No Place like Home: Opinion Formation with Homophily and Implications for Policy Decisions*

Ali Sina Önder† Marco Portmann‡ David Stadelmann §

Abstract

We demonstrate a simple model of opinion diffusion where a local opinion leader acts as the initiator of public discussion. We show the possibility of driving a significant wedge between opinions of two groups that exhibit homophily even though individuals are highly conformist. In particular, we show that there exists an opinion gap between the group which the opinion leader belongs to (referred to as the *residence community*) and the other group; and this opinion gap is increasing in the relative size of the residence community. Using a unique dataset of national referenda in Switzerland from 2008 to 2012, we show that members of parliament (MPs) match referenda outcomes in their residence communities closer than they do in neighboring communities, and this wedge interacts significantly with the relative size of the residence community, thus aligning with our theoretical conjectures. We conclude that observed opinion gaps can actually be overrated to the extent that they are driven by structures that underlie the social web of different groups within the society.

Keywords: *Opinion Leadership; Diffusion; Homophily; Communication in Networks; Voter Preferences; Representation*
JEL Codes: *D72; D85; H79*

*We thank Ivo Bischoff, John Conley, Rainer Hegselmann, Linda G Veiga, Benno Torgler, and the participants of the 2015 Annual Meeting of the European Public Choice Society and the 2015 Congress of the International Institute of Public Finance for discussion and helpful comments.
†Önder: University of Bayreuth. *Email:* ali-sina.oender@uni-bayreuth.de
‡Portmann: University of Fribourg. *Email:* marco.portmann@unifr.ch
§Stadelmann: University of Bayreuth and CREMA. *Email:* david.stadelmann@uni-bayreuth.de

1 Introduction

In Axelrod's own words, *"it has long been recognized that everyday domestic politics of democracies is largely shaped by the nature of the societal cleavages, [and] an important question is whether such cleavages will be ameliorated or reinforced through local interactions"* (Axelrod, 1997, p. 204). We ask in this paper whether it is safe to assume that all of the observed differences in the public opinions of different groups within a society is driven by differences in deeply rooted ideological opinions within the society. Is it possible that observed opinion differences are mere reflections of differences in how various groups are linked rather than fundamental differences in the world views of these groups? In this paper we use a simple model of opinion diffusion, and we simulate this model to show that a significant opinion gap may be observed even within a society with highly susceptible or conformist individuals and with no propaganda for a competing argument.

A crucial step in understanding public opinion formation is understanding how social connections may influence public opinion. The analysis of opinion formation based on social connections goes back to the seminal works of French (1956), Harary (1959), and DeGroot (1974), where individuals' influence on other members of society has been modeled using weighted adjacency matrices to capture interpersonal connectivity and the direction of the influence.[1] Friedkin and Johnsen (1990, 1999) generalized the repeated opinion updating process of the above-mentioned earlier studies by allowing the sustained influence of individuals' initial opinions. Binary opinions and the discrete updating of opinions following stochastic rules on interaction have been modeled by Clifford and Sudburry (1973) and Holley and Liggett (1975) in the well-known voter model. An important finding of all of the above-mentioned studies is that individuals' opinions tend to become more similar upon interaction.

One of the well-established empirical regularities of social networks is the tendency of *birds of a feather to flock together*, also known as *homophily* in the network literature.[2] Homophily describes the tendency of individuals to have disproportionately more social interaction with people who are like themselves; that is, they have similar backgrounds (e.g., same ethnicity, same gender, same education, etc.) or have similar attachments or interests (e.g., same employment, same place of birth, residing in the same location, etc.). Motivated by the observation that significant differences exist within a society as far as beliefs or attitudes are

[1] See Friedkin and Johnsen (2011) for a detailed discussion on these pioneer models.
[2] Homophily as a term originates from Lazarsfeld and Merton (1954). See McPherson et al., (2001) for a detailed overview and discussion.

concerned, Axelrod (1997) introduced a social influence model where the probability of interaction within a society depends on the cultural proximity of individuals. Hence this is where homophily is introduced into the canonical voter model. More recent studies show that opinion updating via repeated averaging of peers' opinions á la DeGroot (1971) is bound to lead to a convergence of opinions, that homophily alone is not enough to sustain societal cleavages (Golub and Jackson, 2010, 2012; Dandekar et al., 2013), and that any divergence from consensus must be a result of confidence thresholds (Hegselmann and Krause, 2002; Lanchier, 2010a), biased assimilation (Dandekar et al., 2013), the stubbornness of agents (Acemoglu et al., 2013; Friedkin, 2015; Duggins, 2017), xenophobia in highly clustered networks (Flache and Macy, 2011), or the amplification of opinions (Baumgaertner et al., 2016).

Currarini et al. (2009) show that social networks with homophily display three characteristics: members of larger groups have on average more connections per capita, larger groups also have a larger fraction of their links to individuals of their own type, and larger groups form their own types of links at rates greater than their relative fraction in the population. Jackson and Yariv (2007) and Golub and Jackson (2012) establish, in line with the findings of Currarini et al. (2009), that diffusion is faster, and time-to-consensus is shorter, among densely connected individuals. Halberstam and Knight (2014) deliver very interesting empirical evidence by investigating information sharing among Twitter users during the 2012 U.S. presidential election, and they show that individuals are disproportionately exposed to like-minded information. Moreover, members of larger groups are exposed to more tweets (hence more information) per capita than members of smaller groups.

Building on these significant findings we contribute to this literature in three dimensions. First, we start our investigation from the state of consensus, and, using stubborn and influential agents, we simulate how opinions diverge in a society with homophilic connections. This analysis is especially relevant considering nationwide political decisions where citizens are asked to side with one of two possible alternatives and that the formation of an opinion requires some exposure to fundamental information, which is provided by stubborn and politically motivated agents. We show that existence of influential and stubborn agents may lead to societal cleavages even if the society is highly susceptive. Second, we provide simulation results showing that groups' size differences have important implications for the observed opinion differences of those groups in a homophilic society. Third, we take our simulation results to data using a unique dataset on the Swiss referendum outcomes between 2008 and 2012 and obtain empirical findings that sustain our simulation results.

We explore a simple opinion diffusion process where an opinion leader initiates the diffusion process by communicating her opinion to her peers. The group to which the opinion leader belongs agrees more with the opinion leader than other groups do. We call such an opinion gap in favor of the opinion leader's own group *resident opinion leader bias*. Our simulations lead to two conjectures. First, there exists positive resident opinion leader bias under homophily; and second, the resident opinion leader bias depends on the relative size of the two groups.

In the empirical part of our study we explore referenda data to present an interesting empirical case of empirical regularities that fall in line with our simulation results. Referendum outcomes are especially valuable for this investigation, because this is where public opinion manifests itself. We use a unique dataset combining community-level outcomes of national referenda on legislative proposals in Switzerland between 2008 and 2012 and the votes of each member of the parliament (MP) on these proposals. This allows us to investigate how an MP's votes on various legislative proposals compare to the majority of voters' opinions in the community where the MP resides (*residence community*), and to the majority of voters' opinions in its neighbor communities. We find that an MP's vote matches the majority opinion in her residence community in 62.4% of all proposals, whereas this ratio falls to 59.7% for neighbor communities. Thus MPs' roll call behavior matches voters' decisions on average by about 2.7 percentage points better in their residence communities than in neighbor communities. This observation repeats itself consistently throughout various subsamples and robustness checks, thus we establish the *resident MP bias* as an empirical regularity. We further investigate size effects and find that the relative size of the residence community significantly interacts with the resident MP bias in our empirical analysis. These empirical findings fall in line with our simulation results.

In the next section we present a model of opinion diffusion and discuss simulation results. In section 3 we discuss the empirical relevance of the model and present our dataset. We show our empirical findings in section 4, and we offer further discussion in the final section.

2 Diffusion of Opinions Within and Across Groups

We present a basic model of opinion diffusion initiated by an opinion leader in a society that is divided into two groups. Without loss of generality we assume the two groups to be geographically separated and refer to them as *communities*. We further assume that there exists geographic homophily so that individuals in either community have a higher probability to be linked to those residing in the same community than to those residing in the other community. Note that this setting can be generalized to any social group exhibiting homophily (based on ethnicity, gender, place of birth, etc.). There exists an *opinion leader* residing in one of the communities. We refer to this community as the *residence community*, and the other community is called the *neighbor community*, denoted by *r* and *n*, respectively. The opinion leader initiates the opinion diffusion process. Our aim is to isolate the effects of homophily and size differences on opinion diffusion within and across groups, and we do not consider the diffusion of competing opinions or consensus properties in our setting.

