Some for the Price of One: Targeted Redistribution and Social Structure

Perry Carter Joseph J. Ruggiero

October 17, 2023

Abstract

Social networks play an important role in distributive politics, yet little is known about how features of societies affect patterns of redistribution. We study a formal model of an election in which candidates may offer excludable transfers to policy-motivated voters connected on a social network. Our model facilitates comparison of societies based on their average features. In equilibrium, transfers are determined by group shares and homophily, while inequalities between groups are driven by a disproportionate targeting of minorities. We also consider heterogeneous information between candidates, clarifying the informational benefits of density and demonstrating that homophily can endogenously produce in-group favoritism. These results highlight the importance of aggregate social structure and suggest new directions for empirical studies of redistributive politics.

Carter: Ph.D. Candidate in Politics, Princeton University. pjcarter@princeton.edu.
Ruggiero: Ph.D. Candidate in Politics, Princeton University. jruggiero@princeton.edu.

Elected officials face a general tradeoff between the unconditional provision of public goods to the entire electorate and targeted redistribution. Untargeted, rule-based provision is typically more consistent with democratic norms and enjoys widespread support among voters (Kitschelt 2000; Vicente and Wantchekon 2009). At the same time, targeted allocation of resources—through illicit practices or more routine pork-barrel politics—can be electorally optimal for politicians despite voter preferences (Groseclose and Snyder 1996; Dixit and Londregan 1996; Banks 2000; Lizzeri and Persico 2001; Persson and Tabellini 2002; Stokes 2005; Dekel, Jackson, and Wolinsky 2008).

However, voter decisions do not only depend on material incentives: social connections also influence choice (Beck, Dalton, Greene, and Huckfeldt 2002; Rolfe 2012; Fafchamps, Vaz, and Vicente 2020; Eubank, Grossman, Platas, Rodden et al. 2021). Social structure—the propensity of social and political groups to adopt particular patterns of interrelationship (Granovetter 2005)—therefore modulates the form and efficacy of targeted redistributive strategies (Stokes, Dunning, Nazareno, and Brusco 2013; Gans-Morse, Mazzuca, and Nichter 2014; Holland and Palmer-Rubin 2015; Cruz, Labonne, and Querubin 2017). Despite the importance of social structure, existing work has focused exclusively on the first-order properties (i.e., degree) of the targeted voters without considering the role of the network as a whole (Finan and Schechter 2012; Cruz 2019). If socially influential voters make the most attractive targets for co-optation, then, the question of which *societies* are more or less favorable to certain redistributive strategies remains unanswered.

In this article, we study a networked model of a large election in which candidates compete to influence policy-motivated voters with targeted and excludable transfers. Each candidate is associated with a policy and can extend private transfers to voters at the expense of a public good, which both candidates and voters care about. Voters belong to a group with distinct preferences and care about the welfare of their neighbors on the network, so that voting decisions are also informed by the electoral preferences of their social connections. Our model can therefore be thought of as an extension of the canonical probabilistic voting model

(Persson and Tabellini 2002) that allows us to study the role played by social transmission of preferences in distributive politics.

The unique equilibrium transfer to a voter is proportional to their network centrality, which captures a core tradeoff between targeting voters who influence others and those who are easily influenced. The sensitivity of any agent's centrality to small changes on the network, however, complicates direct analysis of the role played by a society's underlying structural features. To overcome this problem, we eschew consideration of exact social network realizations in favor of an explicit generative model (specifically, a stochastic block model). By applying techniques from random graph analysis (Chung and Radcliffe 2011) to a class of games on networks for which equilibrium strategies are proportional to the vector of Katz-Bonacich centralities (Katz 1953; Bonacich 1987; Ballester, Calvó-Armengol, and Zenou 2006),[1] we can derive closed-form expressions for *expected* equilibrium strategies across repeated draws from a generative model. In this way, fixed characteristics of society that govern network formation are directly connected to equilibrium behavior.

Although approximations, these strategies are arbitrarily close to those played on any realization in large societies. Since these expressions depend on the parameters that govern network formation, and not on the arbitrary structure of a specific realization, they allow us to study the effects of deep features of societies on the efficacy of targeted redistribution directly. A core advantage of this approach is that it allows systematic comparison across societies (either cross-sectionally or over time) based on stable underlying features of the social environment, without needing to consider a complete network in each case of interest. We focus on three network-level attributes that research on social networks has consistently highlighted:

1. density, or the ratio of realized to potential ties;

2. fractionalization, or the number and relative size of salient social groups;[2] and

[1] This approach has provided a good empirical fit when compared to alternative measures of centrality (Calvó-Armengol, Patacchini, and Zenou 2009; Battaglini and Patacchini 2018).

[2] See Alesina, Devleeschauwer, Easterly, Kurlat, and Wacziarg (2003) for a discussion of appropriate ways

3. homophily, or the relative propensity of agents to form ties with members of their own group.

Our model suggests that diversion of public resources will be most prevalent when a relatively small minority can be targeted and when candidates are well-informed about voter preferences. These predictions, while consistent with the empirical literature's emphasis on the targeting of ethnic minorities by well-informed and socially embedded brokers (Stokes et al. 2013), highlight important channels that have been largely overlooked. In our model, fractionalization matters not because of differences in preferences (Easterly and Levine 1997; Alesina, Baqir, and Easterly 1999) or between-group inequities (Baldwin and Huber 2010), but because it affects the proportion of ties that are within or between groups and, thus, the level of social pressure that can be achieved. Similarly, although the importance of information about individual voters for contract enforcement is well-studied (Finan and Schechter 2012; Stokes et al. 2013), we highlight the importance of information about *neighborhoods* of voters for redistributive outcomes.

Social segregation between groups, however, typically reduces the value of transfers. Since minorities are disproportionately targeted, the incentive to target individuals is moderated as members of minorities become less connected to others on the network. This is consistent with previous work showing that segregation does not lead to in-group favoritism (Franck and Rainer 2012) and suggests a countervailing effect that may partly explain the contradictory findings on the effects of group diversity (Habyarimana, Humphreys, Posner, and Weinstein 2007; Baldwin and Huber 2010).

Finally, when the quality of information is allowed to vary across candidates, we find that spending is increasing in fractionalization when the candidate has low-precision information on minorities and is typically decreasing in homophily. Inequality, on the other hand, tends to decrease in fractionalization and has a nonlinear relationship to homophily.

to conceptualize fractionalization.

Targeted Redistribution and Social Influence

While previous work on targeted redistribution has considered the role of network position (degree or centrality) of individual voters (Cruz 2019; Ravanilla, Davidson Jr, and Hicken 2022; Ravanilla and Hicken 2023), our findings underscore the importance of attending to less obvious network features that may have counterintuitive effects. An important implication of our approach of focusing on *average* features of social networks for applied research is that it does not require time and resource-intensive measurement of complete networks (Breza, Chandrasekhar, McCormick, and Pan 2020). Instead, all of the relevant properties considered in our model—density, fractionalization, and homophily—can be inferred from the average propensity of individuals to form ties with one another.

Further, formal models of targeted redistribution have not considered the role of connections between players—most assume continuous distributions of voters, for which results do not necessarily generalize to finite populations (Groseclose and Snyder 1996; Banks 2000; Dekel, Jackson, and Wolinsky 2008; Lizzeri and Persico 2001). This literature has generated important insights into the possibility of vote buying to induce inefficient supermajority coalitions (Groseclose and Snyder 1996; Banks 2000), the difficulty of overcoming private incentives by providing public goods (Lizzeri and Persico 2001), the role of varying commitment structures and institutions in mitigating inefficiencies in redistribution (Dal Bó 2007; Dekel, Jackson, and Wolinsky 2008), and the institutional factors driving the mix of strategies chosen by clientelist machines (Gans-Morse, Mazzuca, and Nichter 2014). However, variation is driven either by individual factors or by institutional environments. Despite the prominent role afforded to ties between actors in empirical accounts, this aspect of the strategic environment has gone largely unexplored.

Two important exceptions comes from Battaglini and Patacchini (2018) and Dixit and Londregan (1996). Battaglini and Patacchini (2018) study the problem of influencing members of a legislature through campaign contributions using a networked model. The authors find, following Ballester, Calvó-Armengol, and Zenou (2006), that the equilibrium transfers to voters (legislators) are proportional to their Katz-Bonacich centrality (Bonacich 1987) weighted by the equilibrium probability of pivotality. Although they employ a similar modeling technology, their paper focuses on the role of pivotal voting in small legislative elections, while we mainly consider the effects of network-level changes in large general elections. By shifting attention from realized networks to an underlying generative model, we can draw sharp conclusions about the effects of social structure.

In their canonical treatment of targeted redistribution, Dixit and Londregan (1996) consider the role of within-group heterogeneity in interest group preferences in determining party strategies. The paper concludes that the targeting of core or swing voters is determined by the relative efficiency of each party at providing goods to its own supporters. While the authors consider the consequences of voters' social embeddedness, however, they operationalize this factor solely in terms of group membership, without addressing the patterns of *interactions* within and between groups. Our model can thus be seen as extending these insights, introducing social interaction among voters as an additional source of heterogeneity.

Finally, this paper introduces a useful technical innovation that opens up new avenues for the study of social networks in political science. By extending several key results in the spectral theory of random graphs (Chung and Radcliffe 2011), we provide a general framework to study the impact of average social network properties on political outcomes. While we focus in the present paper on the case of targeted redistribution, social network structure is important in many areas of political science, and our approach can easily be extended to other applications.

Model

In this section, we lay out and justify the basic structure of the model, which builds on the framework of probabilistic voting with spillovers developed in Battaglini and Patacchini

(2018) in the context of campaign contributions to legislators. While the basic structure of our model is similar, we introduce flexibility in the structure of social influence, as well as a core trade-off between public and private provision, which facilitates our focus on the role of social structure in large elections. For a more in-depth discussion of the underlying assumptions and their consequences for our findings, refer to the section on Discussion of Assumptions.

We begin by assuming that voters care about policy in a unidimensional space and the provision of a public good, but that they can also be influenced with targeted transfers that alter their likelihood of voting for one candidate over another. Since we are primarily interested in elections where n is sufficiently large that the probability of pivotality is approximately zero, we assume expressive voting based on net preference after transfers. In order to retain our focus on the network-specific elements of the model, we treat transfers from both candidates as credible campaign promises (Dekel, Jackson, and Wolinsky 2008).[3] These offers, therefore, are not a binding contract from the perspective of the voter: voters may receive promises from both candidates and will ultimately vote in accordance with their own preferences, net of the offers made. As such, candidates do not receive a vote with certainty, but only an improved likelihood of support. This setting is in line with existing findings highlighting the importance of voter reciprocity norms and repeated interactions as a means of contract enforcement (Rueda 2017; Cruz 2019).

The key feature of this model is network dependence. In addition to being influenced by direct transfers and campaign promises, voters place some weight on their neighbors' expected votes. This approach to modeling dependency, which typically implies that agents take actions in proportion to their centrality (Ballester, Calvó-Armengol, and Zenou 2006), has been employed in a variety of applications capturing peer effects and found strong empirical support (Battaglini and Patacchini 2018; Fafchamps and Labonne 2020). While our

[3] For instance, candidates may offer voters attractive but unnecessary jobs at a local school for which they are inadequately qualified, potentially reducing the quality of education in the process. Alternatively, they may direct the same resources toward investing in improving the school itself, benefiting the entire community.

primary interest is in the role played by aggregate network properties, in this section we take the network as fixed in order to characterize equilibrium strategies conditional on the network. In the following section, we shift our focus to social structure, assuming that the network is generated according to a stochastic block model and studying the effect of changes in its parameters.

Substantively, the network spillover mechanism has two main interpretations. First, voters can be thought of as communicating with their acquaintances about their intent to vote, which provides information about the candidate's desirability, which in turn leads to a preference for voting for the same candidate. A transfer made to any voter on the network will positively affect *all* voters' likelihood of voting for a candidate on a connected graph, albeit with diminishing returns in social distance. Second, network spillovers may be a consequence of social pressure. Even if voters do not gain any payoff relevant information from their neighbors, they may still be intrinsically motivated to take the same action as a majority of them.

While more-connected voters are exposed to greater influence from the whole network, we normalize the total social influence (i.e., the sum of edge weights) of any voter's *immediate* connections to one so that all voters place equal weight on their neighbors' vote probabilities.[4] This implies that voters can be thought of as making their decisions based on a weighted average of their neighbors' actions, and not a sum. This aspect of the model diverges somewhat from others in the literature who assume that the most connected voters are also the most easily influenced, and is motivated by empirical evidence that the effect of peer pressure is similar at all levels of connectedness (Green and Gerber 2010; Lazer, Rubineau, Chetkovich, Katz, and Neblo 2010; Jang, Lee, and Park 2014).

Both voters and candidates care about the latter's programmatic commitment to provide a public good, but candidates face a potential trade-off between these two goals. The funds for private transfers to voters, which can provide an electoral advantage, must be diverted

[4] See Appendix A.5 for a discussion of the consequences of relaxing this assumption.

from the provision of the public good. Our model is thus designed to capture the basic trade-off between public and private provision that has been a focus of much of the empirical literature (Cruz, Labonne, and Querubin 2020). Crucially, public good provision may *also* benefit from the structure of the network since all voters are equally affected by positive spillovers. The effectiveness of targeted distribution relies on the ability of candidates to exploit the differential benefits of swaying *some* voters relative to the lost utility this induces for all others. At the same time, candidates must also avoid triggering *negative* spillovers arising from the dissatisfaction of influential voters with diversion of funds from public good provision.

Finally, we assume that candidates attempt to maximize vote share rather than to win a majority. This assumption is empirically appropriate in many cases, such as in authoritarian regimes, where incumbents frequently seek to achieve overwhelming vote shares (Reuter and Robertson 2012), or under proportional representation where votes translate directly into influence. It is also worth noting that, while we use the term "candidate," it may be more appropriate in many applications to view the agents as vote brokers, since they are assumed to have accurate knowledge of the local social network structure but are uncertain of voters' final decisions.

Setup

Consider a game with n voters that need to make a choice between two candidates. All voters are located on a network $\mathcal{G}$, which is assumed to be connected.[5] We use the terms network and graph interchangeably throughout the paper to refer to an undirected graph, which is an ordered pair $(\mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ is a set of n vertices and $\mathcal{E}$ is a set of m edges such that $\mathcal{E} \subseteq \{\{x, x'\} : x, x' \in \mathcal{V} \wedge x \neq x'\}$.[6] Each candidate $k = 1, 2$ is associated with a policy $y_k = k$, where the policy space is a subset of $\mathbb{R}$, and each voter $i \in \mathcal{V}$ is endowed with a group

[5] A connected undirected graph is one in which a path exists between any two vertices.

[6] See Jackson (2010) for a general primer on graph theory as it relates to games on social networks.

membership $\ell_i = 1, 2$, which corresponds to an ideal policy $x_i = \ell_i$. Substantively, groups may be interpreted as corresponding to any grouping that is both socially and politically meaningful, such as political parties and ethnic or religious groups.

To gain vote share, candidate k can extend n private offers or *bribes*, $b_{ik} \geq 0$. These bribes, however, come at the expense of a public good, which the candidate also values. A candidate k's problem is to choose $\boldsymbol{b}_k \in \mathbb{R}^n_+$ that solves

$$\begin{aligned} \max_{\boldsymbol{b}_k} \quad & \alpha_k \sum_{i \in \mathcal{V}} \phi_{ik}(\boldsymbol{b}_k, \boldsymbol{b}_{-k}) - \boldsymbol{b}_k \cdot \mathbf{1} \\ \text{subject to} \quad & b_{ik} \geq 0 \text{ for all } i \end{aligned}$$

where $\phi_{ik}(\cdot)$ is the probability voter i votes for candidate k and α_k represents the value placed on one vote by candidate k. We thus normalize the value placed on a unit of public good by the candidate to 1 so that α_k can be interpreted as the candidate's relative degree of office motivation. In particular, an α_k of 0 corresponds to a fully programmatic candidate, who trivially prefers to offer no private transfers and promise the full amount of the public good, while $\alpha_k \to \infty$ implies pure office motivation.[7]

Voters support the candidate that offers them a higher total utility.[8] All voters care about policy according to a standard quadratic loss function and have $\gamma \geq 0$ value for a unit of public good, so that they incur a loss of γ for every unit of bribes offered by a candidate to any voter[9]. Additionally, voters have private information unknown to the candidates and other voters in the form of a private valence shock for each candidate, $\varepsilon_{ik} \in \mathbb{R}$. Without loss of generality, we can normalize $\varepsilon_{i2} = 0$ and define $\varepsilon_i := \varepsilon_{i1}$, which we assume is an

[7] We focus on candidates that can perfectly target individual voters in the main text. Candidates that assign benefits on the basis of a policy rule, as in the case of ethnic favoritism, where goods are provided based on observable group membership (Chandra 2007), is considered in Appendix A.3.

[8] There is no obligation to vote for a candidate who offered them a bribe.

[9] A consequence of this modelling choice is that voters' utility depends on the sum of bribes offered to all other voters. However, this need not imply that voters actually observe the bribes offered to anyone else: they need only be able to infer with reasonable accuracy the total extent of diversion of public resources.

independent, uniformly[10] distributed mean-zero random variable with support on $\left[\frac{1}{-2\theta}, \frac{1}{2\theta}\right]$, where we impose that the parameter $0 < \theta < 1$ to ensure that all solutions are well-defined.[11] We interpret θ as the candidates' information about the utility of transfers to voters, with smaller θ indicating less-informed candidates. θ can also be taken as reflecting the intensity of the voters' commitment problem, as candidates with higher values can be more certain that transfers will actually secure votes.[12]

Social connections also matter. In particular, voters prefer to vote for the same candidate as their neighbors as defined by the network. Denote by $\phi_{ik}(\cdot)$ the probability voter i votes for candidate k given all bribes, but before the realization of the valence shock ε_i. Then, each voter i places weight $w_{ij} > 0$ on voter j's probability of voting for candidate k if i and j are connected, and 0 otherwise. In the graph $\mathcal{G}$, the set of a voter i's social ties is denoted by $\mathcal{T}_i(\mathcal{G}) \subseteq \mathcal{V}$.[13] The total social influence on each voter is normalized to 1, so that the actual influence of each neighbor j on i's utility is equal to $\left(\sum_{h\in\mathcal{T}_i(\mathcal{G})} w_{ih}\right)^{-1} w_{ij}$, implying that more highly connected voters are less influenced by each individual neighbor.[14] An immediate consequence is that network density only plays an indirect role on equilibrium behavior, as adding edges does not increase the total influence on the network. In much of the subsequent analysis, we will assume for simplicity that $w_{ij} \in \{w_L, w_H\}$ with $w_H \geq w_L$, where w_H is the weight placed on within-group connections and w_L on cross-group connections. All fixed-network results hold for the more general case, however.

The expected payoff voter i receives from candidate k can thus be expressed as

$$U_i(k) = -(x_i - y_k)^2 + u(b_{ik}) + \frac{\sum_{j\in\mathcal{T}_i(\mathcal{G})} w_{ij}\phi_{jk}(\boldsymbol{b})}{\sum_{h\in\mathcal{T}_i(\mathcal{G})} w_{ih}} - \gamma \sum_{m\in\mathcal{V}} b_{mk} + \varepsilon_{ik} \tag{1}$$

[10] The assumption of uniformly distributed shocks is done to simplify the calculations and recover closed-form expressions; however, the core mechanisms of the paper do not fundamentally depend on this.

[11] See Appendix A for a more precise statement.

[12] Equivalently, θ can be taken as a reflection of voters' reciprocity norms in the vote-buying context, as it modulates the likelihood of any given transfer actually resulting in a vote. We explore the consequences of systematic variation in this parameter in the Heterogeneous Information section.

[13] Here, $\mathcal{G}$ is a fixed graph and may have any structure provided that it is connected.

[14] See Appendix A.5 for an analysis of the consequences of relaxing this assumption.

where $u(\cdot)$ is voter utility over bribes, which we assume is strictly increasing with diminishing marginal returns and that the rate of diminution is decreasing, with infinite marginal utility as bribes approach zero. Formally, utility over bribes satisfies $u'(\cdot) > 0$, $u''(\cdot) < 0$, and $u'''(\cdot) \geq 0$. Additionally, we assume $\lim_{b \to 0} u'(b) = \infty$ and $\lim_{b \to \infty} u'(b) = 0$. Together, these assumptions ensure that all solution objects are well-defined and rule out the possibility of corner solutions.

Timing

The timing of the game is as follows.

1. Nature randomly chooses a private utility shock for each voter, $\varepsilon_i \sim \mathcal{U}\left[\frac{-1}{2\theta}, \frac{1}{2\theta}\right]$
2. For all voters $i \in \mathcal{V}$, each candidate $k = 1, 2$ offers a bribe $b_{ik} \geq 0$, which determines the residual public good offered
3. Each voter $i \in \mathcal{V}$ casts a ballot for candidate 1, $v_i = 1$, or candidate 2, $v_i = 2$
4. The winning candidate enacts their promised transfer program, and voters' utility is realized.

