, $x_{\tilde{s}}$, the policy that is implemented is one of the policies that the non-targeted voters assign positive probability to. Taking this into account and dropping the subscripts for the candidate, the maximization problem becomes

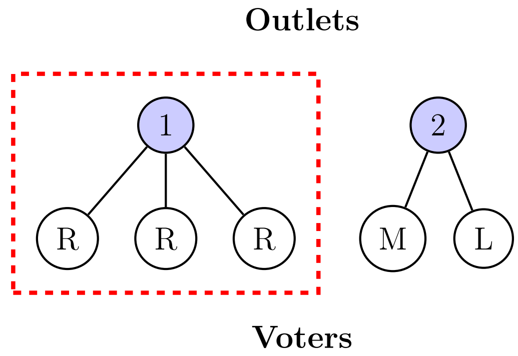
$$\begin{aligned} \max_{x_{\tilde{s}}, T_{\tilde{s}}} \sum_{i \in K(T_{\tilde{s}})} \int_0^1 u(x_{\tilde{s}}|x) g_i(x) dx &+ \sum_{i \in K \setminus K(T_{\tilde{s}})} q_{i\tilde{s}} \int_0^1 u(x_{\tilde{s}}|x) g_i(x) dx \\ &+ \sum_{i \in K \setminus K(T_{\tilde{s}})} \sum_{s \neq \tilde{s}, s \in S} q_{is} \int_0^1 u(x_s|x) g_i(x) dx \end{aligned} \tag{32}$$

Similarly to the proof of Proposition 1, we fix the target set. We can ignore the third part of the previous equation as it is independent of $x_{\tilde{s}}$. We are left with an equation that only depends on the realized state. To keep our notation sparse, we replace $\tilde{s}$ by s for the realized state and obtain equation 10. We then consider an example to illustrate our result. To keep this as simple as possible, we assume that everyone knows the bliss point of the communities, which can take on three possible values. In particular, a community can have bliss point $x_L = 1/2$, $x_M = 6/10$ or $x_R = 7/10$. The uncertainty emerges as the voters do not know about the network structure. We assume that there are three possible states of the world, with each state determining a specific network. The probabilities assigned to the three states are $q_1 = 1/2$, $q_2 = 9/20$ and $q_3 = 1/20$. The networks are depicted in Figure 17.

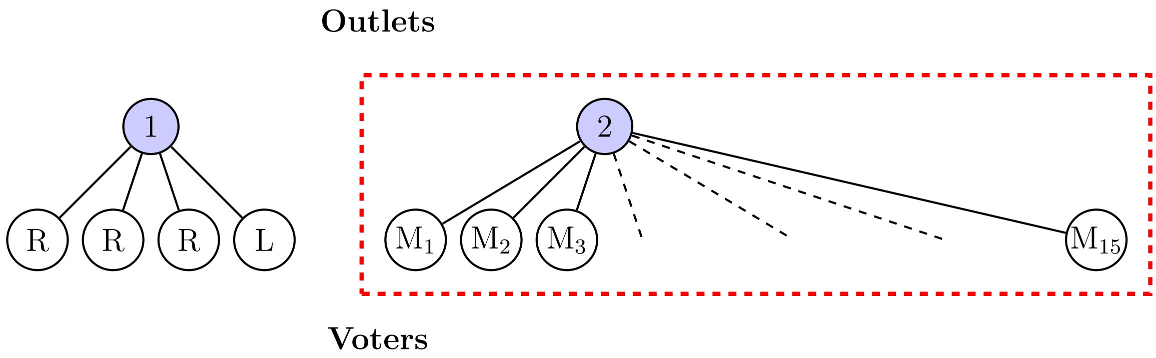
Figure 17: 3 Groups of Voters & 2 Outlets



(a) State 1



(b) State 2



(c) State 3

Note: The target sets in the states are given in the dashed boxes. The communities can have bliss points L, M or R. There are 15 communities with bliss point M in state 3.