2.1 Opinion Formation

Our setting is based on the fundamental structure presented in Friedkin and Johnsen (1999) enhanced with a hierarchical structure as discussed in Lu et al. (2009). The opinion of individual i at time t is denoted as $\alpha_i(t) \in [0,1]$, and this can be interpreted as the probability that individual *i*'s opinion coincides with that of the opinion leader's at time *t*, where time is measured as the number of periods elapsed since the initiation of opinion diffusion by the opinion leader. $\boldsymbol{A}(t) = [\alpha_i(t)]$ denotes the vector of opinions in the society, and opinions are updated according to the following process:

$$\boldsymbol{A}(t) = \boldsymbol{\Omega}\boldsymbol{\Theta}\boldsymbol{A}(t-1) + (I - \boldsymbol{\Omega})\boldsymbol{A}(0), \quad (1)$$

where $\boldsymbol{\Omega}$ is a diagonal matrix with either $\omega_{ii} = 0$ or $\omega_{ii} = 1$, and it captures the stubbornness of individuals. $\boldsymbol{\Theta}$ is an upper triangular matrix with $\boldsymbol{\Theta} = [\theta_{ij}]$ and $\sum_{j \neq i} \theta_{ij} = 1$. θ_{ij}s are weights that individual i assigns to her peers' opinions so that *i*'s opinion in period *t* is a weighted average of her peers' opinions from the previous period. We have $\theta_{ij} \in [0,1)$ for all pairs *ij*.

We employ a stochastic component, namely the probability of link formation between any two given individuals (also called *linking probability*),[3] which is discussed in the next subsection in detail. It is important to note that homophily works through linking probabilities only, that is, an individual's linking probabilities with her own type and different type differ but once a link is established an individual won't discriminate between the opinions of her own

[3] Linking probability originites from the literature on noncooperative network formation games and it denotes the probability that any two randomly chosen nodes in the network are connected by a direct link with one another (or equivalently, they have a geodesic distance of one from each other). See e.g. Jackson (2008), Bala and Goyal (2000), or Slikker and van den Nouweland (2001) for detailed explanations.

type and different type individuals. θ_{ij} takes a strictly positive value if two conditions are fullfilled: θ_{ij} is above the first diagonal of $\boldsymbol{\Theta}$, and the stochastic event of link formation between individuals *i* and *j* has been realized so that *i* and *j* are connected by an edge. Otherwise we have $\theta_{ij} = 0$.

We hold strictly to the textbook definition of homophily that homophily is the increased probability of having a direct link between same-type individuals (Newman, 2010). Jackson and Rogers (2007) show in a detailed model how the observed homophily in linking probabilities is brought about by a repeated process of "meeting friends of friends," which affects the linking probabilities of individuals but not necessarily how they might be influenced by the views of each other. In our setting, homophily works only through linking probabilities in that individuals are more likely to be directly linked to individuals in their own community than those in the other community. Introducing homophily through the influence matrix $\boldsymbol{\Theta}$ is problematic, because this would indicate that individuals assign different weights to information coming from different sources. This is not the case in our model, and individuals assign the same weight to all of their peers whether or not these peers are located in the same community.

We refer to each individual to whom the opinion leader is directly linked as an *influential*. Influentials constitute a fixed proportion of the society, and they can be understood as intermediaries between the opinion leader and the rest of the society; technically they serve as injection points.[4] Influentials may be highly respected and knowledgeable individuals in the society, who provide valuable information relevant for a public discussion topic, or they may be manipulative rent seekers who go out and ruthlessly push for the opinion leader's view. In any case these individuals have strong opinions, they can reach to a larger share of the society than a regular individual can do, and their opinions are taken seriously by the society. In some sense, the opinion leader activates the influentials and sends them out to the society to gain support for her point of view. Watts and Dodds (2007) show that the role of influentials is theoretically marginal in generating cascades, but their model makes use of heterogeneity in how easily individuals get convinced, and this approach is beyond the scope and aim of our paper.

[4] Katz and Lazarsfeld (1955) and Merton (1968) put forward influence models where influentials act as intermediaries. Influentials are assumed to have a larger number of connections than an average citizen, and their opinions are valued for whatever reason by the public.

Influentials are stubborn agents such that $\omega_{ii} = 0$ if individual i is an influential. For any other individual j who is not an influential (we call such members of the society a *regular individual*) we have $\omega_{jj} = 1$, hence j is highly susceptive to others' opinions. We assume that every regular individual has the same initial opinion before the opinion leader initiates the opinion diffusion process. The upper triangular shape of $\boldsymbol{\Theta}$ reveals the hierarchical structure of communication within the society. Rows and columns of $\boldsymbol{\Theta}$ are ordered such that influentials take up the upper left corner of $\boldsymbol{\Theta}$, and regular individuals are placed in the rest of the rows and columns. The upper triangular shape[5] of $\boldsymbol{\Theta}$ is important for the flow of influence, namely between two individuals *i* and *j*—either *i* gets to influence *j*, or *j* gets to influence *i*, but influence does not occur both ways, hence if $\theta_{ij} > 0$ then it must be $\theta_{ji} = 0$. This particular hierarchical structure is useful in emphasizing the role of information in opinion formation, that is, individuals get to hear not only rhetorical arguments from influentials but they receive important information and facts, and they pass these information and facts on to their peers. It is another question whether these facts are objective and true or whether they are rigged on purpose to serve a spesific agenda, but either way they are pieces of information which an individual relies on when updating her opinion. It is important to note that influence is actually equivalent to information provision in this setting.

Opinion updating process runs for two periods, and no influence, i.e. information provision, takes place after two periods. Individuals who are linked to influentials come into contact with information first hand and are able to pass it on to their peers. Individuals, however, who are far away from the source of information are not able to receive convincing amount of information and they are not able to pass it on to their peers. As a result, the diffusion process dies out at this point. This assumption is backed by the empirical finding of Christakis and Fowler (2007), where an individual's influence through her social network is shown to gradually dissipate up to a social distance of three degrees of separation and almost entirely ends beyond that social distance. Additional empirical evidence is provided by Bakshy et al. (2012) who investigate information diffusion in social networks based on an experiment on an online platform and show that individuals who are exposed to information early on are more likely to pass it on sooner. Homophily does not lead to polarization by itself (Dandekar, 2013), and any hierarchical structure where the influence matrix is irreducible will eventually lead to consensus (Lu et al., 2009). To avoid consensus acting as a limiting outcome, opinion updating

[5] This structure of local connections generates a rooted directed spanning tree as shown by Lu et al. (2009) and discussed in detail by Wu (2005).

has to be modeled using, for example, bounded confidence; but, as elaborated in detail by Duggins (2017), this is a rigid approach, and there is not enough empirical evidence at all as to why trust should follow a binary scheme. Homophily, moreover, has a direct effect on how fast consensus is reached (Golub and Jackson, 2012). By halting the process of influence diffusion and opinion updating after two time periods, we aim to focus on opinion differences that are realized along the way to consensus. This is especially empirically relevant when national referenda are considered. In the case of a referendum it is not clear whether the time between the dawn of the idea of a legislative change and the day referendum takes place is enough to reach consensus.

2.2 Geographic Homophily in Linking Probabilities

Linking probabilities make up the stochastic component of our model, and we use them to calculate $\boldsymbol{\Theta} = [\theta_{ij}]$, as will be discussed in this subsection. We denote the population share of influentials in the residence (neighbor) community by π_r^{OL} (π_n^{OL}). Influentials can account for the same proportion of the population in either community, or they can be a larger proportion in the residence community compared to that in the neighbor community. We refer to these two cases as *equal shares of influentials* and *unequal shares of influentials*, respectively. We adopt the following notation to denote linking probabilities between regular individuals: π_{rr} is the linking probability of a regular individual located in the residence community to a regular individual in the same community; π_{rn} is the linking probability of a regular individual located in the residence community to a regular individual in the neighbor community; π_{nn} is the linking probability of a regular individual located in the neighbor community to a regular individual in the same community; and finally π_{nr} is the linking probability of a regular individual located in the neighbor community to a regular individual in the residence community. We further assume that the linking probability of any two randomly chosen individuals is greater if they reside in the same community, thus *geographic homophily* exists.