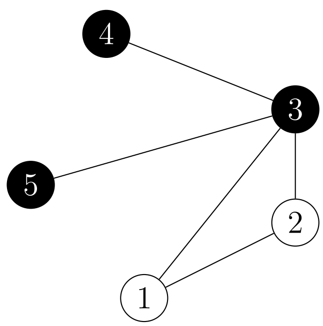
Equilibrium

A voter will cast a ballot for candidate 1 if and only if $U_i(1) \geq U_i(2)$. Here, candidates will not be able to perfectly anticipate voting behavior due to their imperfect information over voter preferences. We can summarize the[15] the candidates' first-order conditions as

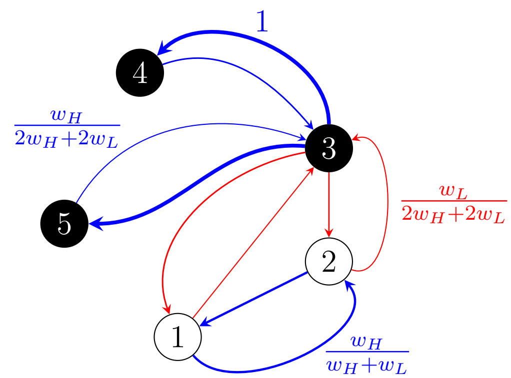
$$\left(\boldsymbol{J}\left[\boldsymbol{u}\right] - \boldsymbol{\Gamma}\right)^{\top} \cdot \left(\boldsymbol{I} - 2\theta\tilde{\boldsymbol{A}}\right)^{-1} \cdot \mathbf{1} = \frac{(1 - \boldsymbol{\lambda})}{\alpha_k \theta} \tag{2}$$

where $\boldsymbol{J}[\cdot]$ is a diagonal matrix with $u'(b_i)$ as the nonzero entries, $\boldsymbol{\Gamma}$ is an $n \times n$ square matrix such that every element of $\boldsymbol{\Gamma}$ is γ, $\boldsymbol{I}$ denotes the identity matrix, $\mathbf{1}$ denotes an n-vector of

[15] See Appendix A for derivation.



(a) Example of a realized graph



(b) Induced weighted directed graph

Figure 1: An example of a realized (undirected) network with $n = 5$ and group labels $\ell_1 = \ell_2 = 1$ and $\ell_3 = \ell_4 = \ell_5 = 2$ and the corresponding induced weighted directed network. In the weighted directed graph, thicker arrows indicate stronger influence, with $w_{ij} = w_H$ if $\ell_i = \ell_j$, $w_{ij} = w_L$ otherwise, and naturally $w_H > w_L$.

1s, $\boldsymbol{\lambda}$ is an n-vector of Lagrange multipliers, and $\tilde{\boldsymbol{A}}$ is a normalized weighted adjacency matrix. Note that by assumption on θ being sufficiently small, $(\boldsymbol{I} - 2\theta\tilde{\boldsymbol{A}})$ is guaranteed to be invertible, so the problem is well-defined.

Definition 1. *Consider a realized graph $\mathcal{G}$ and a corresponding adjacency matrix $\boldsymbol{A}$ such that for all $i, j \in \mathcal{V}$, $A_{ij} = 1$ if $j \in \mathcal{T}_i(\mathcal{G})$ and $A_{ij} = 0$ otherwise. Then, the* ***normalized weighted adjacency matrix*** $\tilde{\boldsymbol{A}}$ *is given by, for all $i, j \in \mathcal{V}$,*[16]

$$\tilde{A}_{ij} = \frac{w_{ij} A_{ij}}{\sum_{m \in \mathcal{V}} w_{im} A_{im}}.$$

By employing the normalized weighted adjacency matrix, we can account for several important features of social interaction. First, a voter i may be more influenced by one social tie than another. Second, it will be more difficult for any one person to influence a highly connected voter than a relatively disconnected one, e.g., an incremental change in the probability that i's neighbor j votes for candidate 1 will have less of an effect on i's vote probability if i has hundreds of neighbors than if j is i's only neighbor.

[16] Note that since all weights are strictly positive, the assumption that $\mathcal{G}$ is connected implies that the normalized weighted graph is strongly connected.

From the candidates' problem in equation (6), we can recover the equilibrium transfer that voter i receives from candidate k,

$$b_{ik} = [u']^{-1} \left(\frac{1}{c_i(\boldsymbol{w}, \theta; \mathcal{G})} \left[\gamma C(\boldsymbol{w}, \theta; \mathcal{G}) + \frac{1}{\alpha_k \theta} \right] \right), \tag{3}$$

where $c_i(\cdot)$ is the ith element of $\boldsymbol{c} = (\boldsymbol{I} - 2\theta \tilde{\boldsymbol{A}})^{-1} \cdot \mathbf{1}$, our measure of centrality, and $C(\cdot) \equiv \sum_{i \in \mathcal{V}} c_i(\cdot)$. The measure $\boldsymbol{c}$ is the Katz-Bonacich centrality on the weighted directed network corresponding to $\tilde{\boldsymbol{A}}$ with attenuation parameter 2θ. The nature of the strategic environment—specifically, the structure of social influence—can therefore be thought of as inducing a latent directed network with connections corresponding to the influence of i on j, which is decreasing in j's weighted degree and greatest when $\ell_i = \ell_j$. The value of a voter to a candidate is thus proportional to their centrality on this latent network, which captures the weighted sum of directed walks of any length that include that voter.

By taking derivatives of the equilibrium bribes defined by equation (3), it is clear that equilibrium transfers offered by candidate k are weakly decreasing in γ and n, while they are weakly increasing in α_k and θ. These results are intuitively consistent with the basic strategic environment: the socially optimal transfers to voters would correspond to a transfer scheme such that the marginal value is equated with γ, weighted by the total centrality on the network, which can be taken as a measure of the additional positive spillover candidates gain from providing a public good. Network spillovers also incentivize candidates to provide additional transfers beyond this level, however. The network induces a tradeoff between the positive effect associated with providing a transfer to i, which are captured by c_i and the concomitant negative effect this induces via *every other voter* being deprived of the public good, captured by C. The relative value of these spillovers is moderated by α_k—that is, more office-motivated candidates place higher value on the additional increase to expected vote share afforded by targeting high-centrality voters—and by θ, which determines the likelihood of their a realized increase in vote share. Finally, it is not necessarily true that an increase in the number of voters results in a decrease in transfers, since it may in general be possible to add another voter in such a way that either or both of (i) centrality increases for some i due to the creation of new paths (ii) C decreases due to the reduction of weights on existing paths.

It is also of note that, if both candidates have identical degrees of office motivation (i.e., $\alpha_1 = \alpha_2 = \alpha$), then they will also choose the same transfer profile in equilibrium. This can be thought of as an analogue to the Median Voter Theorem for spatial models of competition (Downs 1957); it is optimal for both candidates to respond by extending offers to the most valuable voters, who are determined completely by their network positions. A further implication is then that targeted distribution only influences aggregate electoral outcomes if candidates diverge in their motivations or resources. When candidates are perfectly symmetric, their offers perfectly offset one another, so that voting decisions are determined entirely by *ex ante* preferences, which is consistent with findings in the empirical literature (Vicente and Wantchekon 2009).

In this setting, it is straightforward to see why this might not hold true: for instance, if voters respond differently to offers from candidates belonging to their own group, then incentives for in-group favoritism will exist. In a later section of the paper, we will take up the role played by heterogeneous information in this regard, while Appendix C considers the case when information and network structure are correlated.

To study the dependence of equilibrium strategy on social structure, we now transition to considering the underlying generating model that gave rise to the observed network.

Social Structure

This section studies the role of social structure and provides the main results of this paper. In particular, we employ tools from random graph theory to derive closed-form expressions for the centrality of voters in each group, yielding results in terms of the main features of social structure: group fractionalization, density, and homophily. These techniques[17] allow us to consider centrality on the average graph only, permitting analysis of comparative statics explicitly in terms of social structure—that is, the underlying parameters that govern the social network generative process—rather than of a single realized graph.

While the resulting statements relate to expectations and therefore do not *necessarily* apply to specific network realizations, Theorem 1 guarantees that they *will* hold with probability approaching one in any large network under moderate assumptions.[18] These results, therefore, allow us to make empirically applicable predictions about how large societies behave that permit systematic comparison across space and time based on underlying commonalities, without needing to observe the complete centrality vector for every case of interest.

While a potential concern is that the observed network's properties have a higher probability of deviating dramatically from its expectation in small, isolated communities of only a few hundred residents, this issue is mitigated by empirical work that has shown targeted distributional strategies are also prevalent in urban and semi-urban environments where our asymptotic approximations are guaranteed to hold, such as major cities in Argentina (Brusco, Nazareno, and Stokes 2004; Stokes 2005).

We now formalize the concept of an average network in the context of our model, which can be conveniently represented through its average adjacency matrix.

Definition 2. *The* ***average normalized weighted adjacency matrix*** $\bar{\tilde{\mathbf{A}}}$ *is given by, for all* $i, j \in \mathcal{V}$,

$$\bar{\tilde{A}}_{ij} = \frac{w_{ij} p_{ij}}{\sum_{m \in \mathcal{V}} w_{im} p_{im}}.$$

where w_{ij} *is the weight placed on the connection between* i *and* j *and* p_{ij} *is the probability of that connection being realized.*

[17] These are technical in nature and therefore reserved for Appendix D.

[18] Specifically, we require that the minimum expected degree grows at a rate greater than $\ln(n)$, an assumption generally borne out empirically (Eubank, Kumar, Marathe, Srinivasan, and Wang 2004).

For the following results, we assume that the graph is drawn according to a two-group stochastic block model with share $s \geq \frac{1}{2}$ of group 1, a probability p_H of intra-group connection, and a probability $p_L \leq p_H$ of inter-group connection.[19] That is, voters are assumed to be endowed with group membership *ex ante* and each possible dyad forms a tie independently and randomly with a probability that depends only on whether its members belong to the same group. For further discussion and justification for employing a stochastic block model, see the section on assumptions and limitations.

We assume for simplicity that $w_{ij} = w_H$ for in-group voters and $w_{ij} = w_L$ for out-group voters, with the natural assumption that $w_H \geq w_L$. Finally, we denote by $\delta \in (0, 1)$ the ratio $\frac{w_L p_L}{w_H p_H}$, which thus captures the degree of homophily on the network (lower δ corresponds to more homophily). Denoting $w_L p_L$ as $\tilde{p}_L$ for ease of presentation (since weights and probabilities do not have separable effects on average), we can re-parameterize the model by letting $\rho = \tilde{p}_H$ and hence $\tilde{p}_L = \delta\rho$,[20] so that ρ captures the baseline propensity to form ties and δ reflects the extent of differential preference for in-group members. To see how these relate to network density, note that expected density is given approximately by $\rho\,(1 - 2s(1 - s)(1 - \delta))$ for large n.[21] Here, the first term reflects the effect on density of a simple increase in connection probabilities, while the second reflects the attenuating impact of homophily as the degree of fractionalization changes, impacting the proportion of cross-group ties.

The main result, which draws on the asymptotic bounds on the average adjacency matrix derived in Appendix A, allows us to obtain closed-form expressions for each voter's centrality that hold with high probability given large n. First, note the following definition.

Definition 3. *Two sequences of random vectors $\boldsymbol{c}_n$ and $\boldsymbol{c}'_n$ are* ***asymptotically equivalent***

[19] Notably, the parameter s captures all of the information provided by the Herfindahl-Hirschman index, which is a widely used empirical measure of concentration.

[20] For technical reasons we impose that $\rho, \delta > 0$, which implies that all average weights are positive. Although this means that networks where cross-group spillovers are strictly negative fall outside of the scope of the analysis in this section, it is consistent with a distribution on w that allows for negative w to be drawn for certain i, j, provided that the expectation is positive.

[21] See Appendix A for derivation.

if and only if $\boldsymbol{c}_n - \boldsymbol{c}'_n \underset{p}{\to} 0$.

Further define $\psi_{s,\delta} := s(1-s)(1-\delta) < 0$, implying that we can rewrite expected density as $\rho(1-2\psi_{s,\delta})$, so that $\psi_{s,\delta}$ captures the attenuating impact of fractionalization and homophily on density. Then we can state the following proposition.

Proposition 1 (Expected Centrality)**.** *Suppose that the assumptions of Theorem 1 hold, and consider a sequence of random graphs $\mathcal{G}_n$ drawn from a stochastic block model. Then the centrality of a voter in party* 1 *and* 2 *is asymptotically equivalent to*

$$c_1 = \frac{(1-s)\left(\delta - \psi_{s,\delta}(\delta(1+\theta) - 1 + \theta)\right)}{sn\left((1-s)\delta + \psi_{s,\delta}\right)\left(s(\delta+\theta-1) + 1 - \theta\right)} \tag{4}$$

and

$$c_2 = \frac{\delta - \psi_{s,\delta}(\delta(1+\theta) - 1 + \theta)}{n(s - 1 + \psi_{s,\delta})(s(\delta+\theta-1) - \delta)}, \tag{5}$$

respectively.

The proof of this result is presented in Appendix D. Unlike realized networks, the expected network is necessarily complete, since all voters have positive probability of being connected to all others. Note that this need not apply to any specific realization, as all possible networks on n vertices are in the support of the generative model. Instead, the completeness of the *expected* network (more precisely, the strict positivity of the matrix of tie formation probabilities) allows us to study how changes in generative parameters affect equilibrium strategies.[22] While it remains possible that realized networks will be drawn in such a way that the equilibrium strategy differs from these expressions, Proposition 1 guarantees that this will occur with vanishing probability in sufficiently large societies. Thus, the quantities in expressions (4) and (5) serve to approximate voter centralities for *any* large society that we can expect to realize.

[22] In particular, while the adjacency matrix can be arbitrarily large, it only contains four unique values that correspond to directed connections within and between each group. Since this generates the structure of a block matrix, it is therefore possible to derive an explicit formula for its inverse, which in turn determines the value of each voter's centrality.

An immediate conclusion that follows from Proposition 1 is that, while centrality is a function of group sizes, information, and homophily, expected density does not directly relate to equilibrium transfers (but see Appendix A.5). While changes in δ and s influence density via $\psi_{s,\delta}$, which in turn relates to centrality, a "pure" increase in density via ρ does not have any impact on expected centrality, holding these parameters constant. That is, a uniform increase in connection probabilities between all voters would not influence equilibrium transfers in any way, since in this case the impact of increasing the number of others influenced by a voter is exactly offset by a concomitant decrease in the weight of each individual connection. Although their shared dependence on homophily and group share may induce a correlation between network density and transfers, Proposition 1 implies that density *per se* should not be expected to have a direct causal impact. Many informal accounts of network effects assign considerable prominence to density, indicating that it is both a defining feature of ethnic networks and a key facilitator of illicit electoral practices due to its impact on peer pressure, monitoring, and information flows (Fearon and Laitin 1996; Miguel and Gugerty 2005; Stokes 2005; Chandra 2007; Stokes et al. 2013). In contrast, our model suggests this need not be a major determinant of the efficacy of targeted transfers, further highlighting the value in formally considering the entire network structure, and not simply the connections of individual voters in a given realization.

Also of note is that $c_i \leq c_j$ when group i is larger than group j—i.e., each individual member of the minority group will always receive a higher equilibrium transfer than a member of the majority. Intuitively, this is driven by the fact that the influence of each member of the minority group increases as the group becomes smaller, making them more valuable to target.[23]

It is straightforward to examine how the total spending of candidates, as well as the level of between-group inequality, depends on these parameters by taking partial derivatives. Let B denote the sum of bribes across all voters and Q denote the level of inequality, specifically

[23] Assuming any positive degree of homophily. In the case that $\delta = 1$, members of each group are, on average, interchangeable and thus receive identical expected transfers.

$Q \equiv (b_1 - b_2)^2$, which ranges from 0 (perfect equality) to positive infinity. Then, we have the following result.

Proposition 2 (Total Bribes). *$\frac{\partial B}{\partial \alpha} > 0$, $\frac{\partial B}{\partial \gamma} < 0$, $\frac{\partial B}{\partial s} > 0$, $\frac{\partial B}{\partial \theta} > 0$, and for all $\theta \leq \theta^* \approx 0.23$, $\frac{\partial B}{\partial \delta} > 0$.*

Since θ needs to be small for the solution to be well-defined, the condition for these results will always hold for large n.[24] The effect of α, γ, and θ here are intuitive: more weight placed on public good provision, either by candidates or voters, reduces diversion of resources towards private provision. Similarly, better information increases the marginal value of targeting any voter. The other two results are non-obvious, however. The effect of δ is counterintuitive: an increase in δ (i.e., a decrease in homophily) increases total spending. The mechanism for this is straightforward—higher δ corresponds to stronger "weak ties" between groups raising the value of transfers to all voters—and corresponds closely the empirical findings in

Moreover, while it might be expected that more unequal group sizes lead to increased expenditure, the mechanism is somewhat surprising. Increasing the size of the majority group leads to a relative increase in the individual-level transfers offered to members of the *minority* group (see below), which more than compensates for the reduction in its size. It is similarly straightforward to study the effect of network parameters on inequality.

Proposition 3 (Group Size and Inequality). *Let $Q \equiv (b_1 - b_2)^2$ denote the total inequality. Then, $\frac{\partial Q}{\partial \alpha} \leq 0$, $\frac{\partial Q}{\partial s} > 0$, $\frac{\partial Q}{\partial \theta} \leq 0$, and $\frac{\partial Q}{\partial \gamma} \leq 0$. Moreover, $\frac{\partial Q}{\partial \delta} < 0$ only if $\delta \geq \sqrt{1 - 2\theta}$, and is positive otherwise.*

Once again, the effects of information and group size are intuitive, suggesting that better informed candidates in more demographically uneven societies will concentrate their

[24] In general, the cutoff θ^* depends on the other parameters; however, we verify numerically that 0.23 is an approximate lower bound and, in large networks, the invertibility assumption on the centrality matrix binds (see Assumption 2 in the Appendix).

resources more intensely in the groups that provide the highest return. As with total spending, however, the effects of homophily are surprising. While higher homophily (lower δ) can increase inequality, this only holds for networks that exhibit extremely low degrees of homophily (since by assumption θ must be small). Since group membership is the primary determinant of policy ideal points by construction and homophily based on political preferences is generally quite strong (Huber and Malhotra 2017), this is unlikely to occur in real societies. In contrast, at more moderate levels of homophily, increasing the relative influence of voters on members of their own group actually decreases the overall inequality of transfers from candidates.

This result is especially remarkable given that Dasaratha (2020) arrives at the opposite conclusion regarding Katz-Bonacich centrality on an undirected and unweighted network. In fact, the key to understanding this result is the tradeoff faced by candidates between targeting highly-connected voters, who influence many others, and voters whose neighbors are *not* highly connected, since they are more easily influenced. In the extreme, as δ approaches 0, the greater value of transfers to members of the minority is completely offset by their disconnectedness from the majority, such that the equilibrium bribes approach equality.

Heterogeneous Information

A key feature of the model studied thus far is that both candidates have identical and completely homogeneous information about the preferences of all voters, modeled as a single commonly known value of θ. Among the competing candidates, homogeneous information unsurprisingly results in homogeneous behavior. Preference for one group over another is also driven mainly by group share, with candidates tending to favor the (marginally more valuable) minority regardless of their own affiliation. In practice, however, this is unlikely to hold true. Empirical research has consistently emphasized the crucial intermediary role of brokers and local agents who possess superior knowledge about particular groups of voters, and in competitive settings informational asymmetries across candidates may account for divergent strategies (Stokes et al. 2013; Calvo and Murillo 2013; Fafchamps and Labonne 2017; Cruz, Labonne, and Querubin 2017). Formal treatments of targeted redistribution have likewise emphasized the importance of group homogeneity (Persson and Tabellini 2002) and the comparative advantage enjoyed by politicians in targeting their own supporters (Dixit and Londregan 1996; Stokes 2005). The additions to the model here can thus be thought of as extending the canonical probabilistic voting model to consider how group and network heterogeneity interact.

In this section, we study the consequences of relaxing this assumption, allowing the precision of candidates' information to vary arbitrarily across candidates and voters. Variation in information may arise due to systematic differences between the two groups and thus affect both candidates symmetrically. For instance, if partisanship is stronger in one party than another, then candidates may view those associated with the "weaker" group as more likely to be swing voters, since knowing their group label—the only information visible to the candidate in the model—is less informative about their ultimate decision. Intuitively, the first-order effect of this variation is to reduce the value of transfers to members of the less predictable group, as they are associated with a lower marginal value in expectation. Nevertheless, it is unclear *a priori* how this affects the comparative statics derived in the homogeneous case, as the reduced value of members of this group also reduces the significance of all flow-on effects in the network.

In practice, while it still holds true that members of the minority group receive higher average transfers unless the groups are approximately equal in size, the impact of fractionalization (relative group share) on total expenditure is now conditional on the relative information available about both groups, with increased fractionalization associated with higher transfers when the minority group behaves unpredictably. In addition, while homophily continues to depress total transfers, it has a highly contingent effect on inequality.

Equilibrium

We begin from the setup of the baseline model, with the distinction that the information held by candidate k about voter i's preferences is allowed to vary. In particular, voter i's net preference for candidate 1, ε_i, is now drawn from one of two uniform distributions with density parameter $\theta_i \in \{\underline{\theta}, \overline{\theta}\}$ with $\overline{\theta} > \underline{\theta}$. We can think of θ_i as voter i's private type, which is unknown to candidates.