In state 3, there are 15 communities with moderate bliss points. The target sets are depicted in the dashed boxes.

We calculate the optimal policy for each possible target set for each state and compare the payoff to the candidate across target sets. Based on this, we find the optimal target set. On the equilibrium path, beliefs are consistent with selected policies. If a candidate deviates from the equilibrium strategy, then voters assign zero probability to the policy he chooses. Consider as an example that instead of targeting outlet 2 in state 1, the candidate chooses outlet 1. In this case, the left communities that are no longer targeted assign zero probability to the policy associated with this target set, but believe with the same posterior as before that they are in either state 2 or 3.

The results show that in each state of the world a different group of voters is targeted. In state 1, the left voters are targeted, in state 2 those on the right and in state 3 the moderate ones. An overview of the policies in the states of the world, versus the average expected preference is given in Table 5. As can be seen the implemented policies do not match the

x_1	x_2	x_3	$E(X\|s=1)$	$E(X\|s=2)$	$E(X\|s=3)$
.548	.648	.601	.557	.640	0.610

Table 5: Policies Versus Mean of Preferences

expected bliss points of the voters in any state of the world. In particular, when the left voters are targeted in state 1, then the policy chosen is to the left of the mean bliss point and the reverse holds true if the candidate selects right voters in state 2. Then, the implemented policy lies to the right of then mean bliss point. The exact same pattern can be replicated with voters who are not fully rational and different priors. This is shown in Table 6, where the subscripts denote the state. Note however that in the example voters can infer significantly

π_1	π_2	π_3	$E(X\|T_1)$	$E(X\|T_2)$	$E(X\|T_3)$
.6	.5	.65	.5	.7	.6

Table 6: Priors and Policies with Behavioral Voters

more than we would expect in reality. First, in each target set, there is only one type of voter. If we allow, for example, for a state in which there are some left communities that are targeted and some that are not targeted, then this adds additional uncertainty. Further, if voters are not entirely sure what preferences communities have, then they have again a harder time making an inference about what the policy will be. Going back to our initial example of Chris Christie's targeting strategy, it is probably not clear to a viewer of either one of the telenovelas what the political preferences of the audiences are. So, we have shown that even with a relatively low level of uncertainty, targeting specific voters and choosing a policy that caters to their preferences is better than choosing the policy that takes into account all voters to the same extent.

Appendix D: Additional Proofs

Proof of Lemma 12 We remove these assumptions , namely: (1) coverages have no overlap, (2) $j = T \backslash T'$, one after the other. We first consider the case in which all outlets possibly overlap. We adjust the notation in order to make the proof less notationally intensive. Denote by k_1 the number of groups that belong only to outlet l and by $\hat{x}_1 = E(X|l - j - o)$ the bliss point in these groups; by k_2 the number of groups that belong both to outlet l and j and by $\hat{x}_2$ the bliss point in these groups; by k_3 the number of groups that belong both to outlets l and o and by $\hat{x}_3$ the bliss point in these groups; by k_4 the number of groups that belong only to all 3 outlets and by $\hat{x}_4$ the bliss point in these groups. There is one group that forms an additional link, which leads to $k_5 = 1$. The bliss point of this group is denoted by $\hat{x}_5$ and the group belongs to outlet j. The number of groups exclusively associated to outlet j is $k_5 + k_6$. The bliss point of the groups the groups only connected to outlet j, excluding the group that forms a link is $\hat{x}_6$. Last, k_7 is the number of groups that belong both to outlets j and o and $\hat{x}_7$ the bliss point in these groups. Finally, $\hat{x}_8$ gives the bliss point of the k_8 groups that only belong to outlet o.