The linking probability of regular individuals in the same community (*intra-community* linking probability) is assumed to be inversely related to the relative size of that community. The linking probabilitiy of regular individuals in different communities (*inter-community* linking probability) is smaller than the smallest of the two intra-community linking probabilities. Based on these assumptions, we define linking probabilities in the following way:[6] $\pi_{rr} = (1-l_r)^\beta$, $\pi_{nn} = (1-l_n)^\beta$, and $\pi_{rn} = \pi_{nr} = \gamma\min(\pi_{rr}\pi_{nn})$, where β>1, γ∈(0,1), and l_r (l_n) is the relative size of the residence (neighbor) community such that $l_r + l_n = 1$. The linking probability parameter β captures the curvature of the probability function that individuals who are remote from the opinion leader will be exposed to her views. This can be interpreted in the following way: β shows the ability of individuals to convince their peers, or, in other words, β captures how quickly the opinion leader's influence fades away as it diffuses. γ captures the degree of homophily in inter-community linking probabilities.[7]

Since influentials have larger inter- and intra-community linking probabilities compared to regular individuals, we use the following notation whenever we refer to influentials' linking probabilities to regular individuals:[8] π_{rr}^I is the linking probability of an influential located in the residence community to a regular individual in the same community; π_{rn}^I is the linking probability of an influential located in the residence community to a regular individual in the neighbor community; π_{nn}^I is the linking probability of an influential located in the neighbor community to a regular individual in the same community; and π_{nr}^I is the linking probability of an influential located in the neighbor community to a regular individual in the residence community. The linking probabilities of influentials and regular individuals follow patterns similar to those among regular individuals, but these are necessarily larger so as to reflect the well-connectedness of influentials in the society: $\pi_{rr}^I = \pi_{rr}^{1/\beta}$, $\pi_{nn}^I = \pi_{nn}^{1/\beta}$, and $\pi_{rn}^I = \pi_{nr}^I = \pi_{rn}^{1/\beta}$.

Let N_r and N_n denote the population of residence and neighbor communities, respectively. Since the linking probability of two individuals varies only due to their locations, the expected

[6] Functional forms of linking probabilities shown here are chosen in a way such that they embody our basic assumptions on linking probabilities as simple as possible. Aim of this section is not to investigate the exact functional form of linking probabilities but to help visualizing implications of our fundamental assumptions about the patterns of linking probabilities.

[7] γ serves a purely technical purpose. Note that intra-community linking probabilities depend on community size, so that individuals in the smaller community will always have a smaller intra-community linking probability than inter-community, hence geographic homophily holds for individuals in the smaller community even if γ was equal one. The situation is different for individuals in the larger community: if γ was equal one, then there won't be difference between their intra-community and inter-community linking probabilities, which would mean that there is no geographic homophily in linking probabilities of individuals residing in the larger community. Scaling the inter-community linking probability by γ∈(0,1) simply makes sure that geographic homophily holds.

[8] Note that influentials may be linked to each other as well, but since this will not affect their opinion formation in our model, we leave this out for tractability.

number of peers of an individual in either community can be expressed in a uniform way: let d_{rr} and d_{rn} denote the expected number of peers in the residence and neighbor communities, respectively, of a regular individual residing in the residence community. Similarly, let d_{nn} and d_{nr} denote the expected number of peers in the neighbor and residence communities, respectively, of a regular individual residing in the neighbor community. Hence we have

$$d_{rr} = \pi_r^{OL}\pi_{rr}^I(N_r - 1) + (1 - \pi_r^{OL})\pi_{rr}(N_r - 1); \text{ and } d_{rn} = \pi_n^{OL}\pi_{nr}^I N_n + (1 - \pi_n^{OL})\pi_{nr}N_n \quad (2)$$

for individuals residing in the residence community, and

$$d_{nn} = \pi_n^{OL}\pi_{nn}^I(N_n - 1) + (1 - \pi_n^{OL})\pi_{nn}(N_n - 1); \text{ and } d_{nr} = \pi_r^{OL}\pi_{rn}^I N_r + (1 - \pi_r^{OL})\pi_{rn}N_r \quad (3)$$

for individuals residing in the neighbor community.

Assuming an individual gives equal weights to the opinions of her peers, we obtain for every individual residing in the residence community $\theta_{ij} = \theta_r$ and for every individual residing in the neighbor community $\theta_{ij} = \theta_n$, where

$$\theta_r = \frac{1}{d_{rr}+d_{rn}}; \text{ and } \theta_n = \frac{1}{d_{nn}+d_{nr}} \quad (4)$$

so that the weights an individual assigns to her peers' opinions always add up to one, as required in the setting of the equation (1).

A Numerical Example.

Consider a society consisting of 15 individuals, 10 of whom constitute the residence community and the other five constitute the neighbor community. That is, N_r=10 and N_n=5, and the relative sizes of the communities are thus $l_r = 0.67$ and $l_n = 0.33$. Suppose that the share of influentials is 0.2 in both communities (hence $\pi_r^{OL} = \pi_n^{OL} = 0.2$) so that there are two influentials in the residence community and one influential in the neighbor community. We assume β=2 and γ=0.5 so that we obtain the following linking probabilities for the influentials in the society: $\pi_{rr}^I = 0.33$; $\pi_{nn}^I = 0.67$; $\pi_{rn}^I = \pi_{nr}^I = 0.24$; and we obtain the following linking probabilities for regular individuals: $\pi_{rr} = 0.11$; $\pi_{nn} = 0.44$; and $\pi_{rn} = \pi_{nr} = 0.06$. These linking probabilities yield the following expected values for number of peers: $d_{rr} = 2.16$; $d_{rn} = 0.48$; $d_{nn} = 1.94$; $d_{nn} = 1.44$. Hence an individual residing in the residence community is expected to have 2.16 peers in the residence community and 0.48 peers in the neighbor community. An individual residing in the neighbor community is expected to have 1.94 peers in the neighbor community and 1.44 peers in the residence community. It is important that the total number of expected peers of a resident of the neighbor community is larger than that of an individual residing in the residence community. This is driven by the relative size difference of the two communities, which is in line with findings of Currarini et al. (2009). Using the expected numbers of peers we find that $\theta_r = 0.379$ and $\theta_n = 0.296$. An individual residing in the residence community assigns a weight of 0.379 to the views of each of her peers, irrespective of where a respective peer is residing. Similarly, an individual residing in the neighbor community assigns a weight of 0.296 to the views of each of her peers. Hence, as mentioned in the previous subsection, individuals do not discriminate between opinions of their peers based on homophily. Homophily works solely through linking probabilities. Note that all calculations concerning linking probabilities and expected values must be swapped between the two communities if the population numbers are swapped; that is, if N_r=5 and N_n= 10. This is due to the fact that the share of influentials is the same in both communities. In the next subsection we investigate the expected values of the opinions in two communities and how these evolve over different parameter values and with changes in the communities' relative sizes.

2.3 Opinion Difference and Community Size

The baseline level of public opinion is assumed to be α in the residence and neighbor communities, and it remains at this level if the opinion leader does not initiate the opinion diffusion process by communicating her opinion to influentials. We assume that the influentials in the residence (neighbor) community reach a given level of opinion α_r^I (α_n^I) immediately upon their contact with the opinion leader. Individuals who are linked to influentials and individuals who are two degrees away from influentials form their opinions by taking a weighted average of their peers' opinions.[9] Those residing in the residence (neighbor) community who are linked to influentials reach an opinion level of α_r^a (α_n^a) on average, whereas individuals residing in the residence (neighbor) community who are two degrees away from influentials reach an opinion level of α_r^b (α_n^b) on average. All other regular individuals remain at their initial opinion level α.

The average opinion of either community is found by taking the average of existing opinions in the respective community, weighted by the number of individuals holding these opinions. Let S_r^I, S_r^a, S_r^b denote the population shares of influentials (I), individuals one degree away

[9] See the appendix A.1 for the calculation of expected opinions and population shares. We run this analysis using *R*, and our source code is available upon request.

from influentials (a), and individuals two degrees away from influentials (b) residing in the residence community. The average opinion in the residence community is thus given by

$$\alpha_r = \alpha_r^I S_r^I + \alpha_r^a S_r^a + \alpha_r^b S_r^b + \alpha(1 - S_r^I - S_r^a - S_r^b). \qquad (5)$$

Similarly, let S_r^I, S_r^a, S_r^b denote the population shares of influentials (I), individuals one degree away from influentials (a), and individuals two degrees away from influentials (b) residing in the neighbor community. The average opinion in the neighbor community is given by