While the candidates do not know which distribution voter i's net preference was drawn from, they have common priors and receive signals about each voter's type $m_{ik} \in \{\underline{\theta}, \overline{\theta}\}$ such that $m_{ik} = \theta_i$ with a probability (assumed greater than half) that depends on the voter-candidate pair. In other words, candidates receive informative signals about the preferences of voters and those signals may be more precise for some voters than for others. After receiving signals $\boldsymbol{m}_k = (m_{1k}, \ldots, m_{nk})$, candidates form posterior beliefs $\boldsymbol{\mu}_k = (\mu_{1k}, \ldots, \mu_{nk})$ where $\mu_{ik} = \Pr(\theta_i = \underline{\theta} | m_{ik})$ and distribute bribes accordingly.

The full derivation of equilibrium behavior under heterogeneous information is similar to the baseline case and can be found in A.4. Equilibrium bribes can now be expressed as

$$b_{ik} = [u']^{-1} \left(\frac{1}{c_i(\boldsymbol{w}, \underline{\theta}, \overline{\theta}; \mathcal{G})} \left[\gamma C(\boldsymbol{w}, \underline{\theta}, \overline{\theta}; \mathcal{G}) + \frac{1}{\alpha_k \hat{\theta}_{ik}} \right] \right),$$

where the only difference from the previous section is that bribes rely not only on the centrality measure c_i, but also on candidate k's belief about voter i's type, $\hat{\theta}_{ik} := \mathbb{E}_{\mu_{ik}}[\theta_i]$. Since each candidate's posteriors are equal to their priors in expectation (Kamenica and Gentzkow 2011), they will act according to their (common) priors on average. Therefore, it is necessarily true that $\mathbb{E}[\hat{\theta}_{i1}] = \mathbb{E}[\hat{\theta}_{i2}] = \hat{\theta}_i$ for each voter i.

Now, because one candidate may have more informative signals than the other, it no longer holds true that otherwise identical candidates will choose the same offers. For a given network, the candidate with more accurate signals will be more responsive to the realized voter types, which means they will have an advantage over their opponent in the sense that they can better anticipate whether they should spend more or less on specific voters. However, as long as each candidate has well-specified prior beliefs about voter types, then both candidates will spend the same amount on each voter on average.[25]

Further, the baseline model's result on electoral outcomes continues to hold for any realization when candidates have the same quality of information. By introducing an informational advantage to one candidate, the better-informed candidate should be able to improve their electoral performance in expectation by more precisely allocating bribes to the voters with the greatest marginal return. To the extent that signal structure depends on group membership, therefore, candidates can benefit electorally from a greater share of voters in their own group due to the more efficient targeting this permits.

Comparative Statics

A natural question is how the information available to candidates relates to the structure of society. There are two main sources of variation in information: cross-group differences in the prior distribution of types and network-dependent variations in posterior information. In this section, we study the first type of variation analytically, while Appendix C examines the impact of network dependencies through simulations.

We begin by assuming without loss of generality that $\hat{\theta}_1 < \hat{\theta}_2$. This may be interpreted as reflecting a difference in prior distributions across groups, affecting both candidates symmetrically. Applying the same approach as for the baseline model,[26] we can recover explicit expressions for centrality:

Proposition 4 (Centrality Under Heterogeneous Information)**.** *For n sufficiently large, the*

[25] It is straightforward to see that this depends on the assumption of a common prior distribution from which θ_i is drawn. If candidates hold heterogeneous priors about voter types or if types are candidate-specific, then divergence will occur on average. Our results are essentially unaffected by this modification since we focus on a single candidate.

[26] See Supplemental Materials for details and code.

centrality of a voter in group 1 *and* 2 *is asymptotically equivalent to*

$$c_1 = \frac{\delta\left(-\hat{\theta}_1 + \hat{\theta}_2 - 1\right) + (\delta - 1)s^2\left(\delta\hat{\theta}_1 + \delta + \hat{\theta}_2 - 1\right) - s\left(\delta^2(\hat{\theta}_1 + 1) - 2\delta(\hat{\theta}_1 - \hat{\theta}_2 + 1) - \hat{\theta}_2 + 1\right)}{ns((\delta - 1)s - \delta)(-\hat{\theta}_2 + s(\delta + \hat{\theta}_2 - 1) + 1)}$$

and

$$c_2 = \frac{\delta - \left((\delta - 1)s^2(\delta\hat{\theta}_2 + \delta + \hat{\theta}_1 - 1)\right) + s\left(\delta^2(\hat{\theta}_2 + 1) - 2\delta - \hat{\theta}_1 + 1\right)}{n(s - 1)((\delta - 1)s + 1)(s(\delta + \hat{\theta}_1 - 1) - \delta)},$$

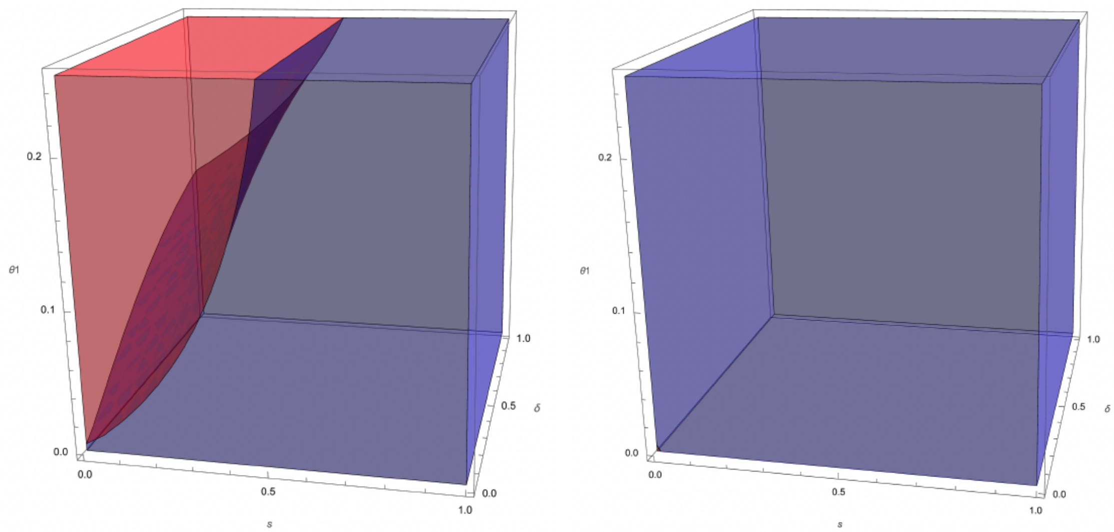
respectively, with probability approaching 1*.*

Intuitively, office motivations once again increase total candidate spending and the public cost of diverting resources reduces total candidate spending.

Proposition 5 (Bribes Under Hetergeneous Information)**.** $\frac{\partial B}{\partial \alpha} > 0$ *and* $\frac{\partial B}{\partial \gamma} < 0$*.*

However, due to the greater complexity of these expressions, it is no longer feasible to provide explicit characterizations for most of the main comparative statics—typically, the sign of the relevant derivatives depends nonlinearly on the four parameters s, δ, $\hat{\theta}_1$, and $\hat{\theta}_2$. In this section, we therefore adopt the approach of evaluating each derivative at a fine grid of points in $(s, \delta, \hat{\theta}_1)$ space, holding $\hat{\theta}_2$ constant at a range of values, assuming without loss of generality that $\hat{\theta}_1 < \hat{\theta}_2$. In Figures 2 and 3, we show the regions over which the main quantities of interest take positive and negative values assuming a moderate value $\hat{\theta}_2 = \frac{1}{4}$, while corresponding figures for other values of $\hat{\theta}_2$ can be found in Appendix B.

Unlike in the constant-information case, changes in parameters no longer have uniform effects. A clear illustration of this change from the baseline result can be seen in Figure 2a, which shows the effect of an increase in the size of group 1 on total expenditure. Whereas under homogeneous information an increase in the size of the majority group always increases expenditures (see Proposition 2), this need not necessarily be the case when information varies by group. In particular, when the less predictable group is in the minority, an increase in its size can now lead to a *increase* in total expenditure. This suggests another channel through which ethnic diversity can impact public goods provision: changing the patterns of



(a) Change in share of group 1, $\frac{\partial B}{\partial s}$

(b) Change in homophily, $\frac{\partial B}{\partial \delta}$

Figure 2: Effects of changes in social structure on total transfers given heterogeneous information by group with $\hat{\theta}_2 = 0.25$. The x-axis is given by $s \in [0,1]$, the y-axis by $\theta_1 \in [0, \theta_2]$, and the z-axis by $\delta \in [0,1]$, with a positive derivative given by blue and a negative derivative by red. A positive change in δ corresponds to a reduction in homophily, so that Panel (b) should be interpreted as showing that greater homophily leads to a reduction in transfers.

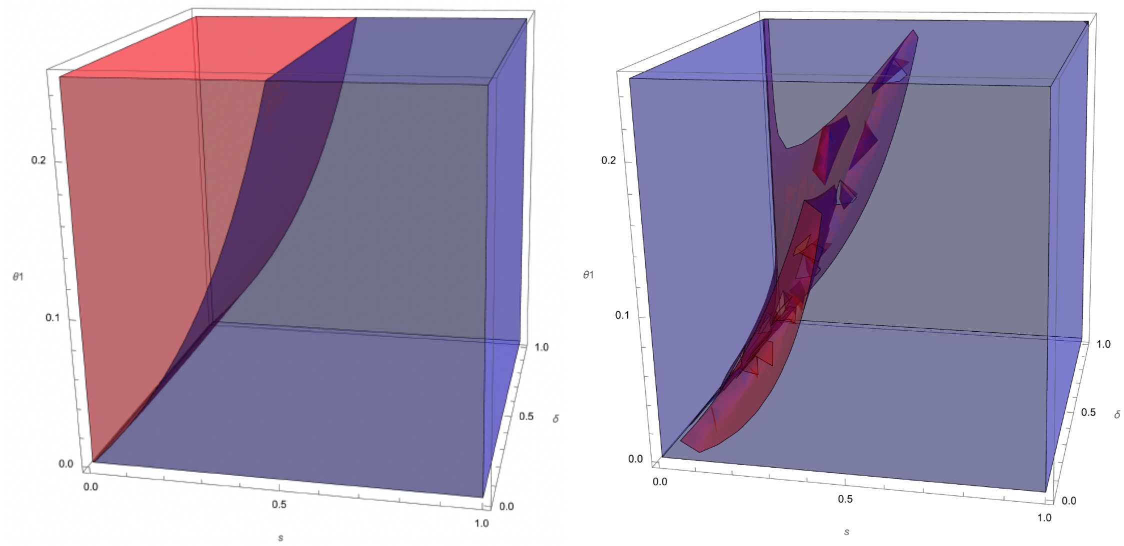
interactions among voters (Habyarimana et al. 2007; Larson and Lewis 2017; Larson 2017).

This effect is driven by a relatively higher rate of substitution into transfers to the majority group: as group sizes approach equality, members of the majority become relatively more valuable, leading to higher expenditure. Provided that homophily is not too strong, the increased number of cross-group ties then increases the value of transfers to all voters, offsetting the reduction in transfers to the minority. Crucially, this relies on $\hat{\theta}_1$ being sufficiently low relative to $\hat{\theta}_2$: when this is the case—that is, the minority group is also much less predictable—the value of transfers to minority voters is relatively lower, so the effect of increasing value of majority voters dominates.

Finally, we consider the effect of a decrease in the level of homophily in the network (increase in δ), shown in Figures 2b and 3b. As in the baseline model, decreasing levels of homophily are uniformly associated with increases in expenditure unless both $\hat{\theta}_\ell$ are sufficiently high and fractionalization is also high (see Appendix B for examples, since this only occurs for higher values of $\hat{\theta}_2$). In other words, for an increase in homophily to be associated with an increase in spending, it is necessary that the corresponding relative decrease in cross-group ties actually be associated with an *increase* in the value of transfers to at least one group. Intuitively, this relies on two conditions: (1) the proportion of possible cross-group ties is sufficiently low that a marginal reduction in their probability does not have too large an effect and (2) the members of at least one group have sufficiently high $\hat{\theta}_\ell$ and are sufficiently numerous that, on average, ties within that group are more valuable than cross-group ties.

The relationship between homophily and inequality, shown in Figure 3b, is contingent. Similarly to the baseline model, when the less-predictable group (here, group 1) is in the majority, increases in homophily are generally associated with decreases in inequality. As the quality of information for group 2 improves, this exceptional range appears mostly when δ and $\hat{\theta}_1$ are both close to 1, corresponding to high predictability and low homophily. In this region, low-level increases in homophily have the effect of further increasing c_2 and decreasing c_1, exacerbating the existing inequality by weakening the equalizing effect of cross-group ties.

By contrast, when group 1 is the minority (and receiving higher average transfers), it is now possible for increases in homophily to cause increases in inequality even at low values of δ and $\hat{\theta}_1$. While increased homophily may increase c_1 if s is sufficiently close to 0 and δ to 1, for most parameter combinations increases in homophily decrease both c_1 and c_2 due to the loss of cross-group ties. However, when group 1 members are unpredictable compared to group 2 members, the net effect is to decrease c_2 by more than c_1, as members of the majority lose more total influence than do members of the minority. There are many parameter combinations that generate this effect, however, indicating that when the smaller group is also the less predictable electorally, the effects of homophily on inequality of transfers are generally ambiguous, while still tending to decrease inequality overall.



(a) Change in share of group 1, $\frac{\partial Q}{\partial s}$

(b) Change in homophily, $\frac{\partial Q}{\partial \delta}$

Figure 3: Effects of changes in social structure on group inequality given heterogeneous information by group with $\hat{\theta}_2 = 0.25$. The x-axis is given by $s \in [0, 1]$, the y-axis by $\theta_1 \in [0, \theta_2]$, and the z-axis by $\delta \in [0, 1]$, with a positive derivative given by blue and a negative derivative by red. A positive change in δ corresponds to a reduction in homophily.

Discussion of Assumptions and Limitations

While the model presented in this paper is general in the sense that it applies to any form of targeted inducements made to connected agents in the presence of positive spillovers, the results depend on a number of assumptions and modeling choices that warrant further discussion. In this section, we consider these assumptions in detail, beginning with the baseline (fixed-graph) model before discussing the network formation process underlying our main results on social structure.

Baseline Model

In the baseline model, we normalize the total incoming social influence of each voter to one. While we consider the consequences of flexibly relaxing this assumption in Appendix A.5, this specification is preferred for several reasons. First, this modeling choice reflects a desire to focus on the effect of the information environment: the standard approach without normalization implies that the marginal impact of an additional social tie is independent of a voter's degree, which is inconsistent with rational learning. Our approach, while short of a game with information transmission, nevertheless reflects an underlying assumption that voters aggregate across all of the information available to them.

The alternative un-normalized formulation has several implications which we view as unnatural in this context. First, as the size of a voter's neighborhood increases, the relative weight placed on their own utility converges to zero, which is inconsistent with findings on peer effects in voting (Green and Gerber 2010) and with studies of online social networks, which typically indicate the reverse relationship (Jang, Lee, and Park 2014). Second, the converse also holds: voters with only a small number of connections are only negligibly influenced by their neighbors, which is again inconsistent with findings of powerful peer effects for *all* voters (Lazer et al. 2010). It is possible, however, that in some applications the total amount of social influence may depend on degree (e.g., if both are driven by latent sociability). We take up this possibility in Appendix A.5.

Further, we choose to focus on a setting without an up-front commitment problem characteristic of one-off exchanges of cash for votes whereby voters have an incentive to defect, taking payments from all candidates and voting for their *ex ante* preferred option regardless of the amounts received. As noted above, we do so because existing research has provided extensive evidence on the ways it is overcome in practice through social norms and relationship maintenance. Moreover, in many forms of clientelistic redistribution, such as the provision of public sector employment, this enforcement problem is essentially absent (Frye, Reuter, and Szakonyi 2019). By partially abstracting away from this aspect of the commitment problem, we are therefore able to focus on the role of social networks in the diffusion of *voter attitudes*, rather than purely as a means for candidates to learn about the identities of reciprocal voters as in Duarte, Finan, Larreguy, and Schechter (2019).

Lastly, we impose that social spillovers are positive. In principle, it is plausible for weights to be negative for some connections. Under ethnic patronage, for example, voters may interpret the intention of an out-group member to vote for a candidate as a sign that they should *not* do so, leading to a negative weight (Boda and Néray 2015). If we were to allow for negative weights, the main results of our model would be essentially unchanged since our centrality measure extends straightforwardly to the case of negative weights (Everett and Borgatti 2014), but we focus on the strictly positive case in the main text to avoid corner solutions, where negative spillovers outweigh positive for some voters.

Network Formation

The goal of this paper is to understand the role of social structure on redistributive strategies and their outcomes. As such, it is necessary to impose a parametric model of network formation that allows us to study the effects of basic elements of social structure in a transparent way. Consequently, we assume that the network is generated according to a stochastic block model.

Although a simplification of real network formation dynamics, stochastic block models have been consistently shown to perform well in approximating real social networks, especially with regard to community structure and local clustering (Ghasemian, Hosseinmardi, and Clauset 2019; Vaca-Ramírez and Peixoto 2022). The key simplification associated with the choice is the assumption of conditional independence of tie formation, which limits its ability replicate micro-level structures found in empirical networks that arise because the existence of a tie between individuals also increases the likelihood of ties between their mutual friends (e.g., clique formation). While micro-structure of this kind is important for individual targeting decisions, however, its relevance to high level strategic decisions made by political actors is less clear in comparison to easily observed factors such as group share or inter-group relations. Importantly for the present application, moreover, they provide a simple and transparent parameterization of two key features of social networks—the baseline propensity to form ties, and the degree of in-group preference—which permits direct consideration of the role of structural features of interest without the introduction of extraneous parameters.

Stochastic block models can be thought of as the paradigmatic case of a class of random network formations that build on the basic Erdos-Renyi model of independent tie formation (Newman, Watts, and Strogatz 2002). A number of other models exist that share this same basic structure, notably as the latent space models that have proven useful in empirical studies of networks (Breza et al. 2020). An advantage of such models over the approach used in this paper is that they are capable of encoding dyadic features that go beyond membership in a single group, and may thus perform better in prediction tasks. However, this comes at the cost of a far more complex parameterization that would obscure our main focus: studying the effect of the core elements of social structure.

Similarly, a variety of network formation models relax the requirement of conditionally independent tie formation, explicitly introducing local dependence into the process. These models can exceed the predictive performance of stochastic block models in capturing real-world social networks (Jackson 2010; Schweinberger and Handcock 2015), but they are far more heavily-parameterized and incorporate additional elements that are not clearly essential to the core features of homophily, fractionalization, and density that form the central focus of this paper. Nevertheless, we view the application of such network formation models, shifting the focus from network-level to local structure, as an interesting and important step for future research.

Discussion

In this article, we provide a formal model of elections in which candidates may offer private transfers to policy-motivated voters connected on a social network, where each voter has a politically salient group identity and transfers come at the expense of a public good. The incorporation of network effects raises the marginal value of transfers to central voters, resulting in diversion of public resources that may greatly exceed the amount voters would prefer absent social pressure. Candidates prefer to make offers to voters who are connected to many easily-influenced neighbors, not simply to those with high degrees.

Our analytical framework allows us to overcome a major limitation of many network models: the need to begin by taking a highly complex discrete graph structure as granted. By employing techniques from spectral random graph theory to study the role of social structure, we are able to explicitly characterize *expected* equilibrium strategies, which almost surely correspond to behavior in large societies. Further, our approach allows us to derive sharp comparative statics from which we can systematically compare variation in outcomes across societies based on stable underlying features of the social environment, without needing to observe complete networks.

Particularly noteworthy are our findings regarding network density and homophily. Contrary to conventional wisdom, density of ties relates to the level of group inequality and total spending only indirectly, through its association with inter-group interactions or candidate information. Homophily, meanwhile, actually decreases both for intermediate parameter values. Similarly, in line with recent reevaluations of the ethnic diversity-public good provision connection (Singh and Vom Hau 2016), we find that candidates actually face the strongest incentive to siphon public resources for targeted inducements under low fractionalization with high levels of social integration. Another implication of the model is that candidates will tend to disproportionately target minority group members regardless of their own group affiliation, although this incentive can be offset by a systematic informational advantage with respect to in-group members.

When candidates have heterogeneous information about voters, fractionalization can promote targeted strategies when not much is known about minorities but reduces inequality in almost all contexts. Homophily, on the other hand, typically reduces spending but has a nonlinear relationship with the degree of inequality among groups. A more general contribution of the model with heterogeneous information is its implications for the dilemma of targeting swing versus core voters (Calvo and Murillo 2004; Stokes 2005). From the candidate's perspective, a strategically significant feature of swing voters is that their behavior on election day is difficult to predict, whether due to inherent features of the voters themselves (such as limited interest in politics) or a lack of knowledge by the candidate.

This perspective becomes particularly relevant when considering the role of the network, as when the type distribution correlates with observable group characteristics, swing or reciprocal voters may not only be high-variance themselves, but also connected to many other such voters. Our model therefore indicates that a previously overlooked tradeoff for candidates is the extent to which they can be certain of positive spillovers from targeting specific social groups. These findings suggest that the social position of the most efficiently influenced voters should also be considered: targeted redistribution may not only rely on individual characteristics, but also on the way voters are embedded in society.

The somewhat surprising result that, holding all else constant, an increase in the number of ties in the network does not alter the optimal strategy comes directly from the tradeoff candidates face between seeking out the best-connected voters and those that have the greatest influence over others, i.e., those that are connected to many relatively isolated voters. A counterfactual increase in density in a given society—that is, a symmetric increase in the connection probabilities of all voters—will not affect centrality (and hence transfers) since the increased number of connections is exactly offset by the reduction in the influence of each individual tie. This remains true even when accounting for diverging quality of candidate information, unless information structures are directly reliant on the underlying social network. The significance of this finding is that density in itself does not necessarily matter for the reasons it is frequently assumed to. Simply adding more ties to a network does not in itself have an effect on strategic behavior unless accompanied by a change in the strength of those ties.