If $T = \{l, j\}$, but $o \in T'$, then it has to hold that

$$\left(\sum_{i=1}^{7} k_i\right)\left(\frac{\sum_{i=5}^{7} k_i\hat{x}_i}{\sum_{i=5}^{7} k_i} - \pi\right)^2 > \left(\sum_{i=1}^{4} k_i\right)\left(\frac{\sum_{i=5}^{7} k_i\hat{x}_i}{\sum_{i=5}^{7} k_i} - \frac{\sum_{i=1}^{4} k_i\hat{x}_i}{\sum_{i=1}^{4} k_i}\right)^2$$

$$\left(\sum_{i=1}^{7} k_i\right)\left(\frac{\sum_{i=6}^{7} k_i\hat{x}_i}{\sum_{i=6}^{7} k_i} - \pi\right)^2 < \left(\sum_{i=1}^{5} k_i\right)\left(\frac{\sum_{i=6}^{7} k_i\hat{x}_i}{\sum_{i=6}^{7} k_i} - \frac{\sum_{i=1}^{5} k_i\hat{x}_i}{\sum_{i=1}^{4} k_i}\right)^2$$

$$\left(\sum_{i=1}^{8} k_i\right)(\hat{x}_8 - \pi)^2 < \left(\sum_{i=1}^{7} k_i\right)\left(\hat{x}_8 - \frac{\sum_{i=1}^{7} k_i\hat{x}_i}{\sum_{i=1}^{7} k_i}\right)$$

$$\left(\sum_{i=1}^{8} k_i - k_6\right)\left(\frac{\sum_{i=7}^{8} k_i\hat{x}_i}{\sum_{i=7}^{8} k_i} - \pi\right)^2 > \left(\sum_{i=1}^{5} k_i\right)\left(\frac{\sum_{i=7}^{8} k_i\hat{x}_i}{\sum_{i=7}^{8} k_i} - \frac{\sum_{i=1}^{5} k_i\hat{x}_i}{\sum_{i=1}^{5} k_i}\right)$$

$$\left(\sum_{i=1}^{7} k_i\right)\left(\frac{\sum_{i=1}^{7} k_i\hat{x}_i}{\sum_{i=1}^{7} k_i} - \pi\right)^2 > \left(\sum_{i=1}^{4} k_i + \sum_{i=7}^{8} k_i\right)\left(\frac{\sum_{i=1}^{4} k_i\hat{x}_i + \sum_{i=7}^{8} k_i\hat{x}_i}{\sum_{i=1}^{5} k_i + \sum_{i=7}^{8} k_i} - \pi\right)^2$$

$$\left(\sum_{i=1}^{8} k_i\right)(\hat{x}_6 - \pi)^2 < \left(\sum_{i=1}^{5} k_i + \sum_{i=7}^{8} k_i\right)\left(\hat{x}_6 - \frac{\sum_{i=1}^{5} k_i\hat{x}_i + \sum_{i=7}^{8} k_i\hat{x}_i}{\sum_{i=1}^{5} k_i + \sum_{i=7}^{8} k_i}\right)^2$$

To further simplify notation I set

$$k_a = \textstyle\sum_{i=1}^{4} k_i \quad \text{and} \quad \hat{x}_a = \textstyle\sum_{i=1}^{4} k_i\hat{x}_i.$$

Similar to before I now define

$$
\begin{aligned}
B_0 &= k_a + 1,\\
B_1 &= k_a + 1 + k_6 + k_7,\\
B_2 &= k_a + 1 + k_7 + k_8,\\
B_3 &= k_a + 1 + k_6 + k_7 + k_8.
\end{aligned}
$$

Again, it has to be the case that

$$\hat{x}_8 > \frac{1}{k_8}\left(B_3\pi - \hat{x}_a - \hat{x}_5 - k_7\hat{x}_7 - \sqrt{\frac{B_3}{B_1}}\left(B_1\pi - \hat{x}_a - \hat{x}_5 - k_7\hat{x}_7\right) + k_6\hat{x}_6\left(-1 + \sqrt{\frac{B_3}{B_1}}\right)\right).$$

as well as

$$\hat{x}_8 < \frac{1}{k_8}\left(B_2\pi - \hat{x}_a - \hat{x}_5 - k_7\hat{x}_7 - \sqrt{\frac{B_2}{B_0}}\left(B_0\pi - \hat{x}_a - \hat{x}_5\right)\right).$$