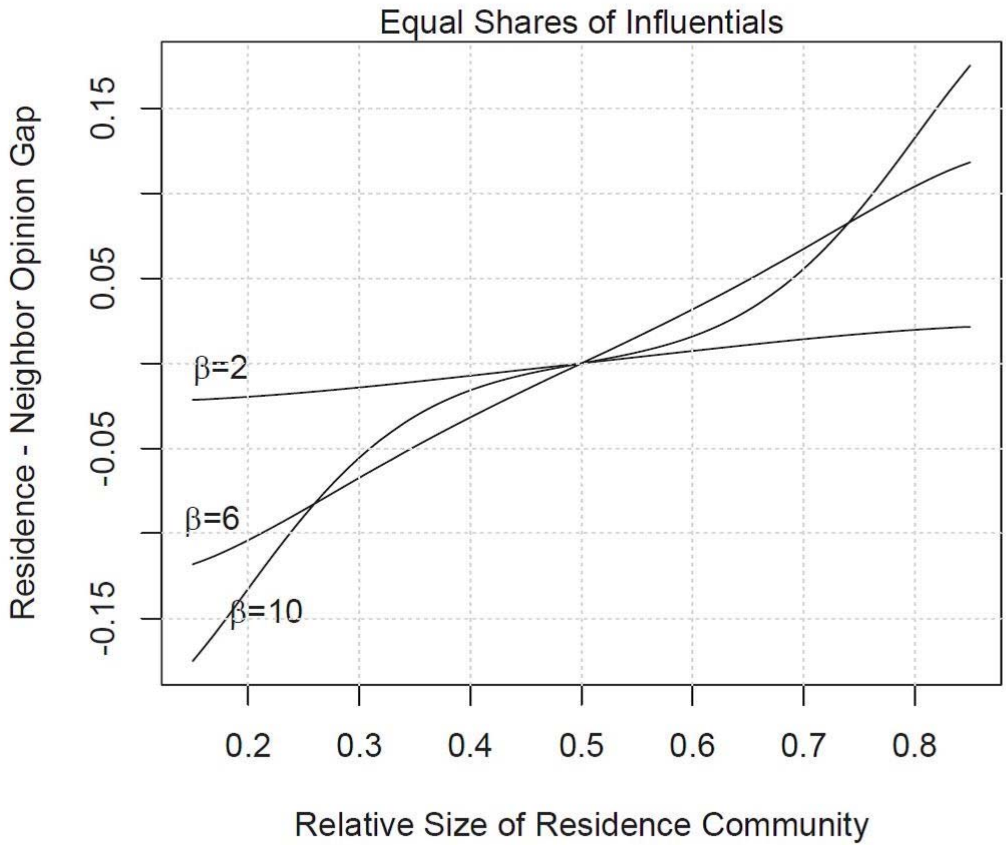
$$\alpha_n = \alpha_n^I S_n^I + \alpha_n^a S_n^a + \alpha_n^b S_n^b + \alpha(1 - S_n^I - S_n^a - S_n^b). \qquad (6)$$

We refer to the difference between these two average opinions, $\alpha_r - \alpha_n$, as the *opinion gap* between the two communities. The size and sign of the opinion gap vary with respect to the linking probabilities and the relative sizes of the communities, thus it is important to understand how the opinion gap changes for different combinations of these key parameters. We simulate opinion gaps under two alternative scenarios: first under the assumption of *equal shares of influentials*, meaning that influentials account for the same proportion of the population in either community, and second under the assumption of *unequal shares of influentials*, meaning that the opinion leader has disproportionately more links to influentials in her own community.

Equal Shares of Influentials.

We assume that shares of influentials are equal in either community, so that $\pi_r^{OL} = \pi_n^{OL} = \max(\pi_{rr}, \pi_{nn})$. Figure 1 shows how the *opinion gap* changes with the relative size of the residence community under the assumption of equal shares of influentials.

Figure 1. Community Size and Opinion Gap-I



We provide simulation[10] results for three different values of the linking probability parameter β in Figure 1. The three results are qualitatively similar: there is no opinion gap when the two communities have the same size. When their sizes are different, then the larger community agrees more with the opinion leader than the smaller community does, and this is independent of where the opinion leader resides. As the linking probability parameter β varies, the opinion gap changes as well. It is, however, not easy to see outright how the opinion gap will change, because the size and direction of the opinion gap depends non-linearly on the wedge between inter-community and intra-community linking probabilities, which in turn is determined by β and the communities' relative sizes. Considering the two specific cases depicted in Figure 1, we observe that a larger opinion gap is obtained for β=6 than that for β=10 when the resident community has a relative size of l_r=0.6; however, this relation is reversed

[10] We take γ=0.8 in our simulations, and other values of γ between zero and one don't yield qualitatively different outcomes. Lower values of γ maintain the overall shape of the curves, making the opinion gap slightly more pronounced when the relative size of the residence community is smaller than 0.3 or greater than 0.7. Alternative figures with different γ values will be made available at request.

when the relative size of the resident community is larger, for example l_r=0.8. The exact value of the opinion gap depends on β; of course, nevertheless, it is important to observe that the overall pattern of the opinion gap does not vary drastically with β.

The basic implications of opinion diffusion under geographic homophily with *equal shares of influentials* (i.e., complete *symmetry* in shares of influentials in the residence and neighbor communities) are summarized in the following conjecture:

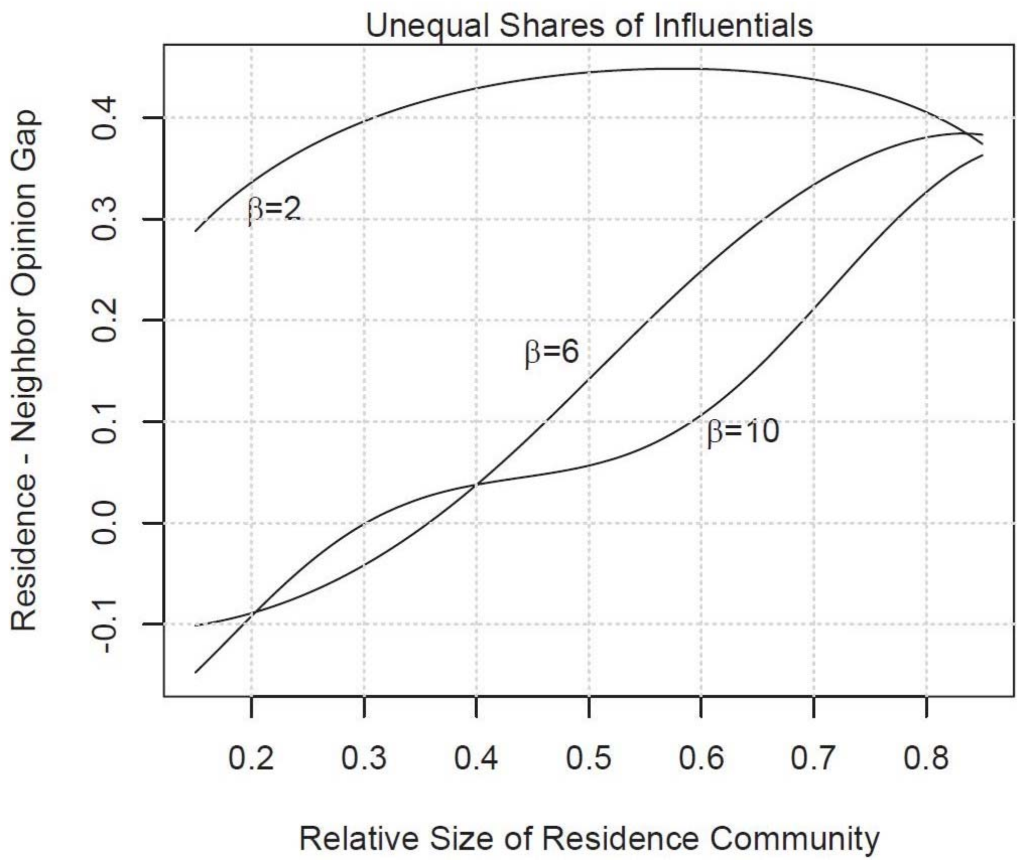
Conjecture 1 *Under geographic homophily with equal shares of influentials, the larger community agrees more with the opinion leader, independent of where the opinion leader resides. The opinion gap increases with the relative size difference. If the two communities have the same size, then no opinion gap exists.*

This conjecture is a direct result of homophily in linking probabilities, which creates social structures similar to those being discussed in the literature of segregation patterns in social networks (e.g., Currarini et al., 2009). Recent research on persistent disagreement under homophily (e.g., Golub and Jackson 2010, 2012) starts from social disagreement and investigates consensus dynamics under homophily, whereas we start from consensus and show how the expected opinions in the two communities can diverge as a result of the structure of inter- and intra-community linking probabilities. If the two communities have the same size, no divergence from consensus is expected. This is simply because the two communities are perfectly symmetric to one another in how the opinion leader's views diffuse through them. If the two communities do not have the same size, then geographic homophily interacts with the relative size difference between the two communities: although influentials constitute the same proportion of the population in both communities, the individuals in the larger community who are one degree away from the influentials have a higher exposure to the influentials' opinions than in the smaller community and update their opinions more heavily (i.e., $\alpha_r^a > \alpha_n^a$). This is the main mechanism driving the opion gap in favor of the larger community. Hence the flow of information becomes amplified in the larger community, and the individuals residing in the larger community are exposed to the opinion leader's views disproportionately more than the individuals in the other community are.

Unequal Shares of Influentials.