These results have direct implications for redistributive strategies in a variety of political settings. For instance, we predict the most intensive targeted redistribution to occur in social contexts with a large majority and small minority but with relatively low levels of social segregation. At the same time, members of the minority group are likely to benefit disproportionately from private provision, especially when candidates are well-informed about voter preferences. In the absence of strong in-group preferences, this will tend to occur regardless of the group affiliation of those dispensing resources and may lead to targeting voters that *ex ante* prefer the opposition.

While we primarily interpret nodes in the model to be individual voters, they can also represent larger aggregates of individuals, ranging from family-sized voting blocs to towns or larger regions. Targeted redistribution frequently occurs in the form of "pork," whereby allocations are made based on geographic location. In this context, spillovers occur not only between individuals, but also geographically (Chen 2010), and it may be more natural to think of politicians as considering allocations at this higher level of aggregation in some settings. Our model extends straightforwardly to this case, with homophily corresponding to geographic clustering of similar locales.

Perhaps the most important implication of these findings is that empirical work on distributive politics should not ignore or take for granted the role of network structure as a determinant of electoral strategy. While it is now well-established that network position affects individual targeting, variations in meso- and macro-level features of social networks, especially homophily and information structures, strongly shape the strategic environment. Moreover, we suggest that density of ties should neither be assumed a defining feature of ethno-political groups nor necessarily a primary determinant of the effectiveness of targeted redistribution, further demonstrating the value of explicitly modeling network dependence.

References

Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic growth* 8(2):155–194.

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public goods and ethnic divisions." *The Quarterly Journal of Economics* 114(4):1243–1284.

Baldwin, Kate, and John D Huber. 2010. "Economic versus cultural differences: Forms of ethnic diversity and public goods provision." *American Political Science Review* 104(4):644–662.

Ballester, Coralio, Antoni Calvó-Armengol, and Yves Zenou. 2006. "Who's who in networks. Wanted: The key player." *Econometrica* 74(5):1403–1417.

Banks, Jeffrey S. 2000. "Buying supermajorities in finite legislatures." *American Political Science Review* 94(3):677–681.

Battaglini, Marco, and Eleonora Patacchini. 2018. "Influencing connected legislators." *Journal of Political Economy* 126(6):2277–2322.

Beck, Paul Allen, Russell J Dalton, Steven Greene, and Robert Huckfeldt. 2002. "The social calculus of voting: Interpersonal, media, and organizational influences on presidential choices." *American political science review* 96(1):57–73.

Boda, Zsófia, and Bálint Néray. 2015. "Inter-ethnic friendship and negative ties in secondary school." *Social Networks* 43:57–72.

Bonacich, Phillip. 1987. "Power and centrality: A family of measures." *American Journal of Sociology* 92(5):1170–1182.

Breza, Emily, Arun G Chandrasekhar, Tyler H McCormick, and Mengjie Pan. 2020. "Using aggregated relational data to feasibly identify network structure without network data." *American Economic Review* 110(8):2454–2484.

Brusco, Valeria, Marcelo Nazareno, and Susan C Stokes. 2004. "Vote buying in Argentina." *Latin American research review* pp. 66–88.

Calvó-Armengol, Antoni, Eleonora Patacchini, and Yves Zenou. 2009. "Peer effects and social networks in education." *The Review of Economic Studies* 76(4):1239–1267.

Calvo, Ernesto, and Maria Victoria Murillo. 2004. "Who delivers? Partisan clients in the Argentine electoral market." *American Journal of Political Science* 48(4):742–757.

Calvo, Ernesto, and Maria Victoria Murillo. 2013. "When parties meet voters: Assessing political linkages through partisan networks and distributive expectations in Argentina and Chile." *Comparative Political Studies* 46(7):851–882.

Chandra, Kanchan. 2007. *Why ethnic parties succeed: Patronage and ethnic head counts in India*. Cambridge University Press.

Chen, Jowei. 2010. "The effect of electoral geography on pork barreling in bicameral legislatures." *American Journal of Political Science* 54(2):301–322.

Chung, Fan, and Mary Radcliffe. 2011. "On the spectra of general random graphs." *The Electronic Journal of Combinatorics* pp. P215–P215.

Chung, Fan RK, and Fan Chung Graham. 1997. *Spectral graph theory*. Number 92 American Mathematical Soc.

Cruz, Cesi. 2019. "Social networks and the targeting of vote buying." *Comparative Political Studies* 52(3):382–411.

Cruz, Cesi, Julien Labonne, and Pablo Querubin. 2017. "Politician family networks and electoral outcomes: Evidence from the Philippines." *American Economic Review* 107(10):3006–37.

Cruz, Cesi, Julien Labonne, and Pablo Querubin. 2020. "Social network structures and the politics of public goods provision: evidence from the Philippines." *American Political Science Review* 114(2):486–501.

Dal Bó, Ernesto. 2007. "Bribing voters." *American Journal of Political Science* 51(4):789–803.

Dasaratha, Krishna. 2020. "Distributions of centrality on networks." *Games and Economic Behavior* 122:1–27.

Dekel, Eddie, Matthew O Jackson, and Asher Wolinsky. 2008. "Vote buying: General elections." *Journal of Political Economy* 116(2):351–380.

Dixit, Avinash, and John Londregan. 1996. "The determinants of success of special interests in redistributive politics." *the Journal of Politics* 58(4):1132–1155.

Downs, Anthony. 1957. "An economic theory of political action in a democracy." *Journal of political economy* 65(2):135–150.

Duarte, Raúl, Frederico Finan, Horacio Larreguy, and Laura Schechter. 2019. Brokering votes with information spread via social networks. Technical report National Bureau of Economic Research.

Easterly, William, and Ross Levine. 1997. "Africa's growth tragedy: policies and ethnic divisions." *The quarterly journal of economics* pp. 1203–1250.

Eubank, Nicholas, Guy Grossman, Melina Platas, Jonathan Rodden et al. 2021. "Viral voting: Social networks and political participation." *Quarterly Journal of Political Science* 16(3):265–284.

Eubank, Stephen, VS Anil Kumar, Madhav V Marathe, Aravind Srinivasan, and Nan Wang. 2004. Structural and algorithmic aspects of massive social networks. In *Proceedings of the fifteenth annual ACM-SIAM symposium on Discrete algorithms*. pp. 718–727.

Everett, Martin G, and Stephen P Borgatti. 2014. "Networks containing negative ties." *Social networks* 38:111–120.

Fafchamps, Marcel, Ana Vaz, and Pedro C Vicente. 2020. "Voting and peer effects: Experimental evidence from Mozambique." *Economic Development and Cultural Change* 68(2):567–605.

Fafchamps, Marcel, and Julien Labonne. 2017. "Do politicians' relatives get better jobs? evidence from municipal elections." *The Journal of Law, Economics, and Organization* 33(2):268–300.

Fafchamps, Marcel, and Julien Labonne. 2020. "Family networks and distributive politics." *Journal of the European Economic Association* 18(4):1697–1725.

Fearon, James D, and David D Laitin. 1996. "Explaining interethnic cooperation." *American Political Science Review* 90(4):715–735.

Finan, Frederico, and Laura Schechter. 2012. "Vote-buying and reciprocity." *Econometrica* 80(2):863–881.

Franck, Raphael, and Ilia Rainer. 2012. "Does the leader's ethnicity matter? Ethnic favoritism, education, and health in sub-Saharan Africa." *American Political Science Review* 106(2):294–325.

Frye, Timothy, Ora John Reuter, and David Szakonyi. 2019. "Vote brokers, clientelist appeals, and voter turnout: Evidence from Russia and Venezuela." *World Politics* 71(4):710–746.

Gans-Morse, Jordan, Sebastian Mazzuca, and Simeon Nichter. 2014. "Varieties of clientelism: Machine politics during elections." *American Journal of Political Science* 58(2):415–432.

Ghasemian, Amir, Homa Hosseinmardi, and Aaron Clauset. 2019. "Evaluating overfit and underfit in models of network community structure." *IEEE Transactions on Knowledge and Data Engineering* 32(9):1722–1735.

Granovetter, Mark. 2005. "The impact of social structure on economic outcomes." *Journal of economic perspectives* 19(1):33–50.

Green, Donald P, and Alan S Gerber. 2010. "Introduction to social pressure and voting: New experimental evidence." *Political Behavior* 32:331–336.

Groseclose, Tim, and James M Snyder. 1996. "Buying supermajorities." *American Political Science Review* 90(2):303–315.

Habyarimana, James, Macartan Humphreys, Daniel N Posner, and Jeremy M Weinstein. 2007. "Why does ethnic diversity undermine public goods provision?" *American Political Science Review* 101(4):709–725.

Holland, Alisha C, and Brian Palmer-Rubin. 2015. "Beyond the machine: Clientelist brokers and interest organizations in Latin America." *Comparative Political Studies* 48(9):1186–1223.

Huber, Gregory A, and Neil Malhotra. 2017. "Political homophily in social relationships: Evidence from online dating behavior." *The Journal of Politics* 79(1):269–283.

Jackson, Matthew O. 2008. Average distance, diameter, and clustering in social networks with homophily. In *International Workshop on Internet and Network Economics*. Springer pp. 4–11.

Jackson, Matthew O. 2010. *Social and economic networks*. Princeton university press.

Jang, S Mo, Hoon Lee, and Yong Jin Park. 2014. "The more friends, the less political talk? Predictors of Facebook discussions among college students." *Cyberpsychology, Behavior, and Social Networking* 17(5):271–275.

Kamenica, Emir, and Matthew Gentzkow. 2011. "Bayesian persuasion." *American Economic Review* 101(6):2590–2615.

Katz, Leo. 1953. "A new status index derived from sociometric analysis." *Psychometrika* 18(1):39–43.

Kitschelt, Herbert. 2000. "Linkages between citizens and politicians in democratic polities." *Comparative political studies* 33(6-7):845–879.

Larson, Jennifer M. 2017. "Networks and interethnic cooperation." *The Journal of Politics* 79(2):546–559.

Larson, Jennifer M, and Janet I Lewis. 2017. "Ethnic networks." *American Journal of Political Science* 61(2):350–364.

Lazer, David, Brian Rubineau, Carol Chetkovich, Nancy Katz, and Michael Neblo. 2010. "The coevolution of networks and political attitudes." *Political Communication* 27(3):248–274.

Lizzeri, Alessandro, and Nicola Persico. 2001. "The provision of public goods under alternative electoral incentives." *American Economic Review* 91(1):225–239.

Miguel, Edward, and Mary Kay Gugerty. 2005. "Ethnic diversity, social sanctions, and public goods in Kenya." *Journal of Public Economics* 89(11-12):2325–2368.

Mostagir, Mohamed, Asuman Ozdaglar, and James Siderius. 2022. "When is society susceptible to manipulation?" *Management Science* .

Mostagir, Mohamed, and James Siderius. 2021. "Centrality in stochastic networks." *Working paper* .

Newman, Mark EJ, Duncan J Watts, and Steven H Strogatz. 2002. "Random graph models of social networks." *Proceedings of the national academy of sciences* 99(suppl_1):2566–2572.

Persson, Torsten, and Guido Tabellini. 2002. *Political economics: explaining economic policy.* MIT press.

Ravanilla, Nico, and Allen Hicken. 2023. "Poverty, social networks, and clientelism." *World Development* 162:106128.

Ravanilla, Nico, Michael Davidson Jr, and Allen Hicken. 2022. "Voting in Clientelistic Social Networks: Evidence From the Philippines." *Comparative Political Studies* 55(10):1663–1697.

Reuter, Ora John, and Graeme B Robertson. 2012. "Subnational appointments in authoritarian regimes: Evidence from Russian gubernatorial appointments." *The Journal of Politics* 74(4):1023–1037.

Rolfe, Meredith. 2012. *Voter turnout: A social theory of political participation.* Cambridge University Press.

Rueda, Miguel R. 2017. "Small aggregates, big manipulation: Vote buying enforcement and collective monitoring." *American Journal of Political Science* 61(1):163–177.

Schweinberger, Michael, and Mark S Handcock. 2015. "Local dependence in random graph models: characterization, properties and statistical inference." *Journal of the American Statistical Association* 77(3):647.

Singh, Prerna, and Matthias Vom Hau. 2016. "Ethnicity in time: Politics, history, and the relationship between ethnic diversity and public goods provision." *Comparative Political Studies* 49(10):1303–1340.

Stokes, Susan C. 2005. "Perverse accountability: A formal model of machine politics with evidence from Argentina." *American Political Science Review* 99(3):315–325.

Stokes, Susan C, Thad Dunning, Marcelo Nazareno, and Valeria Brusco. 2013. *Brokers, voters, and clientelism: The puzzle of distributive politics*. Cambridge University Press.

Vaca-Ramírez, Felipe, and Tiago P Peixoto. 2022. "Systematic assessment of the quality of fit of the stochastic block model for empirical networks." *Physical Review E* 105(5):054311.

Vicente, Pedro C, and Leonard Wantchekon. 2009. "Clientelism and vote buying: lessons from field experiments in African elections." *Oxford Review of Economic Policy* 25(2):292–305.

Online Appendix for "Some for the Price of One: Targeted Redistribution and Social Structure"

Contents

A Additional Model Details

A.1 Additional Assumptions

In order for the solution to be well-defined, we must make two assumptions.

Assumption 1. $\max_{\boldsymbol{b}_k} \theta \left(u(b_{ik}) + 1 - \gamma \sum_i b_{ik}\right) < \frac{1}{2}$ *for all* k.

First, we assume all voters have an interior probability of voting for either candidate for any transfer profile. Specifically, suppose that i is connected to all other voters on the network and that all such voters have probability 1 of voting for the candidate. Our assumptions on $u(\cdot)$ ensure that this is well-defined and is thus equivalent to a condition that θ is sufficiently small.

Assumption 2. $\boldsymbol{I} - 2\theta\tilde{\boldsymbol{A}}$ *is invertible.*

Second, we require that the matrix $\boldsymbol{I} - 2\theta\tilde{\boldsymbol{A}}$ is invertible (see Battaglini and Patacchini (2018) for further discussion of this issue), which is likewise equivalent to assuming that θ is smaller than $\frac{1}{2\lambda_1}$, where λ_1 is the largest eigenvalue of the matrix $\tilde{\boldsymbol{A}}$ (Ballester, Calvó-Armengol, and Zenou 2006), which is guaranteed by normalization to be at most 1. We thus have as a sufficient condition that $\theta \leq \frac{1}{2}$.

A.2 Equilibrium Derivation of the Baseline Model

A voter will cast a ballot for candidate 1 if and only if $U_i(1) \geq U_i(2)$. Here, candidates will not be able to perfectly anticipate voting behavior due to their imperfect information over voter preferences. Using equation (1), we can rewrite this as a condition on the size of the valence shock,

$$\varepsilon_i \leq (-1)^{x_i - 1} + u(b_{i1}) - u(b_{i2}) + \frac{\sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij}(2\phi_j - 1)}{\sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij}} + \gamma \sum_{m \in \mathcal{V}} (b_{m2} - b_{m1}),$$

where we have denoted $\phi_i := \phi_{i1}(\boldsymbol{b}) = 1 - \phi_{i2}(\boldsymbol{b})$ the probability a voter i votes for candidate 1. Noting that $\varepsilon_i \sim \mathcal{U}\left[\frac{-1}{2\theta}, \frac{1}{2\theta}\right]$ implies $\Pr(\varepsilon_i \leq \varepsilon) = \frac{1}{2} + \theta\varepsilon$, we can correspondingly write each voter's probability for voting for candidate 1 (since θ is assumed to be sufficiently small that these probabilities are interior) as

$$\begin{pmatrix} \phi_1 \\ \vdots \\ \phi_n \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + \theta \left((-1)^{x_1 - 1} + u(b_{11}) - u(b_{12}) + \frac{\sum_{j \in \mathcal{T}_1(\mathcal{G})} w_{1j}(2\phi_j - 1)}{\sum_{j \in \mathcal{T}_1(\mathcal{G})} w_{1j}} + \gamma \sum\limits_{m \in \mathcal{V}} (b_{m2} - b_{m1}) \right) \\ \vdots \\ \frac{1}{2} + \theta \left((-1)^{x_n - 1} + u(b_{n1}) - u(b_{n2}) + \frac{\sum_{j \in \mathcal{T}_n(\mathcal{G})} w_{nj}(2\phi_j - 1)}{\sum_{j \in \mathcal{T}_n(\mathcal{G})} w_{nj}} + \gamma \sum\limits_{m \in \mathcal{V}} (b_{m2} - b_{m1}) \right) \end{pmatrix}$$

Here, $\boldsymbol{\phi}$ gives the unique vector of vote probabilities given bribe profiles. While each voter's utility is subject only to their neighbor's vote probabilities, this system of equations necessarily implies that a single voter's probability of supporting candidate 1 is a function of all other voter's probability of supporting 1. This occurs because, for example, a voter i's probability ϕ_i is affected by i's neighbor j's probability ϕ_j, which in turn is affected by j's neighbor m's probability ϕ_m. Since we rule out disconnected components, ϕ_i will both affect and be affected by all other voting probabilities throughout the entire network.

In equilibrium, each candidate chooses a vector of transfers that maximizes their utility, taking the other candidate's strategy as given. This gives rise to the n first-order conditions,

$$\sum_{j=1}^{n} \frac{\partial \phi_j}{\partial b_{ik}} = \frac{1 - \lambda_{ik}}{\alpha_k}$$

where λ_i is the Lagrangian multiplier associated with voter i's nonnegativity constraint (see Battaglini and Patacchini (2018) for an analogous derivation). Differentiating ϕ_i with respect to a bribe from candidate 1 to another voter h, we have

$$\frac{\partial \phi_i}{\partial b_{h1}} = \theta \left(u'(b_{h1}) \mathbb{1}(i = h) + \frac{2 \sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij} \phi'_j}{\sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij}} - \gamma \right).$$

Then we can rewrite the candidates' problem as

$$(\boldsymbol{J}[\boldsymbol{u}] - \boldsymbol{\Gamma})^{\top} \cdot \left(\boldsymbol{I} - 2\theta \tilde{\boldsymbol{A}} \right)^{-1} \cdot \mathbf{1} = \frac{(1 - \boldsymbol{\lambda})}{\alpha_k \theta} \tag{6}$$

where $\boldsymbol{J}[\cdot]$ is a diagonal matrix with $u'(b_i)$ as the nonzero entries, $\boldsymbol{\Gamma}$ is an $n \times n$ square matrix such that every element of $\boldsymbol{\Gamma}$ is γ, $\boldsymbol{I}$ denotes the identity matrix, $\mathbf{1}$ denotes an n-vector of 1s, $\boldsymbol{\lambda}$ is an n-vector of Lagrange multipliers, and $\tilde{\boldsymbol{A}}$ is a normalized weighted adjacency matrix. Note that by Assumption 2, $(\boldsymbol{I} - 2\theta \tilde{\boldsymbol{A}})$ is guaranteed to be invertible, so the problem is well-defined.

A.3 Targeted Redistribution by Policy Rule

Consider a restriction of the model in which candidates simply choose a policy rule for the assignment of excludable ($\boldsymbol{b}$) and non-excludable goods. That is, while candidates observe the aggregate network structure, they lack the technology to make transfers to individual voters, and must instead commit to a policy rule that assigns transfers to groups of voters based on observable characteristics. Formally, candidates choose a vector $\boldsymbol{b} = \{b_1, \ldots, b_L\}$, where b_ℓ is the transfer offered to voters with membership in group ℓ. This case corresponds to a wide variety of clientelistic models in which candidates make promises based on coarse observable factors such as ethnic group or place of residence (Chandra 2007; Vicente and Wantchekon 2009), while lacking the tightly-controlled broker networks necessary to precisely target specific individuals (Cruz 2019).

Applying the same approach as in the baseline model, the candidate's optimal allocation must satisfy

$$\left(\boldsymbol{J}^P\left[\boldsymbol{u}\right] - \boldsymbol{\Gamma}^P\right)^\top \cdot \left(\boldsymbol{I}_n - 2\theta\tilde{\boldsymbol{A}}\right)^{-1} \cdot \mathbf{1} = \frac{(1-\boldsymbol{\lambda})}{\alpha_k\theta} \tag{7}$$

where $\boldsymbol{J}^P\left[\boldsymbol{u}\right]$ is now an $n \times L$ matrix where the i, ℓth entry is equal to $u'(b_\ell)$ if and only if $\ell_i = \ell$ and 0 otherwise, and $\boldsymbol{\Gamma}^P$ is equal to γn_ℓ in every entry in the ℓth row.[27] Then, reading off the first row of equation (7), we have that

$$\sum_{\ell_i=1}(u'(b_{1k}) - \gamma n_1)c_i - \sum_{\ell_j\neq 1}\gamma n_1 c_j = \frac{1-\lambda_{k1}}{\alpha_k\theta}$$

or, denoting $C_\ell(\boldsymbol{w},\theta;\mathcal{G}) = \sum_{\ell_i=\ell} c_i(\boldsymbol{w},\theta;\mathcal{G})$ and $C(\boldsymbol{w},\theta;\mathcal{G}) = \sum_i c_i(\boldsymbol{w},\theta;\mathcal{G})$,

$$b_{\ell k} = [u']^{-1}\left(\frac{1}{C_\ell(\boldsymbol{w},\theta;\mathcal{G})}\left[\gamma n_\ell C(\boldsymbol{w},\theta;\mathcal{G}) + \frac{1}{\alpha_k\theta}\right]\right). \tag{8}$$

It is straightforward to see that in the limiting case where each voter belongs to his or her own group, this model converges to the baseline. Likewise, in the average network, the two models are essentially equivalent because candidates make the same offer to every voter in a group on average, so that are results on social structure are unaffected.