For this to hold it has to be the case that

$$
\begin{aligned}
&\left(B_2\pi - \hat{x}_a - \hat{x}_5 - k_7\hat{x}_7 - \sqrt{\frac{B_2}{B_0}}\left(B_0\pi - \hat{x}_a - \hat{x}_5\right)\right)\\
&> \left(B_3\pi - \hat{x}_a - \hat{x}_5 - k_7\hat{x}_7 - \sqrt{\frac{B_3}{B_1}}\left(B_1\pi - \hat{x}_a - \hat{x}_5 - k_7\hat{x}_7\right) + k_6\hat{x}_6\left(-1 + \sqrt{\frac{B_3}{B_1}}\right)\right)
\end{aligned}
$$

which is equivalent to

$$\frac{1}{\sqrt{\frac{B_3}{B_1}} - 1}\left(\left(B_2 - B_3 + \sqrt{B_1B_3} - \sqrt{B_0B_2}\right)\pi + \left(\sqrt{\frac{B_2}{B_0}} - \sqrt{\frac{B_3}{B_1}}\right)\left(\hat{x}_a + \hat{x}_5\right) - \sqrt{\frac{B_3}{B_1}}k_7\hat{x}_7\right) > k_6\hat{x}_6.$$

We have again a lower bound on $\hat{x}_6$, namely

$$
\begin{aligned}
(k_6 + k_7)k_6\hat{x}_6 > (1 + k_6 + k_7)B_1\pi - (1 + k_6 + k_7)\hat{x}_a - (1 + k_6 + k_7)\hat{x}_5 - (k_6 + k_7)k_7\hat{x}_7\\
-(1 + k_6 + k_7)\sqrt{\frac{B_1}{B_0}}(B_0\pi - \hat{x}_a - \hat{x}_5)
\end{aligned}
$$

To simplify notation once again, we let

$$B_4 = \frac{1 + k_6 + k_7}{k_6 + k_7}$$

Then,

$$k_6\hat{x}_6 > B_4B_1\pi - B_4\hat{x}_a - B_4y5 - k_7\hat{x}_7 - B_4\sqrt{\frac{B_1}{B_0}}(B_0\pi - \hat{x}_a - \hat{x}_5)$$

and so it has to be the case that

$$\frac{1}{\sqrt{\frac{B_3}{B_1}}-1}\left(\left(B_2 - B_3 + \sqrt{B_1B_3} - \sqrt{B_0B_2}\right)\pi + \left(\sqrt{\frac{B_2}{B_0}} - \sqrt{\frac{B_3}{B_1}}\right)(\hat{x}_a + \hat{x}_5) - \sqrt{\frac{B_3}{B_1}}k_7\hat{x}_7\right)$$
$$> B_4(B_1\pi - \hat{x}_a - \hat{x}_5) - k_7\hat{x}_7 - B_4\sqrt{\frac{B_1}{B_0}}(B_0\pi - \hat{x}_a - \hat{x}_5)$$
$$\frac{1}{\sqrt{\frac{B_3}{B_1}}-1}\left(\left(B_2 - B_3 + \sqrt{B_1B_3} - \sqrt{B_0B_2}\right)\pi + \left(\sqrt{\frac{B_2}{B_0}} - \sqrt{\frac{B_3}{B_1}}\right)(\hat{x}_a + \hat{x}_5) - \left(\sqrt{\frac{B_3}{B_1}} - 1\right)k_7\hat{x}_7\right)$$
$$> B_4\left(B_1\pi - \hat{x}_a - \hat{x}_5 - \sqrt{\frac{B_1}{B_0}}(B_0\pi - \hat{x}_a - \hat{x}_5)\right)$$

Comparing this to equations (26), we can see that the left hand side has decreased and the right hand side has increased. Thus, as (26) did not hold, it cannot be that case that this equation holds. Given this result, we can restrict attention to the case without overlap.