An alternative scenario is depicted in Figure 2, where we abandon the assumption of equal shares of influentials in both communities and instead assume unequal shares of influentials; that is, the influentials in the residence community make up a larger proportion of the population in that community than the influentials in the neighbor community do. For this scenario we assume $\pi_r^{OL} = \max(\pi_{rr}, \pi_{nn})$ and $\pi_n^{OL} = \max(\pi_{nr}, \pi_{rn})$. Under this assumption we observe that the residence community agrees more with the opinion leader than the neighbor community does, and this is the case except if the relative size of the residence community is low. For larger values of the linking probability parameter β the relative size of the residence community needs to be larger before a positive opinion gap in favor of the residence community can be observed. We refer to the opinion gap occuring under unequal shares of influentials as the *resident opinion leader bias*. It is important to note that this opinion gap is in favor of the residence community for a wide range of relative community size and parameter values.

Figure 2. Community Size and Opinion Gap-II



With more influentials in the residence community and geographic homophily in the linking probability of individuals, it is no surprise that the opinion leader's views diffuse broader in the residence community so that the residence community tends to agree more with the opinion leader. This leads to a positive resident opinion leader bias even for a small relative size of the residence community. We summarize the implications of opinion diffusion under geographic homophily with *unequal shares of influentials* in the following two conjectures:

Conjecture 2 *There exists a positive resident opinion leader bias under geographic homophily with unequal shares of influentials except for a range of very small relative size of the residence community.*

Conjecture 3 *The resident opinion leader bias increases with the relative size of the residence community for most of its domain, and whether it is increasing at an increasing (convex) or decreasing rate (concave) depends on communities' relative sizes and the linking probability parameter.*

Our findings in this subsection show that observed differences in the opinions of two groups can be driven by the relative size differences of these groups when there is homophily across groups such that individuals are more likely to be linked within their own group. Hence, the main take-home lesson of our simulations is the following: a non-trivial portion of the opinion gap between different groups can be attributed to the structure of social connections within and between these groups as well as to size differences between them and not only to the existence of deeply rooted and competing opinions within the society.

3 Empirical Content of the *Opinion Gap* and Data Description

A plausible setting that would be a good fit for the empirical analysis to find traces of the diffusion process described in the previous section has to satisfy several criteria simultaneously. The data need to contain groups of individuals, including a reasonable candidate for opinion leader, where the groups are clearly separated and yet fairly similar to one another, and the opinions of the different groups and the opinion leader must be measurable—that is, there must be a setting where opinions are revealed. A plausible empirical test that detects traces of the diffusion process described in the previous section must have two aspects: first, it needs to reveal the sign and size of the opinion gap between the groups; and second, it has to capture how the opinion gap interacts with the groups' relative sizes. Such an empirical setting would be less than perfect but a reasonable one by which to investigate whether or not the mechanisms described in the previous section can be empirically relevant.[11]

[11] A perfect setting would be one in which, on top of all mentioned above, individual opinions, individual communication connections, and exchange of opinions are observed.

The existence of an opinion gap between groups (communities) can be a consequence of many different factors; however, an opinion gap that is strictly in favor of the larger group and that increases over some plausible (i.e., empirically relevant) domain of the relative size of the group containing the opinion leader (residence community) hints that the same mechanisms that underlie *conjecture 1* might actually be driving such empirical results. If, on the other hand, it turns out that there exists a mainly positive opinion gap in favor of the group that hosts the opinion leader (namely there exists *resident opinion leader bias*) that increases over some plausible domain of the relative residence community size, then the same mechanisms that underlie *conjecture 2* and *conjecture 3* can be seen as plausible candidates driving such empirical results. It is, however, important to note that our approach does not necessarily provide empirical proof for what exactly drives an opinion gap but rather what *cannot be rejected* as a possible driving factor.

We use a unique dataset combining the votes of the members of the Swiss National Council (the Lower House of the Parliament) on legislative proposals and the results of national referenda in Switzerland between 2008 and 2012. The Swiss National Council consists of 200 MPs who discuss and vote on legislative proposals. Proposals that are agreed upon in the parliament do not automatically become law. If the proposal aims to change the constitution, the citizens have to confirm the change in a national referendum. Proposals that are accepted by the parliament may be challenged in a national referendum if the citizens demand a *facultative* referendum by collecting 50,000 signatures to this end. Citizens also have the right to propose a constitutional amendment upon collecting 100,000 signatures. In such case, MPs vote on the text of the initiative to announce their formal recommendation for the referendum. However, the power to accept or reject the initiative lies solely in the hands of the citizens. A proposal (whether initiated by the National Council or by the citizens) is finally rejected if more than half of the population votes against it.[12]

The voting activities of each MP are recorded and kept by the parliamentary services, and we match these data to another dataset containing the community-level outcomes of national referenda. We draw on the official candidates list of the Swiss federal elections to identify each MP's *residence community*. Hence we observe how an MP votes in the parliament for a legislative proposal and how the majority of the community where she is residing votes in the corresponding referendum on the same proposal. We also collect information on the birthplaces

[12] In addition to the majority of citizens, constitutional amendments require a double majority of citizens and cantons (so called *Ständemehr*).

of MPs and use it in further robustness tests. Voting in the parliament always precedes the popular vote so that the parliamentary vote of each MP about the proposal that is being discussed is revealed well in advance of the referendum. The parliamentary vote of an MP is practically her voting recommendation for the corresponding referendum, and this recommendation cannot be altered (at least, not credibly) later on. This fine-grained data structure allows us to identify how closely preferences expressed by voters coincide with MPs' positions.

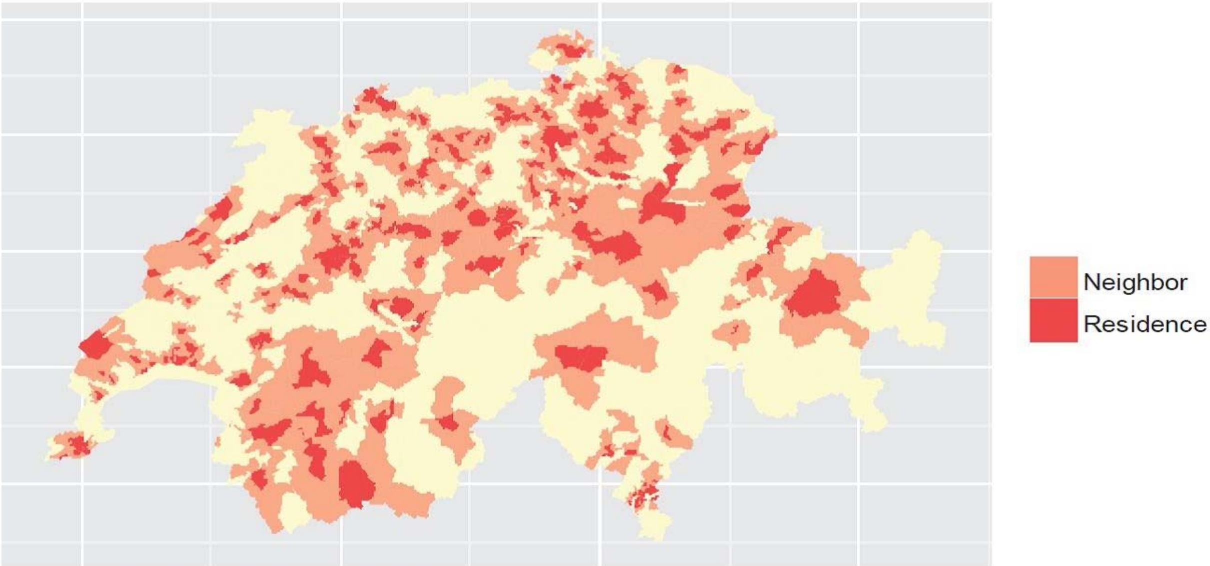
In particular, we compare how an MP's *residence community* matches this MP's position and how communities that share a common border with this *residence community* (referred to as *neighbor communities*) match the same MP's position. Thus, we directly measure whether residence and neighbor communities' revealed opinions in a referendum correspond to MPs' decisions revealed in the parliamentary vote prior to that referendum. We use each MP's vote on a proposal and the outcome of the corresponding referendum in the MP's residence and neighbor communities to create a binary dependent variable, which we refer to as the *opinion match*: if the majority in a community voted yes in a referendum, we take the revealed preference of that community to be in favor of that proposal. If the majority voted no, then the community does not agree with the proposal. The *opinion match* for an MP–community pair for a given proposal is one if both the MP and the community, simultaneously, accepted or rejected the proposal. For proposals where the MPs and communities express different positions the *opinion match* is zero.[13] The binary variable *opinion match* is employed as the dependent variable in our empirical analysis, where the unit of observation is an MP–community–referendum triplet, and community refers to either the MP's *residence community* or *neighbor communities*. A systematic *opinion gap* between the residence and neighbor communities in favor of the residence community will be what we refer to as the *resident opinion leader bias* in the previous section, and we rename it here the *resident MP bias* so as to better convey the setting of our empirical analysis.