On fixed networks, however, the important observation is that the constraint introduces a requirement for candidates to take into account not the value of targeting individual voters, but the *total* centrality of each group. This means that groups that perform a bridging role in the overall society, for instance, by serving as a connection between other groups, may become more attractive targets for redistribution even if no particular member of the group

[27] This reduces to the perfect targeting case when $L = n$.

has especially high centrality.

A.4 Heterogeneous Information (Equilibrium Derivation)

We begin from the setup of the baseline model, with the distinction that the information held by candidate k about voter i's preferences is allowed to vary. In particular, voter i's net preference for candidate 1, ε_i, is now drawn from one of two uniform distributions with density parameter $\theta_i \in \{\underline{\theta}, \overline{\theta}\}$ with $\overline{\theta} > \underline{\theta}$. We can think of θ_i as voter i's private type, which is unknown to candidates.

While the candidates do not know which distribution voter i's net preference was drawn from, they have common priors and receive signals about each voter's type $m_i \in \{\underline{\theta}, \overline{\theta}\}$ such that $m_i = \theta_i$ with a probability (assumed greater than half) that depends on the voter-candidate pair. In other words, candidates receive informative signals about the preferences of voters and those signals may be more precise for some voters than for others. After receiving signals $\boldsymbol{m} = (m_1, \ldots, m_n)$, candidates form posterior beliefs $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ where $\mu_i = \Pr(\theta_i = \underline{\theta} | m_i)$ and distribute bribes accordingly.

Incorporating uncertainty over voter types, we can rewrite the candidates' problem as

$$\begin{aligned} \max_{\boldsymbol{b}_k} \;\; & \alpha_k \sum_{i \in \mathcal{V}} \mathbb{E}_\mu[\phi_{ik}(\boldsymbol{b}_k, \boldsymbol{b}_{-k})] - \boldsymbol{b}_k \cdot \mathbf{1} \\ \text{subject to} \;\; & b_{ik} \geq 0 \text{ for all } i. \end{aligned}$$

Exactly as before, we can write each voter's probability for voting for candidate 1 as a function of all other voter's probability for voting for candidate 1,

$$\begin{pmatrix} \phi_1 \\ \vdots \\ \phi_n \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + \theta_1 \left(V_1(\boldsymbol{b}) + \frac{\sum_{j \in \mathcal{T}_1(\mathcal{G})} w_{1j}(2\phi_j - 1)}{\sum_{j \in \mathcal{T}_1(\mathcal{G})} w_{1j}} \right) \\ \vdots \\ \frac{1}{2} + \theta_n \left(V_n(\boldsymbol{b}) + \frac{\sum_{j \in \mathcal{T}_n(\mathcal{G})} w_{nj}(2\phi_j - 1)}{\sum_{j \in \mathcal{T}_n(\mathcal{G})} w_{nj}} \right), \end{pmatrix}$$

where we denote i's net preference for candidate 1 short of network effects by $V_i(\boldsymbol{b})$ for notational convenience so that

$$V_i(\boldsymbol{b}) := (-1)^{x_i-1} + u(b_{i1}) - u(b_{i2}) + \gamma \sum_{m \in \mathcal{V}} (b_{m2} - b_{m1}).$$

Unlike the baseline case, however, candidates maximize an expected utility that now relies on their posterior beliefs of voter types. In particular, we need to characterize the candidates' expected vote share conditional on their signals. Using the above equation, this can be expressed as

$$\mathbb{E}_\mu[\phi_i] = \frac{1}{2} + \mathbb{E}_\mu[\theta_i] V_i(\boldsymbol{b}) + \frac{\sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij}(2\mathbb{E}_\mu[\theta_i \phi_j] - \mathbb{E}_\mu[\theta_i])}{\sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij}}.$$

For the remainder of this section, we restrict all edge weights to $w_{ij} = 1$ for all $i, j \in \mathcal{V}$. This assumption is without loss of generality as Proposition 1 establishes that, on average for large n, it is only the ratio $\frac{w_L p_L}{w_H p_H}$ that determines outcomes. Hence, as long as tie formation probabilities can vary freely, edge weights have no independent effect in the expected network, which is our focus in this section.

First, note that $\mathbb{E}_\mu[\theta_i] = \overline{\theta} - \mu_i(\overline{\theta} - \underline{\theta})$. This can be thought of as the candidate's net information about voter i, taking into account both first-order uncertainty about i's vote choice and second-order uncertainty over her type. Second, by Lemmas 3 and 4 in the Appendix, we know that unilateral changes in a particular voter's type has a negligible influence on changes in other voter's vote probabilities as the network grows sufficiently large; i.e, $\phi_j(\cdot|\boldsymbol{\theta} : \theta_i = \underline{\theta}) \approx \phi_j(\cdot|\boldsymbol{\theta} : \theta_i = \overline{\theta})$ for $i \neq j$ as $n \to \infty$. Hence we can conclude that, asymptotically, $\mathbb{E}_\mu[\theta_i \phi_j] = \left(\overline{\theta} - \mu_i(\overline{\theta} - \underline{\theta})\right) \mathbb{E}_\mu[\phi_j]$. This allows us to recover n first-order conditions,

$$\frac{\partial \mathbb{E}_\mu[\phi_i]}{\partial b_{h1}} = \left(\overline{\theta} - \mu_i(\overline{\theta} - \underline{\theta})\right) \left(u'(b_{h1}) \mathbb{1}(i = h) + \frac{2 \sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij} [\mathbb{E}[\phi_j]]'}{\sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij}} - \gamma \right).$$

Hence it follows as in the baseline model that the optimality condition

$$(\boldsymbol{J}[\boldsymbol{u}] - \boldsymbol{\Gamma})^{\top} \cdot \left(\boldsymbol{I} - 2\Theta\tilde{\boldsymbol{A}}\right)^{-1} \cdot \mathbf{1} = \Theta^{-1} \cdot \frac{(1-\boldsymbol{\lambda})}{\alpha_k}, \tag{9}$$

where $\Theta := \overline{\theta}\boldsymbol{I} - (\overline{\theta} - \underline{\theta})\boldsymbol{M}$ and $\boldsymbol{M}$ is an $n \times n$ diagonal matrix with posteriors μ_i as nonzero elements. Equilibrium bribes can therefore be explicitly expressed

$$b_{ik} = [u']^{-1}\left(\gamma n + \frac{1}{\alpha_k \hat{\theta}_{ik} c_i(\boldsymbol{w}, \underline{\theta}, \overline{\theta}; \mathcal{G})}\right),$$

where the only difference from the previous section is that equilibrium bribes to a voter i from a candidate k rely not only on their centrality measure c_i, now the ith element of $\boldsymbol{c} = (\boldsymbol{I} - 2\Theta\tilde{\boldsymbol{A}})^{-1} \cdot \mathbf{1}$, but also on candidate k's belief about voter i's type, $\hat{\theta}_{ik} := \mathbb{E}_{\mu}[\theta_i]$. By the law of iterated expectations, since each candidate's posteriors are equal to their priors in expectation, they will act according to their (common) priors on average. Therefore, it is necessarily true that $\mathbb{E}[\hat{\theta}_{i1}] = \mathbb{E}[\hat{\theta}_{i2}] = \hat{\theta}_i$ for each voter i.

Now, because one candidate may have more informative signals than the other, it no longer holds true that otherwise identical candidates will choose the same offers. For a given network, the candidate with more accurate signals will be more responsive to the realized voter types, which means they will have an advantage over their opponent in the sense that they can better anticipate whether they should spend more or less on specific voters. However, as long as each candidate has well-specified prior beliefs about voter types, then both candidates will spend the same amount on each voter on average.

Further, the baseline model's result on electoral outcomes continues to hold for any realization when candidates have the same quality of information. By introducing an informational advantage to one candidate, the better-informed candidate should be able to improve their electoral performance in expectation by more precisely allocating bribes to the voters with the greatest marginal return. Nonetheless, gains in electoral performance remain orthogonal to group membership to the extent the informational structure is also orthogonal to group membership.

A.5 Degree-Dependent Heterogeneity (Equilibrium Derivation)

Although our preferred specification involves normalization of total social influence to unity, it is reasonable to suppose that in some settings more-connected voters may be more susceptible to influence from their peers. For instance, if both the propensity to form ties and the extent to which others' opinions are taken into account when making decisions are correlated with intrinsic sociability, then this pattern would be expected to emerge. Alternatively, if social influence is at least partly a consequence of in-group monitoring and peer pressure, then the total pressure on highly-connected voters may be more than the sum of the individual weights due to an increased need to attend to social conformity.

To account for these possibilities, in this section we augment the baseline model with a weakly increasing *total influence* function $\xi(d_i)$ that parameterizes the impact of degree on susceptibility, where we denote by $d_i = \sum_{j \in \mathcal{T}_i(\mathcal{G})} w_{ij}$ the weighted degree of voter i. It is straightforward to see that $\xi(\cdot) = 1$ corresponds to the baseline case, while $\xi(d) = d$ yields the original un-normalized network. We are therefore interested in the intermediate case where $1 < \xi(d) < d$ for any d. We then have

$$\begin{pmatrix} \phi_1 \\ \vdots \\ \phi_n \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + \theta\left(V_1(\boldsymbol{b}) + \frac{\xi(d_1)\sum_{j \in \mathcal{T}_1(\mathcal{G})} w_{1j}(2\phi_j - 1)}{d_1}\right) \\ \vdots \\ \frac{1}{2} + \theta\left(V_n(\boldsymbol{b}) + \frac{\xi(d_n)\sum_{j \in \mathcal{T}_n(\mathcal{G})} w_{nj}(2\phi_j - 1)}{d_n}\right) \end{pmatrix}$$

It follows from the same argument as before that the optimal solution is determined by

$$(\boldsymbol{J}\left[\boldsymbol{u}\right] - \boldsymbol{\Gamma})^\top \cdot \left(\boldsymbol{I} - 2\theta\hat{\boldsymbol{A}}\right)^{-1} \cdot \mathbf{1} = \frac{(1 - \boldsymbol{\lambda})}{\alpha_k \theta} \tag{10}$$

where $\hat{\boldsymbol{A}}$ is given by

$$\hat{A}_{ij} = \frac{w_{ij}\xi(d_i)A_{ij}}{d_i}.$$

Under the assumption $1 < \xi(d) < d$, it follows from the proof of Lemma 1 and the associated results in Dasaratha (2020) and Mostagir and Siderius (2021) that we have $\lim_{n\to\infty} \Pr(\|\boldsymbol{c}^{(n)}(\hat{\boldsymbol{A}}) - \boldsymbol{c}^{(n)}(\bar{\hat{\boldsymbol{A}}})\| > \epsilon) = 0$. Moreover, it is clear that since on the average network all voters with the same group affiliation share the same expected degree, we have $\xi_i = \xi_{\ell_i}$ for all i. Since ξ depends implicitly on all generative parameters via the average degree, we assume a simple parametric form in order to study how the main results are affected by this change:

$$\xi(d_i) = \xi_0 + \beta d_i$$

or, substituting in for expected degree on the average network,

$$\xi(d_i) = \xi_0 + \beta\rho n\left(s + (1-s)\delta\right).$$

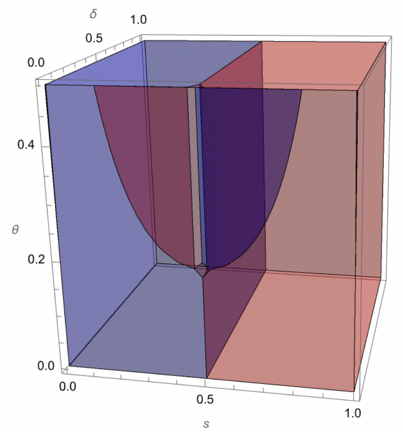
We then follow the same procedure as in the heterogeneous information case to derive comparative statics in terms of the structure parameters, with the distinction that ρ now plays a role. For simplicity of exposition, we fix $\xi_0 = 1$, although results are essentially unchanged by varying this parameter.

Much as in the baseline model, increases in the value placed on the public good generally decrease both spending and inequality, while the degree of office motivation of the candidate also negatively relates to inequality. However, it is no longer necessarily true (see Appendix B) that more office-motivated candidates spend more, as the increase in the relative value of highly-connected voters may actually incentivize *less* spending on targeted transfers in cases with high homophily and groups of very unequal size.

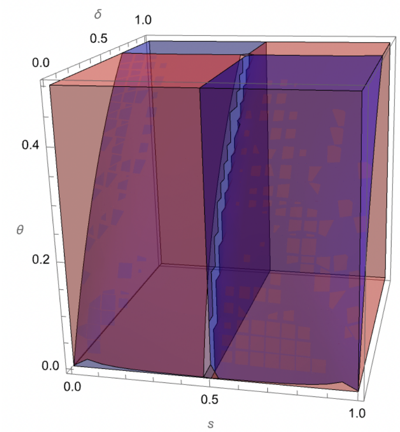
For the remaining results, we present comparative statics at two fixed parameter levels: $\rho = \beta = 0.1$ and $\rho = \beta = 0.9$.[28] These correspond, respectively, to a slight relaxation of the normalization assumption on a sparse network, and to a densely connected network with almost no normalization. Comparison across the two cases thus provides an indication of

[28] Since the two parameters appear only as a product, it is uninformative to consider cases where they are not equal, since, e.g., the solution with $\rho = 0.1$ and $\beta = 1$ is equivalent to that with $\rho = \beta = \sqrt{0.1}$.

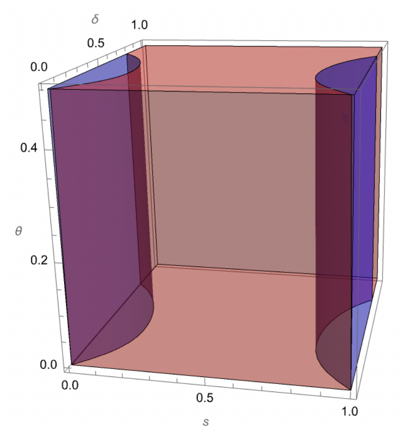
the consequences of relaxing the assumption.



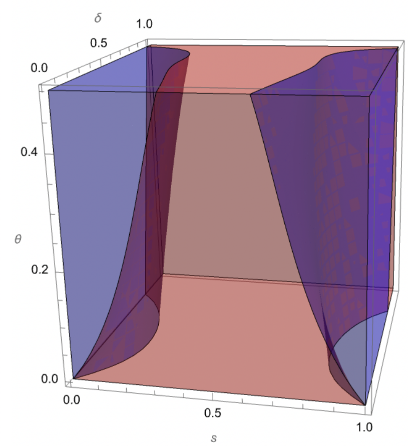
(a) $\frac{\partial B}{\partial s}$ with $\rho = \beta = 0.1$



(b) $\frac{\partial B}{\partial s}$ with $\rho = \beta = 0.9$



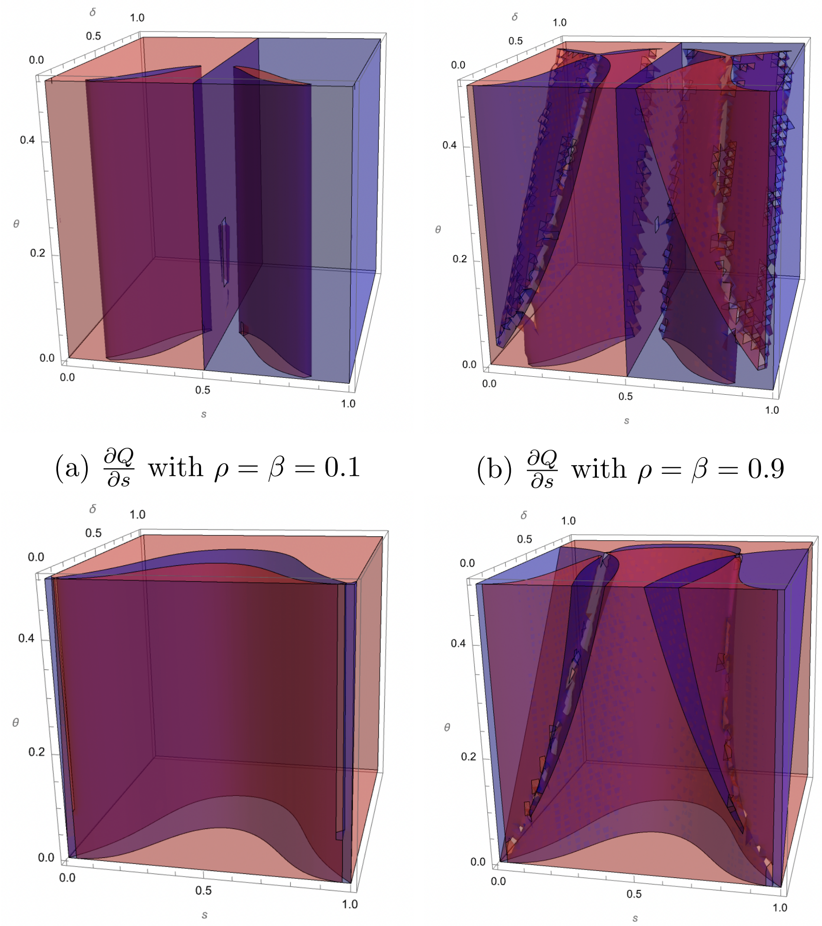
(c) $\frac{\partial B}{\partial \delta}$ with $\rho = \beta = 0.1$



(d) $\frac{\partial B}{\partial \delta}$ with $\rho = \beta = 0.9$

Figure 4: Effects of changes in social structure on total transfers given degree-dependent heterogeneity. Here, $\rho = \beta = 0.1$ in Panels (a) and (c) reflects a minimal relaxation of our normalizing assumption, whereas $\rho = \beta = 0.9$ in Panels (b) and (d) show the results approaching an un-normalized weighting scheme. Recall that a positive change in δ corresponds to a reduction in homophily.

Figure 4 replicates the main results of the baseline model under flexible normalization. Notably, the effect of both increases in homophily and fractionalization is partly reversed from the baseline model with even a minimal level of degree-dependent influence: for most parameter combinations, both homophily and fractionalization now lead to an increase in expenditure. This is mostly driven by the increase in value of highly connected individuals in the absence of normalization—candidates spend the most when they can exploit the division of society into two relatively equal-sized but weakly-connected groups that each have a high probability of containing some highly-connected individuals.

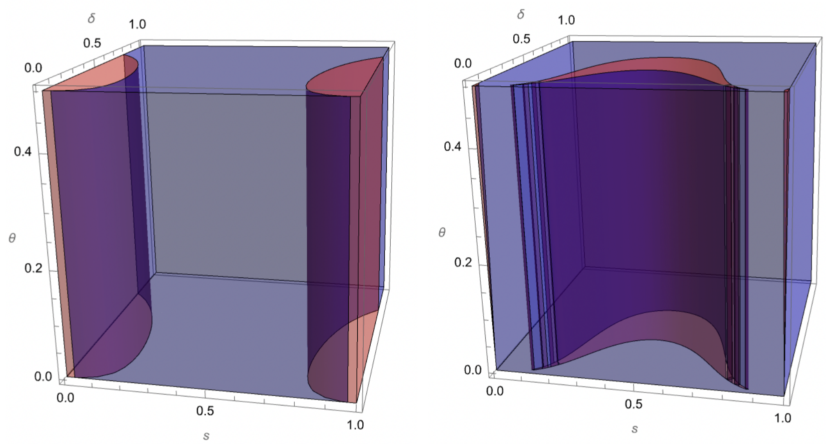


(a) $\frac{\partial Q}{\partial s}$ with $\rho = \beta = 0.1$ (b) $\frac{\partial Q}{\partial s}$ with $\rho = \beta = 0.9$

(c) $\frac{\partial Q}{\partial \delta}$ with $\rho = \beta = 0.1$ (d) $\frac{\partial Q}{\partial \delta}$ with $\rho = \beta = 0.9$

Figure 5: Effects of changes in social structure on group inequality given degree-dependent heterogeneity. Here, $\rho = \beta = 0.1$ in Panels (a) and (c) reflects a minimal relaxation of our normalizing assumption, whereas $\rho = \beta = 0.9$ in Panels (b) and (d) show the results approaching an un-normalized weighting scheme. Recall that a positive change in δ corresponds to a reduction in homophily.

Figure 5 shows the analogous results for inequality. While these overlap at times with those in the baseline model, they are also highly unstable and contingent, reflecting the added complexity. As with total transfers, however, it is noteworthy that over the greater part of the parameter space, the effect of homophily is reversed: more homophilous societies also experience greater inequality, which reflects the superior ability of candidates to exploit dense connections within, but not necessarily across, groups.



(a) $\frac{\partial B}{\partial \rho}$ (b) $\frac{\partial Q}{\partial \rho}$

Figure 6: Effects of changes in density of ties on total spending and group inequality given degree-dependent heterogeneity.