Consider last the case where an additional outlet is also dropped from the target set. Then we simply need to redefine the bliss points of the groups connected to outlet o and the logic remains the same and the results as well.

Proof of Lemma 5 (1)

We assumed that $E(X|T^5) \neq E(X|T^3)$. Then again, similar to the previous step, it now has to hold that

$$k(T^3)\left(E(X|T^3) - E(X|T^2)\right)^2 > k(T^5)\left(E(X|T^5) - E(X|T^2)\right)^2$$
$$k(T^5)\left(E(X|T^5) - E(X|T^4)\right)^2 > k(T^3)\left(E(X|T^3) - E(X|T^4)\right)^2$$

which by Lemma 10 **C.2** is the case if and only if

$$E(X|T^4) < E(X|T^5) < E(X|T^3) \tag{33}$$

Additionally, it has to hold that

$$k(T^3)\left(E(X|T^3) - E(X|T^2)\right)^2 > k(T^4)\left(E(X|T^4) - E(X|T^2)\right)^2$$
$$k(T^4)\left(E(X|T^4) - E(X|T^3)\right)^2 > k(T^5)\left(E(X|T^5) - E(X|T^3)\right)^2,$$

which together with the previous equations is fulfilled if and only if

$$E(X|T^2) < E(X|T^5) < E(X|T^3) \tag{34}$$

We then have restricted the region in which $E(X|T^5)$ can lie and established the following ordering

$$E(X|T^1) < E(X|T^4) < E(X|T^2) < E(X|T^5) < E(X|T^3) \tag{35}$$

We now consider $E(X|T^6)$. Again, if $E(X|T^6) = E(X|T^4)$ we have established cycling between exactly two alternatives. Otherwise, it has to be the case that the following five equations have to be fulfilled

$$\begin{aligned}
k(T^2)\left(E(X|T^2) - E(X|T^1)\right)^2 &> k(T^5)\left(E(X|T^5) - E(X|T^1)\right)^2 \\
k(T^3)\left(E(X|T^3) - E(X|T^2)\right)^2 &> k(T^6)\left(E(X|T^6) - E(X|T^2)\right)^2 \\
k(T^4)\left(E(X|T^4) - E(X|T^3)\right)^2 &> k(T^2)\left(E(X|T^2) - E(X|T^3)\right)^2 \\
k(T^5)\left(E(X|T^5) - E(X|T^4)\right)^2 &> k(T^3)\left(E(X|T^3) - E(X|T^4)\right)^2 \\
k(T^6)\left(E(X|T^6) - E(X|T^5)\right)^2 &> k(T^4)\left(E(X|T^4) - E(X|T^5)\right)^2
\end{aligned}$$

For these equations to hold it must be that $E(X|T^1) < E(X|T^6) < E(X|T^5)$. Further, it cannot be the case that $E(X|T^4) < E(X|T^6)$, by Lemma 10 **C.1**. It therefore must be the case that $E(X|T^4) > E(X|T^6) > E(X|T^1)$. We then have the following order,

$$E(X|T^1) < E(X|T^6) < E(X|T^4) < E(X|T^2) < E(X|T^5) < E(X|T^3) \tag{36}$$

Then, continuing this reasoning it must be $E(X|T^2) < E(X|T^7) \leq E(X|T^5)$, $E(X|T^1) < E(X|T^8) \leq E(X|T^6)$ etc. This implies that there are two bounded sets, with the policy chosen in odd periods moving closer to $E(X|T^1)$ and the policy chosen in even periods moving closer to $E(X|T^2)$, that is the bliss points of the optimal target sets move to the left over time. It cannot be that one of the bliss points moves to the right, as $E(X|T^6) < E(X|T^5)$. This implies that for any target set with bliss point $E(X|T^t) > E(X|T^2)$, $E(X|T^{t+1}) < E(X|T^t)$. Eventually, it must be the case that either $E(X|T^t) = E(X|T^{t+2})$, or that the last possible target set to the right of one of the boundaries, that is of $E(X|T^2)$ or $E(X|T^1)$ has been reached, which implies cycling between two alternatives.