Each of the 26 Swiss cantons is an electoral district so that the constituency of each MP is made of voters in their respective canton. Communities are subdivisions of cantons, and an MP's residence community as well as neighbor communities make up a rather small fraction of the MP's constituency,[14] since the size of most communities is often below 3,000 inhabitants. In Figure 3 we map the locations of MPs' residence communities and their neighbors.

[13] MPs' abstentions are excluded from the analysis.

[14] Number of communities contained in each canton and distribution of population in communities are documented in figure A1 and table A1 in the appendix.

Figure 3. Residence Communities of MPs and their Neighbor Communities



Note: Map based on swissBOUNDARIES 2013.

Legislative proposals[15] that are covered in our dataset concern solely national policy with potential differential ramifications at the district (canton) level only, but there are no differential ramifications at the community level for residence and neighbor communities. As a result, an MP's residence community and neighbor communities are expected to be symmetrically affected by the national policy change in question. Since we observe referenda outcomes separately for each community within the same constituency, we are able to investigate the degree of the *opinion match* between MPs and their residence and neighbor communities. Another advantage in focusing on neighbor communities is that we minimize the heterogeneity across communities that are being compared to one another for their opinion match with the local MP. It can rightfully be argued that no two neighboring communities are perfect substitutes for one another, of course; but it can very well be expected that variations originating from geographical, cultural, and locational traits (e.g., being close to an important industrial district, being close to an airport, etc.) are minimized in the case of two neighboring communities as opposed to two randomly chosen communities. One can also argue that these communities differ in their opinions and preferences and that an MP chooses to reside in the

[15] A complete list of proposals and initiatives taken to referenda between 2008 and 2012 (hence covered in our dataset) is shown in table A2 in the appendix.

residence community in the first place, because this community better fits her own opinions and preferences. Nevertheless, our methodology to compare only neighboring communities serves to minimize such differences. Arguing that such differences drive our results would be equivalent to saying that initial opinion and preference differences between the residence community and its neighbor communities must be so severe and rich (going in many different dimensions) that they are carried over to each public debate and reveal themselves in most of the proposals that are being voted upon. We find such wide and deep differences between any two communities that share a common border an extreme case, which cannot meaningfully plague our results.[16]

Divergence between voters' and politicians' preferences has received substantial attention in the literature, and it has been discussed in great detail how MPs' roll call behavior does not fully reflect voters' preferences.[17] We take an alternative stand here and claim that MPs are not merely passive representatives of the median voter's views but that MPs are rather opinion leaders, even ideological and political innovators, who set the tone for and push their opinions in society. An important merit in focusing on Switzerland is that Swiss MPs reside in their initial residential address and are not required to move to the seat of the parliament[18] during their tenure. Thus, an MP is in touch with the voters in her residence community in close proximity, and these voters are likely to be informed about the opinions of the MP better than voters in other communities in the same electoral district. In the same way, voters interact with MPs and get to know their take on national political issues. This does not eliminate a possible reverse causation that MPs who are well-connected within their residence communities will be better informed about what the voters want and hence cater to their will. Nevertheless, the incentives that are needed for the significance of such reverse causation are not apparent in this setting, especially since there is no "pork barrel," as MPs are elected in electoral districts that comprise their residence and the neighboring communities. Hence, it remains a valid argument that MPs act as opinion leaders in their residence communities.

[16] We further rule out such explanations in robustness tests by analyzing MPs who were born in their residence community.

[17] There is rich literature documenting and discussing to what extend politicians fail to represent voters' preferences, and how this can be explained (see Gerber and Lewis 2004, for example). Ågren et al., (2006), Grofman (2004), Matsusaka (2010), Padovano (2013), Portmann et al., (2012), Giger and Klüver (2015), provide an overview of related literature in economics as well as in political science.

[18] The Swiss National Parliament is located in Bern which can be reached from most parts of the country within two to three hours by train.

The data on Swiss national referenda offer a unique opportunity for empirical investigation due to at least five distinct features. First, individuals reveal their preference on policy issues in a referendum, since they are able to answer a decisive and clear-cut question with a *yes* or *no* (e.g., Schneider et al., 1981; Frey 1994, Carey and Hix 2013; Hessami 2016). Second, MPs vote on precisely the same proposals in the parliament as voters do in referenda such that both decisions are directly comparable (see Stadelmann et al., 2013). Third, MPs vote in the parliament before voters get to vote in referenda so that MPs can credibly be claimed to have sufficient time to initiate the local opinion diffusion process. Fourth, MPs actually reside in the communities that they report as their *residence* so that we have good candidates for local opinion leadership in these communities. Fifth, residence and neighbor communities together represent only a small part of an MP's electoral district, and national policies decided in referenda do not have differential effects on the two communities such that a pork barrel situation is ruled out as a possible cause for opinion gaps. We describe the details of this interesting setting in the next subsection and discuss how the resident opinion leader bias may show up in this context.

4 Searching for the *Resident MP Bias*: Residence vs. Neighbor Communities

We use a logit model to regress the binary dependent variable *opinion match* on the indicator variable *residence community,* which distinguishes MPs' residence from neighbor communities, and on fixed effects for referenda and electoral districts (cantons) of MPs such that

$$\textit{Opinion match}=\Lambda(\alpha+\textit{ResidenceCommunity}+\varrho+\mu). \qquad (7)$$

The independent variable *residence community* is one if the *opinion match* refers to the MP–community pair where the community at hand is the MP's residence community, and it is zero if the *opinion match* refers to pairs of the MP and any of the neighbor communities. $\Lambda(.)$ is the logistic function, ϱ denotes referendum fixed effect, and μ denotes canton or MP fixed effect.[19]

Our baseline results are presented in Table 1. In panel A we consider all MP–community–referendum triplets such that each neighboring community of an MP's residence community

[19] Instead of MP fixed effects, we use electoral district fixed (canton) in some estimations. MPs in our sample did not change their electoral district during the time period analyzed.

enters the regression as a separate observation. Coefficient estimations for the logit model where the dependent variable is the *opinion match* are shown in columns (1) to (4), and columns (5) and (6) show the results obtained from the ordinary least squares (OLS) estimations (linear probability models).

Positive and significant coefficients for *residence community* reveal an interesting relation: in every specification across panel A in Table 1, the *opinion match* of MPs and their residence communities is significantly larger than that of MPs and their neighbor communities. Hence, there exists a significantly higher probability that a community agrees with an MP if the MP is residing in that community. The discrete effects of the *residence community* (and its coefficient in the case of linear probability models) in panel A show that the probability that an MP's vote and the majority vote of her residence community coincide is about 3.2 to 3.9 percentage points larger than the probability that an MP's vote and the majority vote of a neighboring community coincide.[20] These effects are not only statistically significant but also economically relevant, considering that three of the last four U.S. presidential elections were decided by a margin of less than four percentage points.

[20] We calculate discrete effects applying the Delta-Method as suggested by Ai et al., (2003).

Table 1. Representation of the Resident Community - Baseline Regressions

Dependent variable: match between MP's vote and referendum outcome

	A. Residence Community vs. All Neighbor Communities					
	(1)	*(2)*	*(3)*	*(4)*	*(5 OLS)*	*(6 OLS)*
Residence Community	0.148***	0.161***	0.160***	0.154***	0.034***	0.032***
	(0.039)	(0.042)	(0.040)	(0.042)	(0.009)	(0.009)
Intercept	0.375***	0.418	0.287	0.653**	0.574***	0.653***
	(0.037)	(0.261)	(0.293)	(0.261)	(0.069)	(0.059)
Referendum FE	no	yes	yes	yes	yes	yes
Canton FE	no	no	yes	no	yes	no
MP FE	no	no	no	yes	no	yes
Discrete Effect: Neighbor to Residence	0.035***	0.038***	0.039***	0.034***		
(Pseudo) R-squared	0.001	0.133	0.14	0.18	0.095	0.125
Observations	53403	53403	53403	53403	53403	53403

	B. Residence Community vs. a Single "Synthetic" Neighbor Community					
	(1)	*(2)*	*(3)*	*(4)*	*(5 OLS)*	*(6 OLS)*
Residence Community	0.147***	0.164***	0.165***	0.174***	0.035***	0.035***
	(0.039)	(0.043)	(0.043)	(0.045)	(0.009)	(0.009)
Intercept	0.377***	0.663***	0.523*	1.322***	0.630***	0.783***
	(0.034)	(0.233)	(0.274)	(0.264)	(0.060)	(0.053)
Referendum FE	no	yes	yes	yes	yes	yes
Canton FE	no	no	yes	no	yes	no
MP FE	no	no	no	yes	no	yes
Discrete Effect: Neighbor to Residence	0.035***	0.036***	0.038***	0.027***		
(Pseudo) R-squared	0.002	0.155	0.161	0.216	0.112	0.151
Observations	11908	11908	11908	11908	11908	11908

Note: Dependent var. is the matching between MP's vote and ref. outcome.
Standard errors (in parenthesis) are clustered by residence-neighbor-pair.
****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

When all neighbor communities are considered separately for an MP, as we do in panel A of Table 1, an important potential problem arises: since there is no control for the size of the communities in the calculation of the *opinion match*, the majority votes of smaller neighbors will be over-represented in the sample. Hence it is possible that the coefficients of the *residence community* are either over-estimated or under-estimated depending on its relative size. A solution to this potential problem and a conservative test for our conjectures is to create *synthetic* neighbor communities. We pool together the citizens of all neighbor communities of an MP's residence community and create a single *synthetic* neighbor community, which also becomes considerably larger in terms of population size than the residence community in most cases.[21] The *opinion match* of the MP and the newly created *synthetic* neighbor community is one if the majority of the pooled population voted in the same way as the MP and zero otherwise.