Finally, Figure 6 shows the impact of changes in density, parameterized by ρ, on total spending and group inequality.[29] While, intuitively, an increase in the total proportion of ties results in greater spending at moderate values, it is important to note that this need not hold true when the group sizes are unbalanced and a significant degree of homophily is observed. This is because, in this case, the benefits of additional ties accrue disproportionately to the majority group, while the small, disconnected minority remains isolated and, reducing the value of targeted transfers.

While the predictions from the flexibly-normalized model agree with those of the baseline for many parameter combinations, then, they are also diametrically opposed on several key points, particularly the impact of homophily on total transfers. Ultimately, which version of the model's predictions is more plausible is an empirical question and depends on the extent to which social influence depends on degree in practice. A key benefit of this extension is thus to draw attention to the potentially critical importance of this seemingly inconsequential feature of social interaction for the relationship between non-programmatic spending and transfers.

[29] We show here only the case of $\beta = \rho = 0.1$, since they are similar in nature at other values.

A.6 Derivation of Expected Density

The expected density is the expected number of ties within and across each group out of the total number of possible ties on the network, i.e.

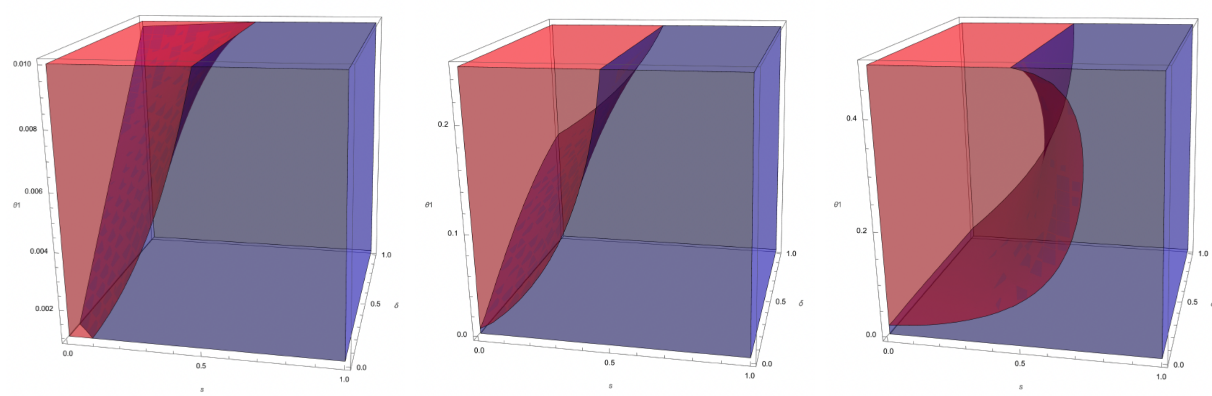
$$\frac{(n-1)\left(sn(sn-1)\tilde{p}_H + (1-s)n((1-s)n-1)\tilde{p}_H + 2s(1-s)n^2\tilde{p}_L\right)}{n}$$

Substituting for $\rho = \tilde{p}_H$ and $\delta\rho = \tilde{p}_L$, and observing that as $n \to \infty$, $\frac{n}{n-1} \approx 1$, we then have that

$$\frac{\rho}{n-1}\left(2ns(\delta - \delta s - 1 + s) + n - 1\right) \approx \rho\left(2s(1-s)(\delta-1)+1\right).$$

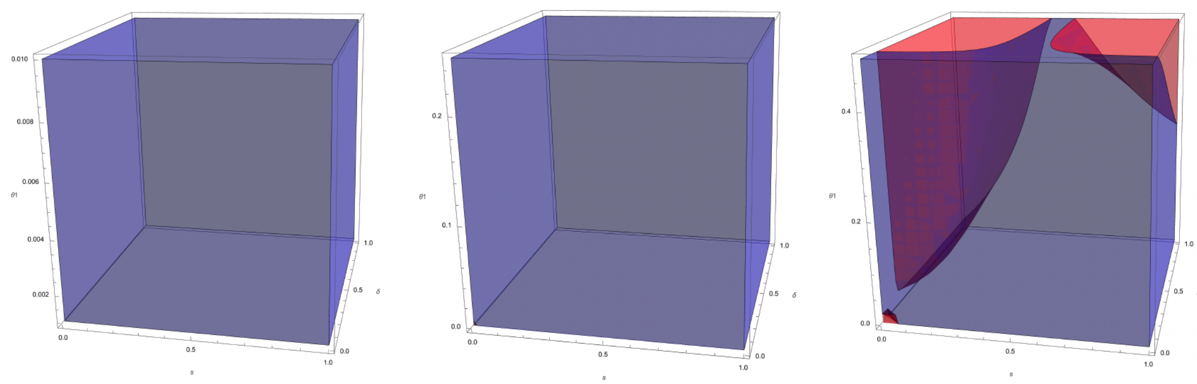
B Additional Tables and Figures

This section includes additional tables and figures for the case of heterogeneous information.



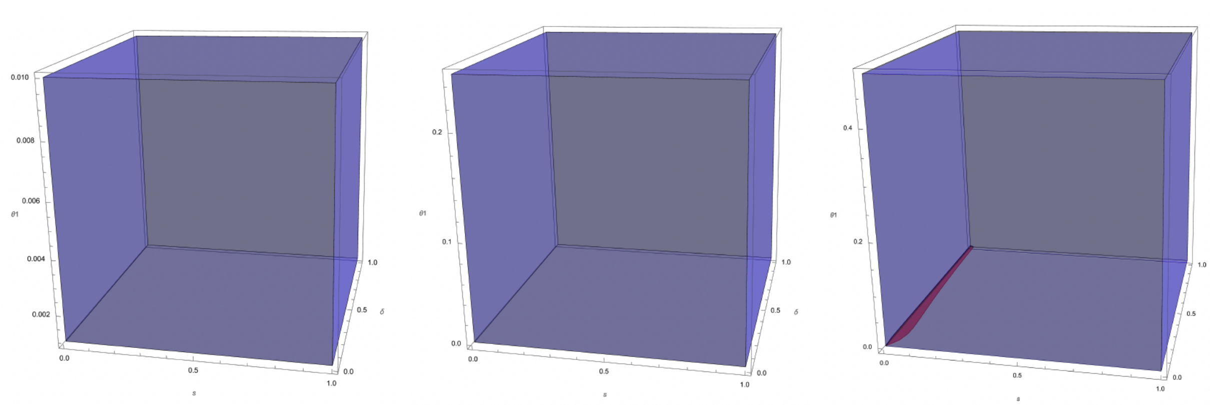
(a) $\hat{\theta}_2 = 0.01$ (b) $\hat{\theta}_2 = 0.25$ (c) $\hat{\theta}_2 = 0.5$

Figure 7: $\frac{\partial B}{\partial s}$ as a function of parameters for varying values of $\hat{\theta}_2$. The x-axis is given by $s \in [0,1]$, the y-axis by $\theta_1 \in [0, \theta_2]$, and the z-axis by $\delta \in [0,1]$. Color shading is blue if $\frac{\partial B}{\partial s} > 0$ and red if $\frac{\partial B}{\partial s} \leq 0$.

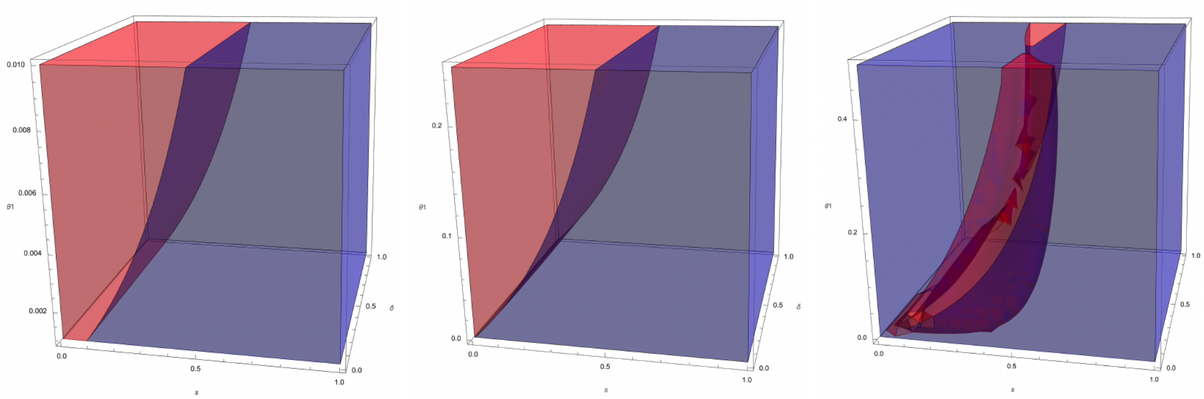


(a) $\hat{\theta}_2 = 0.01$ (b) $\hat{\theta}_2 = 0.25$ (c) $\hat{\theta}_2 = 0.5$

Figure 8: $\frac{\partial B}{\partial \delta}$ as a function of parameters for varying values of $\hat{\theta}_2$. The x-axis is given by $s \in [0, 1]$, the y-axis by $\theta_1 \in [0, \theta_2]$, and the z-axis by $\delta \in [0, 1]$. Color shading is blue if $\frac{\partial B}{\partial \delta} > 0$ and red if $\frac{\partial B}{\partial \delta} \leq 0$.

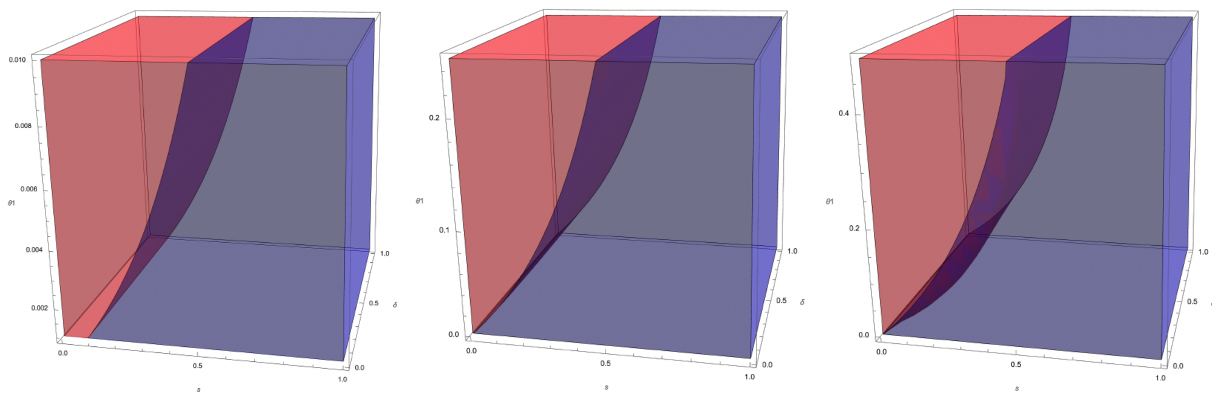


(a) $\frac{\partial B}{\partial \hat{\theta}_1}$, $\hat{\theta}_2 = 0.01$ (b) $\frac{\partial B}{\partial \hat{\theta}_1}$, $\hat{\theta}_2 = 0.25$ (c) $\frac{\partial B}{\partial \hat{\theta}_1}$, $\hat{\theta}_2 = 0.5$



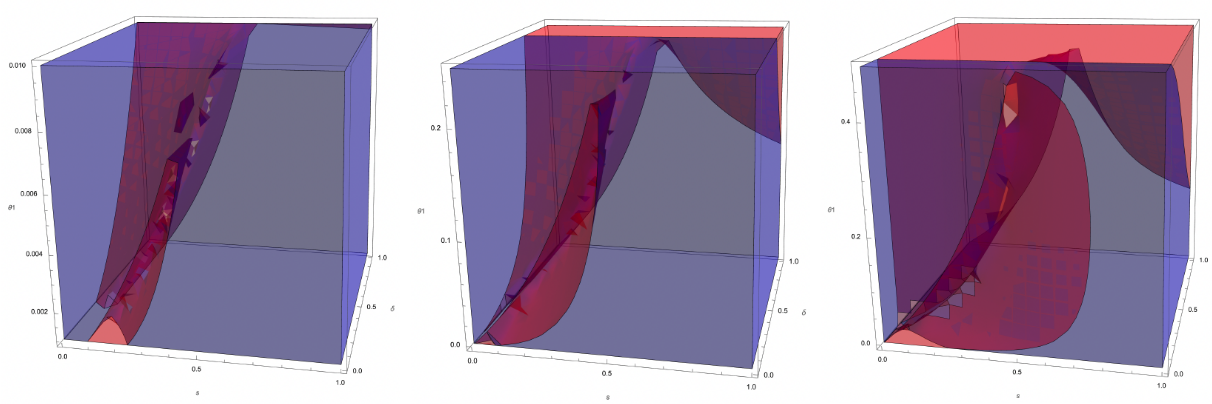
(d) $\frac{\partial B}{\partial \hat{\theta}_2}$, $\hat{\theta}_2 = 0.01$ (e) $\frac{\partial B}{\partial \hat{\theta}_2}$, $\hat{\theta}_2 = 0.25$ (f) $\frac{\partial B}{\partial \hat{\theta}_2}$, $\hat{\theta}_2 = 0.5$

Figure 9: Derivatives of total transfers with respect to information as a function of parameters. The x-axis is s, the y-axis is δ, and the z-axis is θ_1.



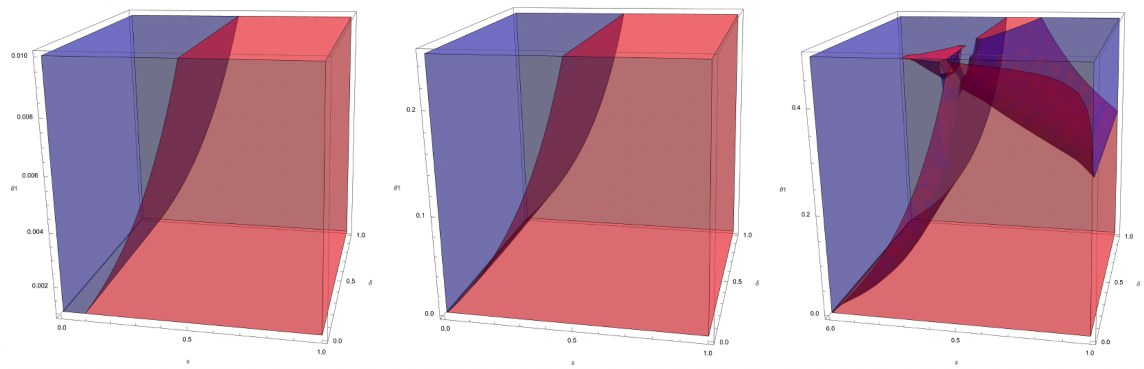
(a) $\hat{\theta}_2 = 0.01$ (b) $\hat{\theta}_2 = 0.25$ (c) $\hat{\theta}_2 = 0.5$

Figure 10: $\frac{\partial Q}{\partial s}$ as a function of parameters for varying values of $\hat{\theta}_2$. The x-axis is given by $s \in [0,1]$, the y-axis by $\theta_1 \in [0,\theta_2]$, and the z-axis by $\delta \in [0,1]$. Color shading is blue if $\frac{\partial Q}{\partial s} > 0$ and red if $\frac{\partial Q}{\partial s} \leq 0$.



(a) $\hat{\theta}_2 = 0.01$ (b) $\hat{\theta}_2 = 0.25$ (c) $\hat{\theta}_2 = 0.5$

Figure 11: $\frac{\partial Q}{\partial \delta}$ as a function of parameters for varying values of $\hat{\theta}_2$. The x-axis is given by $s \in [0,1]$, the y-axis by $\theta_1 \in [0,\theta_2]$, and the z-axis by $\delta \in [0,1]$. Color shading is blue if $\frac{\partial Q}{\partial \delta} > 0$ and red if $\frac{\partial Q}{\partial \delta} \leq 0$.



(a) $\hat{\theta}_2 = 0.01$ (b) $\hat{\theta}_2 = 0.25$ (c) $\hat{\theta}_2 = 0.5$

Figure 12: $\frac{\partial Q}{\partial \hat{\theta}_1}$ as a function of parameters for varying values of $\hat{\theta}_2$. The x-axis is given by $s \in [0, 1]$, the y-axis by $\delta \in [0, 1]$, and the z-axis by $\theta_1 \in [0, \theta_2]$. Color shading is blue if $\frac{\partial Q}{\partial \hat{\theta}_1} > 0$ and red if $\frac{\partial Q}{\partial \hat{\theta}_1} \leq 0$.

C Network-Dependent Information

It is natural to think network dependence and candidate information are mutually reinforcing: social connections between brokers and voters are important precisely because they provide information about voter preferences, facilitating more accurate targeting (Finan and Schechter 2012; Stokes et al. 2013). In this section, we therefore consider a special case of the heterogeneous information model in which candidates' posterior beliefs about voter behavior depend on their position in the network.

We now place the candidates on the network and assume that a voter's type θ_i depends on their position on the network. If voters whose shortest-path distance from a candidate is lower have a higher probability that $\theta_i = \overline{\theta}$, then a systematic relationship will exist between i's probability of connecting with candidates p_{ik} and their average posterior information $\hat{\theta}_i$.[30]

In contrast to previous results, density may have a direct effect, as increases in edge density will increase the likelihood of all voters being type $\overline{\theta}$ through a decrease in their expected distance from any candidate, which will directly affect candidates' posterior beliefs and the corresponding equilibrium bribes. Consequently, changes in density will now affect expected equilibrium behavior by changing the prior distribution of types. Although $\theta_i = \overline{\theta}$ symmetrically increases a voter's value to both candidates, equilibrium electoral outcomes may be affected if candidates enjoy systematically more precise information about socially proximate voters. In the two-candidate case, candidates will not only target their own close neighbors but also their opponent's, since both are associated with higher posteriors.

For this reason, we focus on the one-candidate case, which arises naturally in many applications. For example, vote brokers typically rely heavily on personal connections with voters, as the preferences of more socially distant voters are less legible, and only one agent typically makes offers in a given locality (Stokes et al. 2013; Holland and Palmer-Rubin 2015).

[30] Since θ_i measures the predictability of a voter's behavior given their observable characteristics, it is plausible that voters who are more connected to political candidates will vote more consistently.

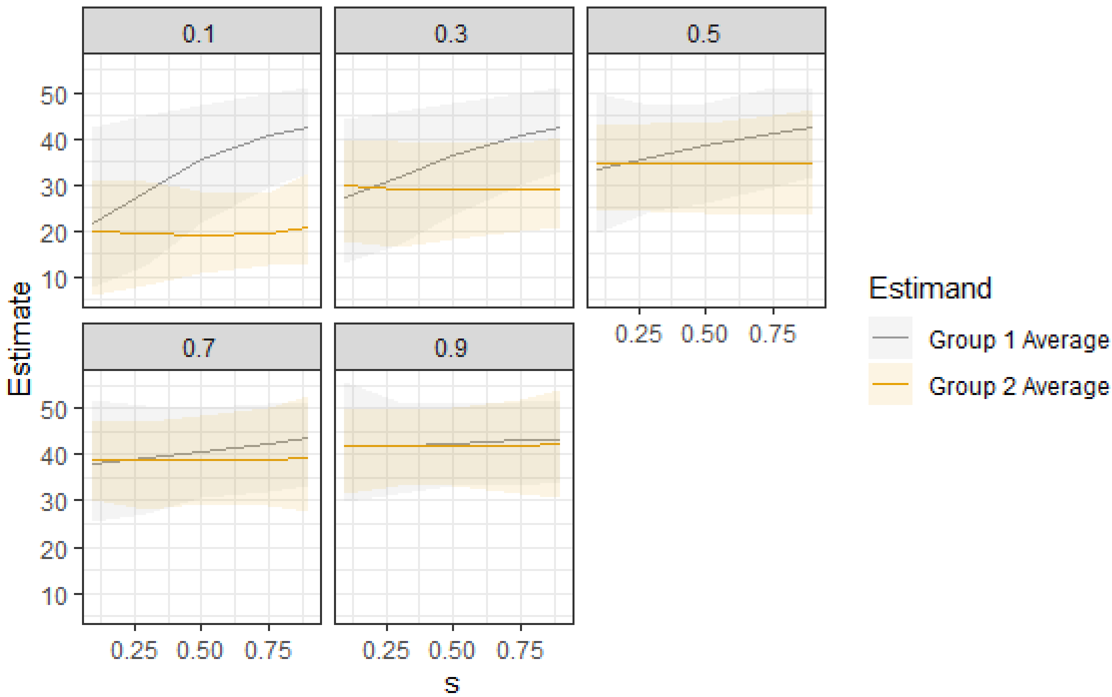


Figure 13: Average transfers from candidate 1 by group as a function of group 1's share (s), by homophily (δ).

Figures 13 and 14 show the results of model simulations. Since the asymptotic results in the preceding section require independence between θ_i and p_{ij}, we take the approach of estimating $\mathbb{E}[b_i]$ on a finite network directly through 1,000 repeated draws of networks from the stochastic block model. For each combination of parameters, we calculate the average transfers to members of each group, $\bar{b}_\ell = \mathbb{E}[b_i|\ell_i]$, the overall inequality, $Q = \left(\bar{b}_1 - \bar{b}_2\right)^2$, and the total transfers $B = \sum_i b_i$ at each draw. Estimates are then calculated as the average across all draws, along with bootstrapped 95% percentile confidence intervals.