Suppose next that $E(X|T^1) < E(X|T^2)$, but $E(X|T^3) < E(X|T^2)$. We need to distinguish between two cases, namely, $E(X|T^3) < E(X|T^1)$ and $E(X|T^1) < E(X|T^3) < E(X|T^2)$.

Step 1 We first let $E(X|T^3) < E(X|T^1)$. We again ask where $E(X|T^4)$ can lie. If $E(X|T^4) = E(X|T^2)$, the proof is completed. We therefore proceed to the case of $E(X|T^4) \neq E(X|T^2)$ By Lemma 10 **C.1**, it must be that $E(X|T^4) < E(X|T^2)$. We can further show that $E(X|T^1) < E(X|T^4)$. The following equations have to be fulfilled:

$$\begin{aligned}
&k(T^2)\left(E(X|T^2) - E(X|T^1)\right)^2 > k(T^3)\left(E(X|T^3) - E(X|T^1)\right)^2, \\
&k(T^3)\left(E(X|T^3) - E(X|T^2)\right)^2 > k(T^4)\left(E(X|T^4) - E(X|T^2)\right)^2, \\
&k(T^4)\left(E(X|T^4) - E(X|T^3)\right)^2 > k(T^2)\left(E(X|T^2) - E(X|T^3)\right)^2,
\end{aligned}$$

and they hold only if $E(X|T^1) < E(X|T^4)$. Therefore, it must be that

$$E(X|T^1) < E(X|T^4) < E(X|T^2),$$

which results in the following overall ordering,

$$E(X|T^3) < E(X|T^1) < E(X|T^4) < E(X|T^2) \tag{37}$$

We then turn to $E(X|T^5)$. If $E(X|T^5) = E(X|T^3)$, we have established that there is cycling between two policy platforms. Therefore, let $E(X|T^5) \neq E(X|T^3)$. Then there are two possible cases, namely (i) $E(X|T^5) < E(X|T^4)$ and (ii) $E(X|T^5) > E(X|T^4)$.

(i)$E(X|T^5) < E(X|T^4)$**:** By Lemma 10 **C.1**, it must be that $E(X|T^5) < E(X|T^3)$. Continuing this reasoning, it must be that $E(X|T^6) \leq E(X|T^4)$. This implies that both platforms gradually move left. As $0 \leq E(X|T)$, at some point we must reach cycling between two alternatives.

(ii) $E(X|T^5) > E(X|T^4)$**:** By Lemma 10 **C.1**, $E(X|T^5) > E(X|T^2)$, which yields the following ordering,

$$E(X|T^3) < E(X|T^1) < E(X|T^4) < E(X|T^2) < E(X|T^5)$$

We then turn to $E(X|T^6)$, assuming that $E(X|T^6) \neq E(X|T^4)$. It cannot be the case that

$E(X|T^6) > E(X|T^5)$, by Lemma 10 **C.3**. Then, it has to be the case that $E(X|T^6) < E(X|T^5)$. By Lemma 10 **C.2** it must hold that $E(X|T^6) > E(X|T^3)$. Further, it must be the case that $E(X|T^6) < E(X|T^4)$ as for $E(X|T^6) > E(X|T^4)$ the following two equations are never fulfilled:

$$k(T^4)\left(E(X|T^4) - E(X|T^3)\right)^2 > k(T^6)\left(E(X|T^6 - E(X|T^3)\right)^2$$
$$k(T^6)\left(E(X|T^6 - E(X|T^5)\right)^2 > k(T^4)\left(E(X|T^4) - E(X|T^5)\right)^2,$$