Table 1's panel B presents the regression results using *synthetic* neighbor communities. The *residence community*'s discrete effect and significance in each of the six specifications are very similar to those in panel A: the discrete effect turns out to be between 3.5 and 3.8 percentage points except for the specification (4) where the discrete effect is 2.7. Thus the opinion gap remains positive and statistically significant even after introducing *synthetic* neighbor communities to take into account any problems related to size difference and aggregation effects.

Our empirical analysis thus far establishes the existence of a statistically significant opinion gap between communities in favor of MPs' residence communities as an empirical regularity.[22] It is important to note that this empirical regularity cannot yet be linked to our opinion diffusion model discussed in the simulation section of this paper, because the relative size of communities needs to be controlled for as well. Before we introduce relative size differences into our empirical analysis, we provide in the next subsection a brief discussion on the robustness of the empirical regularity established above.

4.1 Robustness: Opinion Gap Across Subsamples

We analyze the opinion gap across different subsamples of our dataset to check for the robustness of our findings documented in the previous subsection. The results are shown in Table 2. The *opinion match* of every single neighbor community with the corresponding MP is used as a separate observation in panel A, and we use a single *synthetic* neighbor community for each residence community in panel B. We restrict our analysis in column (1) to the residence communities of MPs that reside in the same community as their birth place. This setting minimizes any possible self-selection of MPs of their residence communities due to differences in preferences, since the cost of moving will still be contained in such a decision process. In

[21] Synthetic communities as well as matching representative's votes with local referenda outcomes are discussed in depth in Hermann and Leuthold (2007).

[22] We provide further evidence on the significance of the opinion gap by using the actual share of votes in residence and neighbor communities that match MP's vote instead of using the binary variable *opinion match*. Although point estimates for the coefficient of *Residence Community* are smaller than those in Table 1, they remain positive and significant. Estimation results are shown in Table A5 in the appendix.

column (3) we look at MPs who hold a local (political) office in their residence community. These two columns represent cases where the alignment of the residence community with the MP can be expected to be more likely due to long-established personal and political relationships. In column (2) we eliminate the neighbor communities in another canton.[23] The coefficients for *residence community* remain statistically significant and positive across these subsamples. The discrete effects when all neighbor communities are included separately are between 3.1 and 3.4 percentage points, and those from the estimation based on synthetic neighbor communities are between 2.5 and 2.9 percentage points. Thus, we obtain not only qualitatively but also quantitatively similar results in cases where the MP possibly has strong local connections compared to those in Table 1.

Column (4) shows a specific subset of referenda such that the time between when MPs vote in the parliament and voters vote in referenda is above the median of such time for all referenda covered in our dataset. When there is more time between the parliamentary and the popular vote, we still obtain a significant and positive coefficient for *residence community* where estimations using separate neighbor communities and a single synthetic neighbor yield discrete effects of 2.6 and 2.7 percentage points, respectively. A plausible explanation for the slight decrease in discrete effects is that the convergence process between communities may have begun to have an effect such that the residence and neighbor communities reach the later stages of the opinion diffusion process so that they are getting closer to a consensus.

[23] Cantons constitute MPs' electoral districts, and it is quite possible that an MP is more responsive to her residence community if the neighbors lie outside the electoral district.

Table 2. Representation of the Resident Community - Robustness

Dependent variable: match between MP's vote and referendum outcome

A. Residence Community vs. All Neighbor Communities

	Residence is Birthplace	*No Border*	*Local Office*	*Long Time before Voting*	*Resident MPs w/same View*	*Neighbor MPs w/same View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Residence Community	0.150**	0.153***	0.134***	0.105***	0.689***	0.172
	(0.060)	(0.043)	(0.039)	(0.037)	(0.149)	(0.134)
Intercept	0.668*	0.629**	0.450	0.292	1.707***	1.205***
	(0.377)	(0.283)	(0.310)	(0.234)	(0.005)	(0.006)
Referendum FE	yes	yes	yes	yes	no	no
Canton FE	no	no	no	no	no	no
MP FE	yes	yes	yes	yes	yes	yes
Discrete Effect: Neighbor to Residence	0.033**	0.034***	0.031***	0.026***	0.070***	0.029
(Pseudo) R-squared	0.195	0.184	0.176	0.120	0.267	0.102
Observations	18202	48148	31317	18028	6265	12445

B. Residence Community vs. a Single "Synthetic" Neighbor Community

	Residence is Birthplace	*No Border*	*Local Office*	*Long Time before Voting*	*Resident MPs w/same View*	*Neighbor MPs w/same View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Residence Community	0.191***	0.160***	0.165***	0.128***	1.319***	0.255
	(0.060)	(0.046)	(0.045)	(0.034)	(0.475)	(0.203)
Intercept	1.366***	1.312***	1.173***	0.755***	-0.074	0.543
	(0.419)	(0.261)	(0.287)	(0.238)	(1.579)	(0.855)
Referendum FE	yes	yes	yes	yes	yes	yes
Canton FE	no	no	no	no	no	yes
MP FE	yes	yes	yes	yes	yes	no
Discrete Effect: Neighbor to Residence	0.029**	0.025**	0.029***	0.027***	0.295	0.057
(Pseudo) R-squared	0.226	0.216	0.200	0.170	0.668	0.545
Observations	3528	11970	6840	4014	1082	2344

Note: Dependent var. is the matching between MP's vote and ref. outcome.
Standard errors (in parenthesis) are clustered by residence-neighbor-pair.
****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

The subsample in column (5) is restricted to residence communities where several MPs reside, and these MPs have voted in the same way for the given legislative proposal, hence they have the same opinion about the subject at hand. The coefficient for *resident community* is very high and highly significant This is an expected outcome, because with several residing MPs sharing the same opinion the opinion gap is expected to be amplified. Finally, we restrict our analysis in column (6) to a subsample where there is at least one MP residing in the neighbor community and moreover, she has (or they have) the same opinion as the MP in the residence community. The coefficient of *resident community* turns out insignificant in this case, as expected, because MPs in either community promote the same view, and as a result we do not observe a significant opinion gap. Additional robustness results based on further subsamples are shown in table A6 in the appendix. These additional robustness checks further support the existence of a significant and economically relevant opinion gap between residence and neighbor communities in favor of the residence community.

In the appendix we provide further robustness checks. Tables A3 and A4 are counterparts of tables 1 and 2 using an alternative depedent variable. The dependent variable in tables A3 and A4 is the actual share of votes in residence and neighbor communities that match MP's vote, e.g. if the MP voted "no" on a proposal and 55% of votes in a community is "no" in the corresponding referendum, then the *percentage match* between the MP and that community is 55%. Similarly, if 48% of a community voted in the same way as the MP then the *percentage match* is 48%. Although point estimates for the coefficient of *Residence Community* are smaller in table A3 than those in table 1, they remain positive and highly significant.

4.2 Relative Community Size and the *Resident MP Bias*

We introduce the size difference between residence communities and their neighbors as an additional independent variable and also control for the interaction between the size difference and the indicator variable for the residence community. This allows us to test the empirical relevance of the three conjectures formulated following our simulations. The regression results are shown in Table 3, panel A, where each neighbor is considered separately, and panel B, where neighbor communities are aggregated into a single *synthetic* community.