Since our focus is primarily on social structure, we repeat the process over a grid of values between 0 and 1 for each of $\delta = \frac{p_L}{p_H}$ and $s = \frac{n_1}{n}$, which govern homophily and fractionalization. To reduce the computational burden, the remaining parameters are held constant at moderate values of $n = 200, \alpha_1 = \alpha_2 = 1, \gamma = \frac{1}{400}, u(b) = 200\ln(b)$, while we assume for simplicity that candidate k's information decays in social distance according to a power law, $\hat{\theta}_i = 2^{-d(i,k)}$.

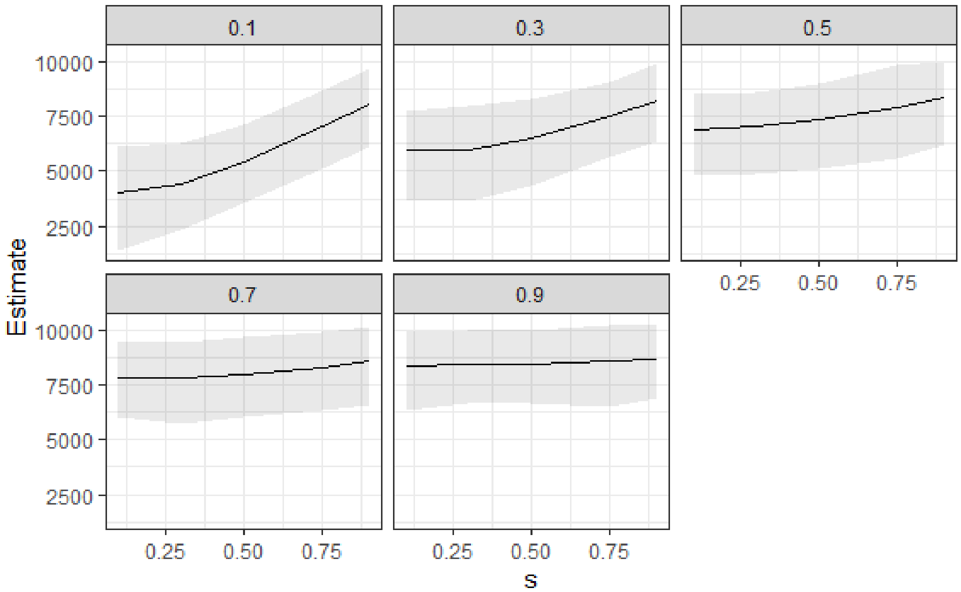


Figure 14: Total transfers from candidate 1 as a function of group 1's share (s), by homophily (δ)

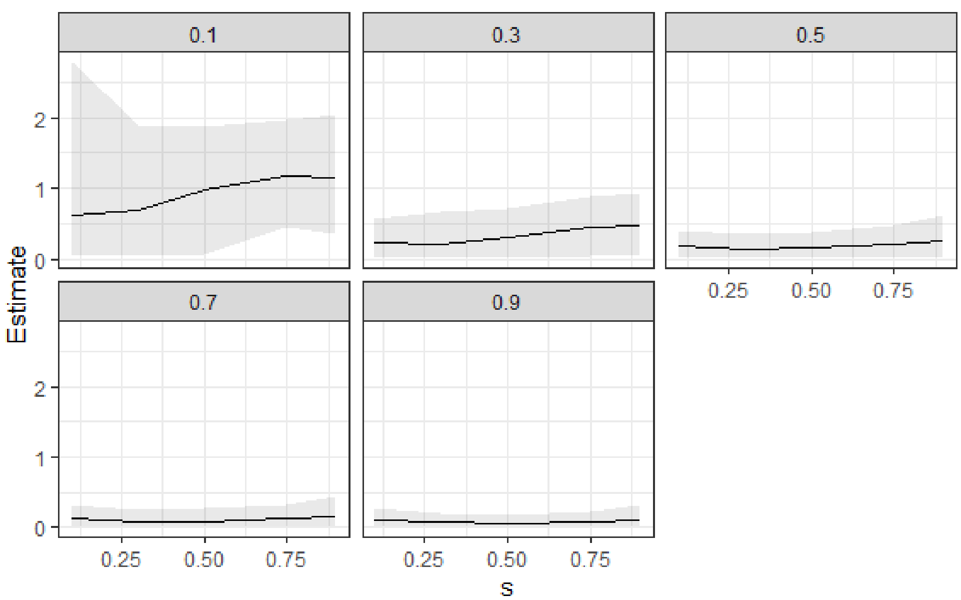


Figure 15: Inequality of average transfers from candidate 1 as a function of group 1's share (s), by homophily (δ)

As Figure 13 demonstrates, the introduction of network dependence in a one-candidate environment introduces a distinct element of in-group favoritism, consistent with empirical observations across a variety of settings. As might be expected, this favoritism is entirely driven by homophily: when δ is close to 1 (low homophily), the two groups receive almost identical amounts on average, with consistently large differences emerging only when homophily is quite extreme. Intuitively, preference for the candidate's own group members is driven by lower social distance on average, amplified by greater separation between the groups, which leads to higher confidence that bribes to them will lead to an increase in vote share. In contrast to the independent case, this tendency to favor in-groups completely outweighs the preference for minorities. While the relative size of the groups has little effect on average on transfers to out-group members, when homophily is large transfers to the in-group are sharply increasing in group size. Under homophily, an increase in the size of the in-group further decreases the expected shortest path distance, thus raising the value of transfers to all group members. As a consequence, this also tends to increase the level of inequality (Figure 15) as more funds are diverted to the candidate's in-group members.

As can be seen in Figure 14, however, this effect is insufficient to completely offset the overall loss in network spillovers induced by an increase in homophily. While, for a given value of δ, expenditure on private transfers is highest when $s \to 1$, driven by increases in in-group spending, it is also increasing in δ for all values of s. In fact, as s approaches 1, the probabilities converge to the same values as when $\delta = 1$, so that in either case the candidate is, on average, as close as possible to all voters.

When information depends on network structure in this way, therefore, we would expect the greatest diversion of funds from public to private goods in societies that are either perfectly homogeneous or have minimal levels of social segregation.

D Supplementary Propositions and Proofs

The following lemma allows us to make asymptotic statements that will hold with high probability and therefore justifies the analysis of the expected, rather than the realized, network in the context of large elections.

Lemma 1. *Under the assumptions of Theorem 1, in addition to assuming that the minimum expected degree* $\underline{d}_n$ *grows at a rate greater than* $\ln(n)$ *and that the expected Laplacian matrix has second-smallest eigenvalue bounded away from zero, for any* $\epsilon > 0$, $\lim_{n\to\infty} \Pr(\|\boldsymbol{c}^{(n)}(\tilde{\boldsymbol{A}}) - \boldsymbol{c}^{(n)}(\bar{\tilde{\boldsymbol{A}}})\| > \epsilon) = 0$.

Proof of Lemma 1. Let $\mathcal{G}^{(n)}$ be a sequence of random graphs over n vertices, and denote by $\delta_{(n)}$ the smallest expected weighted degree, i.e., $\delta_{(n)} = \min_i \sum_j w_{ij}^{(n)} p_{ij}^{(n)}$. Further, let $\bar{w}_{(n)} = \max_{i,j} w_{ij}^{(n)}$ and $\underline{w}_{(n)} = \min_{i,j} w_{ij}^{(n)}$ be the largest and smallest individual weights, satisfying $\frac{\bar{w}_{(n)}}{\underline{w}_{(n)}} \leq \omega$ for some $\omega > 0$ for all n. Then if there exists a non-decreasing sequence of $k_{(n)} > 0$ such that $\delta_{(n)} \geq k_{(n)} \ln(n)$ and $\underline{w}_{(n)} \cdot \bar{w}_{(n)} = o\left(\sqrt{\frac{\delta_{(n)}}{\ln(n)}}\right)$, then the realized centrality vector centrality vector $\boldsymbol{c}^{(n)}(\tilde{\boldsymbol{A}})$ is with high probability close to the centrality of the average graph $\boldsymbol{c}^{(n)}(\bar{\tilde{\boldsymbol{A}}})$ for large n.

Under the stated assumptions, we can apply Theorem 1 to conclude that for any $\xi > 0$, for all n we have

$$\Pr\left(||\boldsymbol{L}_W - \bar{\boldsymbol{L}}_W|| \leq 4\sqrt{\frac{3\omega \ln(4n/\xi)}{\delta}} \right) \geq 1 - \xi$$

Furthermore, by the assumption on the growth rate of the minimum degree, $\lim_{n\to\infty} 4\sqrt{\frac{3\omega \ln(4n/\xi)}{\delta}} = 0$ regardless of the ξ chosen, so that under the 2-norm,

$$\boldsymbol{L}_W \underset{p}{\to} \bar{\boldsymbol{L}}_W$$

Now, for convenience call $\boldsymbol{B} = \boldsymbol{I} - \boldsymbol{L}_W$ and $\bar{\boldsymbol{B}}$ the expected equivalent. Now clearly also

$\boldsymbol{B} \underset{p}{\rightarrow} \bar{\boldsymbol{B}}$, and furthermore we can write $\boldsymbol{B} = \boldsymbol{D}^{-1/2}\boldsymbol{A}_W\boldsymbol{D}^{-1/2} = \boldsymbol{D}^{-1/2}\tilde{\boldsymbol{A}}\boldsymbol{D}^{1/2}$. So we can write, using properties of matrix norms (abusing notation in the second step slightly so that the maximum is over the norm of the matrices) and the above result,

$$\begin{aligned}
\limsup_{n\to\infty} \|\tilde{\boldsymbol{A}} - \bar{\tilde{\boldsymbol{A}}}\| &= \limsup_{n\to\infty} \|\boldsymbol{D}^{1/2}\boldsymbol{B}\boldsymbol{D}^{-1/2} - \bar{\boldsymbol{D}}^{1/2}\bar{\boldsymbol{B}}\bar{\boldsymbol{D}}^{-1/2}\| \\
&\leq \limsup_{n\to\infty} \| \max\left\{\boldsymbol{D}^{1/2}, \bar{\boldsymbol{D}}^{1/2}\right\}(\boldsymbol{B} - \bar{\boldsymbol{B}}) \max\left\{\boldsymbol{D}^{-1/2}, \bar{\boldsymbol{D}}^{-1/2}\right\} \| \\
&\leq \limsup_{n\to\infty} \max\left\{\|\boldsymbol{D}^{1/2}\|, \|\bar{\boldsymbol{D}}^{1/2}\|\right\} \|\boldsymbol{B} - \bar{\boldsymbol{B}}\| \max\left\{\|\boldsymbol{D}^{-1/2}\|, \|\bar{\boldsymbol{D}}^{-1/2}\|\right\} \\
&\leq \limsup_{n\to\infty} \xi \max\left\{\|\boldsymbol{D}^{1/2}\|, \|\bar{\boldsymbol{D}}^{1/2}\|\right\} \max\left\{\|\boldsymbol{D}^{-1/2}\|, \|\bar{\boldsymbol{D}}^{-1/2}\|\right\}
\end{aligned}$$

Now since ξ can be chosen to be arbitrarily small, it is sufficient to establish that both $\|\boldsymbol{D}^{1/2}\|_2\|\bar{\boldsymbol{D}}^{-1/2}\|_2$ and $\|\boldsymbol{D}^{1/2}\|_2\|\bar{\boldsymbol{D}}^{-1/2}\|_2$ are bounded by a constant almost surely. To see that they are, observe that

$$\|\boldsymbol{D}^{1/2}\|_2\|\bar{\boldsymbol{D}}^{-1/2}\|_2 = \sqrt{\frac{\max_i \sum_j w_{ij}a_{ij}}{\max_i \sum_j w_{ij}p_{ij}}} \leq \sqrt{\frac{\bar{w}_{(n)}}{\underline{w}_{(n)}}} \max_i \sqrt{\frac{\sum_j a_{ij}}{\sum_j p_{ij}}} \leq \sqrt{\omega} \max_i \sqrt{\frac{\sum_j a_{ij}}{\sum_j p_{ij}}}$$

and similarly

$$\|\boldsymbol{D}^{1/2}\|_2\|\bar{\boldsymbol{D}}^{-1/2}\|_2 = \sqrt{\frac{\max_i \sum_j w_{ij}p_{ij}}{\max_i \sum_j w_{ij}a_{ij}}} \leq \sqrt{\frac{\bar{w}_{(n)}}{\underline{w}_{(n)}}} \max_i \sqrt{\frac{\sum_j p_{ij}}{\sum_j a_{ij}}} \leq \sqrt{\omega} \max_i \sqrt{\frac{\sum_j p_{ij}}{\sum_j a_{ij}}}$$

But since the a_{ij} are distributed Bernoulli(p_{ij}), (see, e.g., Mostagir and Siderius (2021)) both $\max_i \sqrt{\frac{\sum_j p_{ij}}{\sum_j a_{ij}}}$ and $\max_i \sqrt{\frac{\sum_j a_{ij}}{\sum_j p_{ij}}}$ converge in probability to 1, so that we have for any $\xi > 0$

$$\limsup_{n\to\infty} \|\tilde{\boldsymbol{A}} - \bar{\tilde{\boldsymbol{A}}}\| \leq \xi\sqrt{\omega}$$

That is, the weighted adjacency matrix can be made arbitrarily close to its expected counterpart.

We now wish to show that, for arbitrary $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr(\|(\boldsymbol{I} - \theta\tilde{\boldsymbol{A}})^{-1} - (\boldsymbol{I} - \theta\bar{\tilde{\boldsymbol{A}}})^{-1}\| \geq \epsilon) = 0$$

The key observation is that the above result implies that for any $\mu > 0$, there exists sufficiently large n such that with probability approaching 1, $\|\tilde{\boldsymbol{A}}^k - \bar{\tilde{\boldsymbol{A}}}^k\| \leq \mu$ for all k. Then it is straightforward to note that (since by model assumptions we have $\theta < 1$, so the formula for infinite geometric series can be applied),

$$\begin{aligned}
\limsup_{n\to\infty} \|(\boldsymbol{I} - \theta\tilde{\boldsymbol{A}})^{-1} - (\boldsymbol{I} - \theta\bar{\tilde{\boldsymbol{A}}})^{-1}\| &= \limsup_{n\to\infty} \left\| \sum_{k=0}^{\infty} \theta^k \left(\tilde{\boldsymbol{A}}^k - \bar{\tilde{\boldsymbol{A}}}^k \right) \right\| \\
&\leq \limsup_{n\to\infty} \sum_{k=0}^{\infty} |\theta^k| \left\| \left(\tilde{\boldsymbol{A}}^k - \bar{\tilde{\boldsymbol{A}}}^k \right) \right\| \\
&\leq \sum_{k=0}^{\infty} \mu|\theta^k| \\
&= \frac{\mu}{1-\theta}
\end{aligned}$$

Since μ was chosen arbitrarily, this implies that for any $\epsilon > 0$,

$$\lim_{n\to\infty} \Pr(\|\boldsymbol{c}^{(n)}(\tilde{\boldsymbol{A}}) - \boldsymbol{c}^{(n)}(\bar{\tilde{\boldsymbol{A}}})\| > \epsilon) = 0$$

Finally, note that $\theta > 0$ and the assumption on the eigenvalues of the Laplacian guarantee that the expected adjacency matrix has non-vanishing spectral gap, implying that the network is connected with high probability, so that the centrality is well-defined (Dasaratha 2020; Mostagir and Siderius 2021), completing the proof. □

Lemma 2 (Chung and Radcliffe (2011)). *Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_m$ be bounded independent random Hermitian matrices and set $M > 0 : ||\boldsymbol{X}_i - \mathbb{E}(\boldsymbol{X}_i)||_2 \leq M \ \forall i = 1, \ldots, m$. Then for any*

$a > 0$,

$$\Pr(||\boldsymbol{X} - \mathbb{E}(\boldsymbol{X})||_2 > a) \leq 2n \exp\left(-\frac{a^2}{2v^2 + 2Ma/3}\right)$$

where $\boldsymbol{X} = \sum_{i=1}^{m} \boldsymbol{X}_i$ *and* $v^2 = ||\sum_{i=1}^{m} \mathbb{V}(\boldsymbol{X}_i)||$.[31]

The following theorem is an extension of Theorem 2 from Chung and Radcliffe (2011), which applies only to unweighted graphs, to symmetrically weighted graphs. Theorem 1 allows us to place tight bounds on the deviation of the realized normalized weighted adjacency matrix from its expected counterpart. A limitation of this result is that the bound depends on n and can thus be arbitrarily loose in large societies.

Theorem 1. *Let* $\mathcal{G}$ *be an undirected random graph such that all edge formation probabilities are jointly independent. Denote by* $\boldsymbol{A}$ *the adjacency matrix,* $\boldsymbol{W}$ *a symmetric matrix of weights, and* $\boldsymbol{A}_W$ *the weighted adjacency matrix, such that* $\boldsymbol{A}_W = \boldsymbol{W} \odot \boldsymbol{A}$.[32] *Let* $\boldsymbol{D}_W$ *be the diagonal degree matrix such that* $\{\bar{\boldsymbol{D}}_W\}_{ii} = \sum_j w_{ij} a_{ij}$, *and denote by* $\bar{\boldsymbol{A}}, \bar{\boldsymbol{D}}_W$ *the expected equivalents. Finally, let* $\boldsymbol{L}_W = \boldsymbol{I} - \boldsymbol{D}_W^{-1/2} \boldsymbol{A}_W \boldsymbol{D}_W^{-1/2}$ *denote the normalized weighted Laplacian of* $\mathcal{G}$, $\omega = \max_{i,j} w_{ij}$ *be the largest total weight,* $\alpha = \min_{i,j} w_{ij}$ *the smallest, and* $\delta = \min_i \{\bar{\boldsymbol{D}}_W\}_{ii}$ *the smallest expected degree.*

For any $\epsilon > 0$, *there exists a* $k > 0$ *such that, for all* i,

$$\Pr\left(||\boldsymbol{L}_W - \bar{\boldsymbol{L}}_W|| \leq 4\sqrt{\frac{3\omega \ln(4n/\epsilon)}{\delta}}\right) \geq 1 - \epsilon$$

if $\underline{d} > k \ln(n)$ *and* $\alpha\omega \leq \sqrt{\frac{\delta}{3\ln(4n/\epsilon)}}$, *where* α *is the smallest total weight.*[33]

Proof of Theorem 1. Denote $\bar{d}_i$ as the expected (weighted) degree of node i. By the triangle inequality, for any matrix $\boldsymbol{C}$,

$$||\boldsymbol{L}_W - \bar{\boldsymbol{L}}_W|| \leq ||\boldsymbol{C} - \bar{\boldsymbol{L}}_W|| + ||\boldsymbol{L}_W - \boldsymbol{C}||$$

[31] See Theorem 5 in Chung and Radcliffe (2011) for the proof.
[32] Note that $\odot$ indicates the Hadamard (element-wise) product.
[33] Note that this is distinct from α in candidate utility.