Last, we show that $E(X|T^1) < E(X|T^6) < E(X|T^4)$. It must be the case that

$$k(T^2)\left(E(X|T^2) - E(X|T^1)\right)^2 > k(T^5)\left(E(X|T^5) - E(X|T^1)\right)^2$$
$$k(T^3)\left(E(X|T^3) - E(X|T^2)\right)^2 > k(T^6)\left(E(X|T^6) - E(X|T^2)\right)^2$$
$$k(T^4)\left(E(X|T^4) - E(X|T^3)\right)^2 > k(T^2)\left(E(X|T^2) - E(X|T^3)\right)^2$$
$$k(T^5)\left(E(X|T^5) - E(X|T^4)\right)^2 > k(T^3)\left(E(X|T^3) - E(X|T^4)\right)^2$$
$$k(T^6)\left(E(X|T^6) - E(X|T^5)\right)^2 > k(T^4)\left(E(X|T^4) - E(X|T^5)\right)^2$$

These equations only hold for $E(X|T^1) < E(X|T^6)$ and thus it must be

$$E(X|T^3) < E(X|T^1) < E(X|T^6) < E(X|T^4) < E(X|T^2) < E(X|T^5)$$

Based on this, the policy platform on the left $E(X|T_L) \in (E(X|T^1), E(X|T^4))$ and $E(X|T_R) \in (E(X|T^2), E(X|T^5))$ and $E(X|T^{t+2}) \leq E(X|T^t)$. As the set is bounded, eventually cycling between two platforms must be reached.

<u>*Step 2*</u> We now turn to the case where $E(X|T^1) < E(X|T^3) < E(X|T^2)$. Note that if $E(X|T^4) < E(X|T^3) < E(X|T^2)$, the setting is as in the previous case with $E(X|T^1) < E(X|T^2) < E(X|T^3)$ and we have shown that this results in cycling between two alternatives. We therefore restrict attention to the case where $E(X|T^4) > E(X|T^3)$. By Lemma 10 **C.1**, it must be the case that $E(X|T^4) > E(X|T^2)$ and thus $E(X|T^1) < E(X|T^3) < E(X|T^2) < E(X|T^4)$. By Lemma 10 **C.2**, $E(X|T^5) \leq E(X|T^3)$ and by Lemma 10 **C.3**, $E(X|T^5) \leq E(X|T^4)$. Additionally, $E(X|T^2) > E(X|T^5)$ as the following three equations must be

fulfilled

$$k(T^3)\left(E(X|T^3) - E(X|T^2)\right)^2 > k(T^4)\left(E(X|T^4) - E(X|T^2)\right)^2,$$
$$k(T^4)\left(E(X|T^4) - E(X|T^3)\right)^2 > k(T^5)\left(E(X|T^5) - E(X|T^3)\right)^2,$$
$$k(T^5)\left(E(X|T^5) - E(X|T^4)\right)^2 > k(T^3)\left(E(X|T^3) - E(X|T^4)\right)^2.$$

They only hold for $E(X|T^2) > E(X|T^5)$ which yields

$$E(X|T^1) < E(X|T^3) < E(X|T^5) < E(X|T^2) < E(X|T^4)$$

Again, there are two possibilities. It is either the case that $E(X|T^6) > E(X|T^5)$. By Lemma 10 **C.1**, $E(X|T^6) > E(X|T^4)$. It then must also be that $E(X|T^7) \geq E(X|T^5)$. This implies that both platforms gradually move right. As $1 \geq E(X|T)$, at some point we must reach cycling between two alternatives. Alternatively, it can be the case that at some point it becomes better to switch to a platform on the left again, namely $E(X|T^t) > E(X|T^{t+1}) > E(X|T^{t+2})$, for which we have shown that there has to be cycling between exactly two platforms.

School of Economics and Finance

Queen Mary
University of London

This working paper has been produced by
the School of Economics and Finance at
Queen Mary University of London

School of Economics and Finance
Queen Mary University of London
Mile End Road
London E1 4NS
Tel: +44 (0)20 7882 7356
Fax: +44 (0)20 8983 3580
Web: www.econ.qmul.ac.uk/research/workingpapers/