In particular, let $\boldsymbol{C} = \boldsymbol{I} - \bar{\boldsymbol{D}}_W^{-1/2} \boldsymbol{A}_W \bar{\boldsymbol{D}}_W^{-1/2}$. Then since the degree matrices are diagonal, we have $\boldsymbol{C} - \bar{\boldsymbol{L}}_W = \bar{\boldsymbol{D}}_W^{-1/2}(\boldsymbol{A}_W - \bar{\boldsymbol{A}}_W)\bar{\boldsymbol{D}}_W^{-1/2}$. Denoting by $\boldsymbol{A}^{ij}$ the matrix that is equal to 1 in the i,jth and j,ith positions and 0 elsewhere, we can use the symmetry of weights to write the i,jth entry of $\boldsymbol{C} - \bar{\boldsymbol{L}}$ as

$$\boldsymbol{X}_{ij} = \bar{\boldsymbol{D}}_W^{-1/2}(w_{ij}(a_{ij} - p_{ij})\boldsymbol{A}^{ij})\bar{\boldsymbol{D}}_W^{-1/2} = \frac{w_{ij}(a_{ij} - p_{ij})}{\sqrt{\bar{d}_i \bar{d}_j}} \boldsymbol{A}^{ij}$$

Then clearly $\boldsymbol{C} - \bar{\boldsymbol{L}} = \sum \boldsymbol{X}_{ij}$, so Lemma 2 applies. Since $\mathbb{E}(a_{ij}) = p_{ij}$, we have that $\mathbb{E}(\boldsymbol{X}_{ij}) = \boldsymbol{0}$, so that $v^2 = ||\sum \mathbb{E}(\boldsymbol{X}_{ij}^2)||$. Also, each $\boldsymbol{X}_{ij}$ is bounded above by $\|\boldsymbol{X}_{ij}\| \leq \frac{\omega}{\delta}$ Now clearly

$$\mathbb{E}(\boldsymbol{X}_{ij}^2) = \begin{cases} \frac{w_{ij}^2}{\bar{d}_i \bar{d}_j}(p_{ij})(1 - p_{ij})(A^{ii} + A^{jj}) & i \neq j \\ \frac{w_{ii}^2}{\bar{d}_i^2}(p_{ij})(1 - p_{ij})A^{ii} & i = j \end{cases}$$

Now we can write

$$\begin{aligned} v^2 &= \left\| \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{w_{ij}^2}{\bar{d}_i \bar{d}_j}(p_{ij})(1 - p_{ij})A^{ii} \right\| \\ &= \max_i \left(\sum_{j=1}^{n} \frac{w_{ij}^2}{\bar{d}_i \bar{d}_j}(p_{ij})(1 - p_{ij}) \right) \\ &\leq \max_i \left(\frac{\omega}{\delta} \sum_{j=1}^{n} \frac{w_{ij}}{\bar{d}_i}(p_{ij}) \right) \\ &= \frac{\omega}{\delta} \end{aligned}$$

For notational convenience denote $a = \sqrt{\frac{3\omega \ln(4n/\epsilon)}{\delta}}$ and δ so that $a < 1$. In particular, we must have $\delta > 3\omega(\ln(4) + \ln(n) - \ln(\epsilon))$, so that if $k \geq 3\omega(1 + \ln(4/\epsilon))$, $\delta \geq k \ln(n)$ guarantees the result. Then from Lemma 2

$$\begin{aligned}
\Pr(\|\boldsymbol{C} - \bar{\boldsymbol{L}}_W\| > a) &\le 2n \exp\left(-\frac{\frac{3n\omega^2 \ln(4n/\epsilon)}{n\delta}}{2n\omega^2/\delta + 2an\omega^2/3\delta}\right) \\
&= 2n \exp\left(-\frac{\frac{3n\omega^2 \ln(4n/\epsilon)}{\delta}}{2n\omega^2(3+a)/3\delta}\right) \\
&= 2n \exp\left(-\frac{9\ln(4n/\epsilon)}{6+2a}\right) \\
&\le 2n \exp\left(-\frac{9\ln(4n/\epsilon)}{9}\right) \\
&= \frac{\epsilon}{2}
\end{aligned}$$

Now for the second term, note that d_i is a sum of random variables that are bounded between 0 and ω. Then by Hoeffding's Inequality, we have that, for any t,

$$\Pr(|d_i - \bar{d}_i| > t\bar{d}_i) \le 2\exp\left(-\frac{t^2 \bar{d}_i^2}{n\omega^2}\right) \le 2\exp\left(-\frac{t^2\delta^2}{n\omega^2}\right)$$

Now in particular let $t = \sqrt{\frac{n\omega^2 \ln(4n/\epsilon)}{\delta^2}} = \sqrt{\frac{n\omega}{3\delta}}a$. We have $t < a < 1$ if $\delta > \frac{n\omega}{3}$. In our application, $\omega = \rho_H, \delta = n_0 p_L \rho_L + n_1 p_H \rho_H$., so that for all i we obtain

$$\Pr(|d_i - \bar{d}_i| > t\bar{d}_i) \le \frac{\epsilon}{2n}$$

Now note that

$$\left\|\bar{\boldsymbol{D}}_W^{-1/2}\boldsymbol{D}_W^{1/2} - \boldsymbol{I}\right\|_2 = \max_i \left|\sqrt{\frac{d_i}{\bar{d}_i}} - 1\right|$$

To bound this, note that from (D) we can conclude that $\Pr\left(\left|\frac{d_i}{\bar{d}_i} - 1\right| > t\right) \le \frac{\epsilon}{2n}$ and hence with probability at least $1 - \frac{\epsilon}{2n}$,

$$\left\|\bar{\boldsymbol{D}}_W^{-1/2}\boldsymbol{D}_W^{1/2} - \boldsymbol{I}\right\|_2 < \sqrt{\frac{n\omega^2 \ln(4n/\epsilon)}{\delta^2}}$$

Finally, note that since $\|\boldsymbol{L}\|_2 \leq 2$ (Chung and Graham 1997), we have $\|\boldsymbol{I}-\boldsymbol{L}\|_2 \leq 1$. Now consider

$$
\begin{aligned}
\|\boldsymbol{L}_W - \boldsymbol{C}\| &= \|\boldsymbol{I} - \boldsymbol{D}_W^{-1/2}\boldsymbol{A}_W\boldsymbol{D}_W^{-1/2} - \boldsymbol{I} + \bar{\boldsymbol{D}}_W^{-1/2}\boldsymbol{A}_W\bar{\boldsymbol{D}}_W^{-1/2}\| \\
&= \|(\boldsymbol{I}-\boldsymbol{L}_W)\bar{\boldsymbol{D}}_W^{-1/2}\boldsymbol{D}_W^{1/2}\boldsymbol{D}_W^{-1/2}\boldsymbol{A}_W\boldsymbol{D}_W^{-1/2}\boldsymbol{D}_W^{1/2}\bar{\boldsymbol{D}}_W^{-1/2}\| \\
&= \|(\boldsymbol{I}-\boldsymbol{L}_W)\bar{\boldsymbol{D}}_W^{-1/2}\boldsymbol{D}_W^{1/2}(\boldsymbol{I}-\boldsymbol{L})\boldsymbol{D}_W^{1/2}\bar{\boldsymbol{D}}_W^{-1/2}\| \\
&= \|(\bar{\boldsymbol{D}}_W^{-1/2}\boldsymbol{D}_W^{1/2}-\boldsymbol{I})(\boldsymbol{I}-\boldsymbol{L}_W)\boldsymbol{D}_W^{1/2}\bar{\boldsymbol{D}}_W^{-1/2} + (\boldsymbol{I}-\boldsymbol{L})(\boldsymbol{I}-\boldsymbol{D}_W^{1/2}\bar{\boldsymbol{D}}_W^{-1/2})\| \\
&\leq \|\bar{\boldsymbol{D}}_W^{-1/2}\boldsymbol{D}_W^{1/2}-\boldsymbol{I}\|\|\boldsymbol{D}_W^{1/2}\bar{\boldsymbol{D}}_W^{-1/2}\| + \|\boldsymbol{I}-\boldsymbol{D}_W^{1/2}\bar{\boldsymbol{D}}_W^{-1/2}\| \\
&\leq t^2 + 2t
\end{aligned}
$$

Hence, finally,

$$
\begin{aligned}
||\boldsymbol{L}_W - \bar{\boldsymbol{L}}_W|| &\leq ||\boldsymbol{C}-\bar{\boldsymbol{L}}_W|| + ||\boldsymbol{L}_W - \boldsymbol{C}|| \\
&\leq a + \frac{n\omega}{3\delta}a^2 + \sqrt{\frac{4n\omega}{3\delta}}a \\
&= a\left(\frac{\sqrt{3\delta}+2\sqrt{n\omega}}{\sqrt{3\delta}} + \frac{n\omega}{3\delta}a\right) \\
&= a\left(1 + \frac{2\sqrt{3n\omega\delta}+n\omega a}{3\delta}\right)
\end{aligned}
$$

Now, choose $k > 1$ such that

$$
\delta \geq \frac{1}{3}\left(2n\omega\frac{\sqrt{k}+k+1}{(k-1)^2}\right).
$$

$\square$

Proof of Proposition 1. By the result established in Lemma 1, it is sufficient to consider centrality on the average network. Under the stochastic block model, letting s_i denote the share of i's group without loss of generality, we have that the expected degree of i can be written as

$$\sum_{j=1}^{n} w_{ij}p_{ij} = s_i n w_H p_H + (1-s_i) n w_L p_L = n\left(w_L p_L + s_i(w_H p_H - w_L p_L)\right),^{34}$$

For notational convenience, we denote $w_H p_H = \rho, w_L p_L = \delta\rho$ for some $0 < \delta < 1$. The key observation is that the actual value of ρ is irrelevant, since it appears in both the denominator and numerator of each entry of the expected adjacency matrix. Thus, all results depend only on δ, the *relative* expected weight placed on out-group connections.

Note now that we can write the matrix $\boldsymbol{I} - 2\theta\bar{\tilde{\boldsymbol{A}}}$ as a 2×2 block matrix with blocks $\bar{\tilde{\boldsymbol{A}}}_{11} = \boldsymbol{I} - \frac{2\theta}{n(\delta+s_1(1-\delta))}\left(\mathbf{1}_{s_1 n\times s_1 n} - \boldsymbol{I}\right)$, $\bar{\tilde{\boldsymbol{A}}}_{12} = -\frac{2\theta\delta}{n(\delta+s_1(1-\delta))}\mathbf{1}_{s_1 n\times s_2 n}$, $\bar{\tilde{\boldsymbol{A}}}_{21} = -\frac{2\theta\delta}{n(\delta+s_2(1-\delta))}\mathbf{1}_{s_2 n\times s_1 n}$, and $\bar{\tilde{\boldsymbol{A}}}_{22} = \boldsymbol{I} - \frac{2\theta}{n(\delta+s_2(1-\delta))}\left(\mathbf{1}_{s_2 n\times s_2 n} - \boldsymbol{I}\right)$. To apply the formula for block inversion, we first want to identify $\bar{\tilde{\boldsymbol{A}}}_{11}^{-1}$. We conjecture that

$$P = \bar{\tilde{\boldsymbol{A}}}_{11}^{-1} = \begin{bmatrix} a_1 & b_1 & \cdots & b \\ b & a & \cdots & b \\ \vdots & \cdots & \ddots & \vdots \\ b & \cdots & \cdots & a \end{bmatrix}$$

Then we have that

$$\begin{bmatrix} 1 & -(n_1-1)\frac{\theta}{n(\delta+s_1(1-\delta))} \\ -\frac{\theta}{n(\delta+s_1(1-\delta))} & \left(1-(n_1-2)\frac{\theta}{n(\delta+s_1(1-\delta))}\right) \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

which indeed has a unique solution. The inverse of the bottom-right block is identical, swapping group indices. Hence, we can construct the centrality vector according to the

[34] Technically this is an approximation since $p_{ii} = 0$, but the loss is insignificant for large n, which is assumed here.

formula:

$$\boldsymbol{c} = \begin{bmatrix} \left(\bar{\tilde{\boldsymbol{A}}}_{11} - \bar{\tilde{\boldsymbol{A}}}_{12}\bar{\tilde{\boldsymbol{A}}}_{22}^{-1}\bar{\tilde{\boldsymbol{A}}}_{21}\right)^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \left(\bar{\tilde{\boldsymbol{A}}}_{22} - \bar{\tilde{\boldsymbol{A}}}_{21}\bar{\tilde{\boldsymbol{A}}}_{11}^{-1}\bar{\tilde{\boldsymbol{A}}}_{12}\right)^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{I} & -\bar{\tilde{\boldsymbol{A}}}_{12}\bar{\tilde{\boldsymbol{A}}}_{22}^{-1} \\ -\bar{\tilde{\boldsymbol{A}}}_{21}\bar{\tilde{\boldsymbol{A}}}_{11}^{-1} & \boldsymbol{I} \end{bmatrix} \tag{A1}$$

The main-diagonal blocks in the first matrix again have the same structure, with a single value on the main diagonal and another value on the off-diagonal. This has a similar structure to the previous matrix, and the inverse can thus be calculated analogously by solving for main and off-diagonal elements a_i', b_i'.[35]

Remark. There is a substantive interpretation of a_i' and b_i': a_i' is the weighted average of the number of paths back to a voter in group i through the network, while b_i' is the weighted average of the number of paths to someone else in your group through the network. Because in the expected network, all voters are connected to all others, i's centrality does not depend on paths to the other group, because a linear dependence is induced (all paths within group essentially correspond to an equivalent cross-party path).

Substituting these values into equation (A1), we then have that

$$c_i = (1 + n_{-i})a_i' + (1 + n_{-i})(n_i - 1)b_i'$$

or, without loss of generality,

$$\begin{aligned} c_1 = n\Big(& -\delta\theta - \delta n + (\delta - 1)ns^2(\delta\theta + \delta + \theta - 1) \\ & - (\delta - 1)ns(\delta\theta + \delta + \theta - 1) + (\delta - 1)\theta s\Big) \\ & \cdot \Big(-\theta^2 + n^3 s((\delta - 1)s - \delta)(-\theta + s(\delta + \theta - 1) + 1) \\ & + \theta n^2 s(-\delta - \theta + s(2\delta + \theta - 2) + 1) + \theta n\left(\theta + \delta^2\theta s - \delta s + s - 1\right)\Big)^{-1}. \end{aligned}$$

[35] A unique solution again exists, but we suppress the exact expression as it is extremely complex. The mathematica file used to calculate these values is available upon request from the authors.

Since n is by assumption large, this expression is asymptotically equivalent to its leading term. We can thus simplify further, writing

$$c_1 \sim \frac{-\delta + (\delta - 1)s^2(\delta\theta + \delta + \theta - 1) - (\delta - 1)s(\delta\theta + \delta + \theta - 1)}{sn((\delta - 1)s - \delta)(-\theta + s(\delta + \theta - 1) + 1)}.$$

Simplifying and substituting for ψ yields the required expressions.

□

Proof of Proposition 2. Since we have closed-form expressions for c_1 and c_2 and given assumptions on $u(\cdot)$, it is sufficient to simply take derivatives of total centrality (which is invariant in n) with respect to each parameter, and then substitute these into Equation 3. Since these are highly complex objects, they cannot be signed by inspection, so that we instead solve numerically over the permissible parameter space, yielding the given results. The exact expression for θ^* is a highly complex function of other parameters, but it is straightforward to determine by substitution that it is always at least 0.23 in the permissible parameter space. See the accompanying Wolfram Mathematica code for full derivation. □

Proof of Proposition 3. Since we have closed-form expressions for c_1 and c_2 and given assumptions on $u(\cdot)$, it is sufficient to simply take derivatives of the ratio of centralities (which is invariant in n) with respect to each parameter. Since these are highly complex objects, they cannot be signed by inspection, so that we instead solve numerically over the permissible parameter space, yielding the given results. See the accompanying Wolfram Mathematica code for full derivation. □

Proof of Proposition 4. It is straightforward to see that, replacing θ with Θ, the modified centrality of agents i is now equal to their DeGroot centrality (Mostagir, Ozdaglar, and Siderius 2022) on the normalized network. It then follows from Theorem 1 and from the proof of Theorem 1 in Mostagir and Siderius (2021) that we can again consider the expected network only. The result then follows from an analogous argument to the proof of the Proposition 1, calculated using the accompanying Wolfram Mathematica code. □

Lemma 3. *Define the following terms.*

1. $\Delta_{ij} := \left|\phi_j(\cdot|\boldsymbol{\theta} : \theta_i = \underline{\theta}) - \phi_j(\cdot|\boldsymbol{\theta} : \theta_i = \overline{\theta})\right|$
2. W_{ij}^k *the (weighted) sum of all length k walks beginning with i and ending with j*
3. $\mathcal{I}_{ij} := \frac{\sum_{k=1}^{\infty} \overline{\theta}^k W_{ij}^k}{\sum_h \sum_{k=1}^{\infty} \underline{\theta}^k W_{hj}^k}$

Then, for any sequence of graphs $\mathcal{G}^{(n)}$ such that $(\bar{d}_n/\overline{\theta})^{\text{diam}(\mathcal{G})} = o(n)$, there exists a $z_n > 0$ such that $\Delta_{ij}(\mathcal{G}^{(n)}) < z_n \mathcal{I}_{ij}(\mathcal{G}^{(n)})$ which satisfies $z_n = o(n)$.[36]

Proof. Take a graph $\mathcal{G}$, a corresponding Δ_{ij}, and choose a constant $z_n \in \mathbb{R}$ where

$$z_n > z_n^* := \frac{\Delta_{ij}}{\mathcal{I}_{ij}}.$$

Note that the numerator is bounded by 1 and the denominator must be greater than 0, so it must be that z_n^* is finite. Then, z_n must satisfy $\Delta_{ij} < z_n \mathcal{I}_{ij}$. Moreover, it is clear that $z_n^* < \frac{1}{\mathcal{I}_{ij}}$.[37]

Consider now the denominator of $\mathcal{I}_{ij}$. Since this reflects the sum of all walks of length k ending in j, $\sum_h \sum_{k=1}^{\infty} \underline{\theta}^k W_{hj}^k$, it is straightforward to show by induction on the length of walks k that this must always equal $\underline{\theta}/(1 - \underline{\theta})$.

To see this, note that first when $k = 1$, it is trivially true that there are d_j such walks (starting from each of j's neighbors), each with weight $\frac{1}{d_j}$, so that $\sum_h W_{hj}^1 = 1$.

Now assume that for arbitrary $k \geq 2$, $\sum_h W_{hj}^k = 1$. Then for any walk of length $k + 1$ with start vertex h, the deletion of h from the walk produces a (not necessarily unique) walk

[36] This assumption is consistent with the actual characteristics of "well-behaved" social network graphs, both theoretically and empirically (Jackson 2008).

[37] In fact, this inequality is necessarily extremely loose, since $\Delta_{ij} = 1$ would imply that a change in θ_i moves j from never voting for candidate 1 to voting for them with certainty, which cannot be true by construction.

of length k that begins from some h' adjacent to h. In particular, for each walk w_{hj}^k of length k beginning with vertex h there are d_h walks of length $k+1$, each with weight $\frac{1}{d_h} w_{hj}^k$.[38] The sum of all of these walks can then be expressed

$$\sum_{h' \in \mathcal{T}_h(\mathcal{G})} \frac{1}{d_h} w_{ij}^k = w_{ij}^k.$$

But note that this enumeration is exhaustive; that is, there is no walk of length $k+1$ that cannot be constructed in this fashion, and no two distinct w_{ij}^k can be extended to produce the same w_{ij}^{k+1}. Hence, we have that $W^{k+1} = \sum_i w_{ij}^k = 1$ by the induction assumption.

Moreover, this implies that, regardless of n,

$$\begin{aligned}
\sum_h \sum_{k=1}^{\infty} \underline{\theta}^k W_{hj}^k &= \sum_{k=1}^{\infty} \underline{\theta}^k \sum_h W_{hj}^k \\
&= \sum_{k=1}^{\infty} \underline{\theta}^k \\
&= \frac{\underline{\theta}}{1-\underline{\theta}},
\end{aligned}$$

which implies

$$\mathcal{I}_{ij} = \frac{1-\underline{\theta}}{\underline{\theta}} \sum_{k=1}^{\infty} \overline{\theta}^k W_{ij}^k.$$

Now observe that, defining by $d(i,j)$ the unweighted shortest path distance from i to j, for any $k < d(i,j)$, there by definition exist no walks from i to j of length k, so that clearly $W_{ij}^k = 0$. Then a trivial lower bound on $\mathcal{I}_{ij}$ is $\overline{\theta}^{d(i,j)} W_{ij}^{d(i,j)}$. Moreover, by definition there exists at least one path of length $d(i,j)$, which has minimal weight if all vertices on the path have the maximum degree on the network. Hence, a uniform lower bound for any pair i, j

[38] Note that here w_{hj}^k refers to a particular k-length walk from h to j, which should not be confused with w_{hj}, the weight of j's influence on h's voting behavior from other sections of the paper.

on $\mathcal{G}$ is given by $\left(\overline{\theta}/\bar{d}_n\right)^{\mathrm{diam}(\mathcal{G})}$. Hence, we have that

$$z_n < \frac{\underline{\theta}}{1-\underline{\theta}}\left(\frac{\bar{d}_n}{\overline{\theta}}\right)^{\mathrm{diam}(\mathcal{G})} = o(n)$$

by assumption. □

Lemma 4. *Suppose that $\mathcal{G}$ is generated according to the stochastic block model and satisfies the assumptions of Theorem 1 and of Lemma 3 almost surely. Then Δ_{ij} converges in probability to 0 for all i, j as $n \to \infty$.*

Proof. We aim to bound Δ_{ij} using Lemma 3. Now fix some $\epsilon > 0$. Then it follows from the preceding proof that we want to show that for any i, j, there exists N such that for any $N > n$, the probability that

$$\max_{i,j} \sum_{k=1}^{\infty} \overline{\theta}^k W_{ij}^k > \frac{\underline{\theta}\epsilon}{1-\underline{\theta}} \tag{A2}$$

is arbitrarily small.

Now consider the left-hand side of (A2). Note that it follows from the proof of Lemma 1 that, under the given assumptions, we need only consider the average network, since the sum of walks from i to j is arbitrarily close with probability approaching unity for sufficiently large n. Now note that the expected network, defined as before, is simply a weighted complete graph K_n. Consequently, it is straightforward to enumerate all paths from i to j. We show that for all i, j, k, $\bar{W}_{ij}^K = O(n^{-1})$.

In general, the maximal weight for each connection i, j occurs when $\ell_i = \ell_j$ and for all other h, $\ell_h \neq \ell_i$, so that $w_{ij} < \frac{p_H}{(n-1)p_L}$. Since all walks are at their greatest when all weights are maximal, a (very loose) upper bound for $\bar{W}_{ij}^K$ is the corresponding entry of the k^{th} power of $\frac{p_H}{(n-1)p_L}\left(\mathbf{1}_{n\times n} - \boldsymbol{I}_n\right)$. Now note that

$$\begin{aligned}
\left[\left(\frac{p_H}{(n-1)p_L}\left(\mathbf{1}_{n\times n} - \boldsymbol{I}_n\right)\right)^k\right]_{ij} &= \left[\frac{p_H^k}{(n-1)^k p_L^k}\left(\mathbf{1}_{n\times n} - \boldsymbol{I}_n\right)^k\right]_{ij} \\
&\leq \left[\frac{p_H^k}{(n-1)^k p_L^k}\left(\mathbf{1}_{n\times n}\right)^k\right]_{ij} \\
&= \frac{n^{k-1} p_H^k}{(n-1)^k p_L^k} \\
&= O(n^{-1})
\end{aligned}$$

Hence, $p\text{-}\lim_{n\to\infty} \mathcal{I}_{ij} = 0$, and so by Lemma 3, it follows also that $p\text{-}\lim_{n\to\infty} \Delta_{ij} = 0$, completing the proof.

$\square$