Table 3. Resident MP Bias and Size Difference of Resident and Neighbor Communities

Dependent variable: match between MP's vote and referendum outcome

A. Residence Community vs. All Neighbor Communities

	All	*All*	*Local Office*	*Long Time before Voting*	*Resident MPs w/same View*	*Neighbor MPs w/same View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Residence Community	0.201***	0.299***	0.246***	0.220***	0.946***	0.264
	(0.066)	(0.057)	(0.059)	(0.060)	(0.313)	(0.212)
Size Difference (Residence-Neighbor)	-0.048	0.059	0.057	0.058	0.0008	-0.021
	(0.053)	(0.053)	(0.061)	(0.055)	(0.356)	(0.076)
*Residence*SizeDifference*	0.222***	0.265***	0.199**	0.186**	1.123***	0.368
	(0.086)	(0.084)	(0.088)	(0.079)	(0.411)	(0.329)
Intercept	0.306	0.594**	0.394	0.235	1.698***	1.218***
	(0.289)	(0.247)	(0.293)	(0.234)	(0.323)	(0.069)
Referendum FE	yes	yes	yes	yes	no	no
Canton FE	yes	no	no	no	no	no
MP FE	no	yes	yes	yes	yes	yes
Discrete Effect: (3rdQuartile)	0.095***	0.109***	0.096***	0.091***	0.130***	0.091
Discrete Effect: (1stQuartile)	0.053***	0.070***	0.067***	0.057***	0.120***	0.058
(Pseudo) R-squared	0.14	0.181	0.176	0.120	0.268	0.102
Observations	53403	53403	31317	18028	6265	12445

B. Residence Comm. vs. a Single "Synthetic" Neighbor Comm.

	All	*All*	*Local Office*	*Long Time before Voting*	*Resident MPs w/same View*	*Neighbor MPs w/same View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Residence Community	0.253***	0.267***	0.224***	0.167***	1.538***	0.414
	(0.063)	(0.065)	(0.056)	(0.044)	(0.555)	(0.295)
Size Difference (Residence-Neighbor)	-0.049	-0.513	-0.805*	-0.541	-33.593	-0.161
	(0.092)	(0.417)	(0.425)	(0.675)	(21.460)	(0.269)
*Residence*SizeDifference*	0.231**	0.244**	0.172	0.103	1.028	0.566
	(0.098)	(0.102)	(0.112)	(0.078)	(0.999)	(0.476)
Intercept	0.522*	1.298***	1.184***	0.762***	1.743	0.576
	(0.273)	(0.269)	(0.288)	(0.242)	(1.720)	(0.959)
Referendum FE	yes	yes	yes	yes	yes	yes
Canton FE	yes	no	no	no	no	yes
MP FE	no	yes	yes	yes	yes	no
Discrete Effect: (3rdQuartile)	0.058***	0.042***	0.039***	0.036***	0.332	0.097
Discrete Effect: (1stQuartile)	0.019*	0.011*	0.012	0.016*	0.00004	0.002
(Pseudo) R-squared	0.162	0.216	0.200	0.170	0.673	0.546
Observations	11908	11908	6840	4014	1082	2344

Note: Dependent var. is the matching between MP's vote and ref. outcome.
Standard errors (in parenthesis) are clustered by residence-neighbor-pair.
****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

The first two columns of Table 3 display estimation results based on our complete dataset, and columns (3) to (6) correspond to the same numbered columns in Table 2. Residence community remains significant and positive throughout all specifications except for column (6), as expected, where the subsample contains residence communities that have neighbors with an

MP of the same opinion as the MP residing in the residence community. The interaction of the relative size difference and resident community has positive and significant coefficients in all of the five columns where resident community is also significant. Hence the opinion gap in favor of the residence community increases with the relative size difference of residence and neighbor communities: for MPs residing in residence communities of similar size we expect a larger resident MP bias for the residence community that has smaller neighbors. Similarly, for residence communities with neighbor communities of comparable size, we expect a larger resident MP bias for the larger residence community.[24]

According to Table 3, the probability that a residence community agrees with the resident MP is significantly higher than that for its neighbor community if these communities are of equal size—this is namely the baseline effect of the variable *residence community*. When the residence community is smaller than its neighbor, we still expect a positive resident MP bias but not as large as in the case when the two communities are of equal size. A negative total effect of the *residence community* is possible if the residence community is smaller than one-ninth of the neighbor community, according to the coefficients listed in column (1), whereas such a negative total effect is impossible according to the specification used in column (2) as well as according to most of the restricted subsamples investigated in columns (3) to (6). This empirical finding corresponds to the *resident opinion leader bias* discussed in the simulation section under the assumption of *unequal shares of influentials*, and hence we refer to its empirical counterpart as the *resident MP bias*.

Third quartile and first quartile discrete effects show the level of the resident MP bias when the total effect of the variable *residence community* is evaluated at the third and first quartiles, respectively, of the size difference between residence and neighbor communities. Pairs of residence and neighbor communities are ranked from the largest difference in favor of residence community to the largest difference in favor of neighbor community. Using our complete sample as reported in columns (1) and (2), the discrete effect (i.e., the resident MP bias) evaluated at the third quartile of this ranking is about 9.5 to 10.9 percentage points, whereas that evaluated at the first quartile is about 5.3 to 7 percentage points.

When we consider synthetic neighbor communities, the coefficient of the interaction effect between the relative size and the resident community turns out to be significant and positive for the whole sample. The corresponding discrete effects in third and first quartiles in the case of a

[24] As shown in detail in the appendix, using a percentage match variable (table A5) instead of a binary match indicator (table 3) yields qualitatively very similar results.

single synthetic neighbor community range from 4.2 to 5.8 percentage points and from 1.1 to 1.9 percentage points, respectively. No significant interaction effect is obtained for the subsamples that we investigate in columns (3) to (6). The relative size variation of the residence communities that make up our restricted subsamples in columns (3) to (6) most likely does not vary enough to obtain statistical significance for these interaction effects. Nevertheless, the levels of the resident MP bias that are significantly different from zero are 3.6 to 3.9 percentage points at the third quartile and 1.6 percentage points at the first quartile.

We take the existence of the resident MP bias and size effects discussed above to hint at the following: if at least some part of these empirical results is driven by the mechanisms described in the simulation section of this paper, then the *unequal shares of influentials* (as depicted in Figure 3) across the two communities must be the case and not the *equal shares of influentials*, thus confirming geographical homophily in inter-community links, especially including those of the resident MP. Overall, our empirical analysis does not prove that opinion diffusion properties under homophily drive our empirical results; however, the simulation setting corresponding to the case of *unequal shares of influentials* cannot be rejected as a possible driving factor of our empirical results, especially considering that our empirical setting is such that inherent opinion differences between communities are minimized (as explained in detail in the data section) and that MPs' motivation to favor their residence communities (pork barreling, maximizing re-election chances, etc.) is eliminated.

Further results about communities' size differences and interaction with the residence community using various subsamples are shown in Table A7 in the appendix, where similar and robust empirical results are obtained.

5 Discussion and Conclusion

We investigate the diffusion of opinions within and between groups where intra-group and inter-group linking probabilities display homophily, and we model a simple diffusion process of influence. We assume that there exists an opinion leader in one of the two groups and that the opinion leader acts as the initiator of the diffusion process to spread her opinions. We show that relative size differences between groups are always associated with an opinion gap. For a wide range of parameter values we obtain what we call the *resident opinion leader bias*, that is, the group to which the opinion leader belongs (*residence community*) agrees on average more with the opinion leader than the other group does. Moreover, the resident opinion leader bias increases with the relative size of the residence community for most of the empirically relevant range of its relative size.

Taking this conjecture to data requires a highly detailed dataset, where each individual's opinions, personal connections, and communications are revealed. Previous studies have focused on the diffusion of views and news using data based on social media interactions; however, they lack observations on individuals' personal opinions and, in particular, whether these opinions translate into actual behavior. We employ a dataset of Swiss referendum decisions between 2008 and 2012 in our empirical analysis where referenda decisions on legislative proposals are recorded at the community level, the votes of MPs on corresponding proposals are recorded as well, and MPs' residence communities are known. This dataset has the unique advantage of matching the actual decisions of MPs and voters at the community level on exactly the same national policy issues. Such a direct measure of actual political representation at the local level is new in the literature.

Our baseline regressions reveal that the referendum decision in a community where an MP resides tends to agree significantly more with the vote of the MP on the corresponding legislative proposal: the probability that the residence community agrees with the resident MP is about three to four percentage points greater than that of neighbor communities. An analysis of the interaction between communities' relative size and the opinion gap reveals that significant interaction effects exist, the opinion gap is in favor of smaller residence communities, and the opinion gap increases with the relative size of the residence community. We refer to this effect as the *resident MP bias*, which is simply the empirical counterpart to the *resident opinion leader bias* introduced in the simulation section. This finding is statistically robust and economically relevant throughout different subsamples and alternative specifications. Although it is not straightforward to empirically pinpoint all effects observed in our simulations, our empirical investigation reveals patterns that align closely with the conjectures obtained from the simulations, especially *conjecture 2* and *conjecture 3* derived under the assumption of *unequal shares of influentials* (i.e., when the opinion leader is assumed to be better connected within her residence community than in the neighbor community). The phenomenon of residence community bias is empirically highly robust even though politicians are elected in the whole district comprised of numerous communities, and this phenomenon has not been observed in such a direct way before nor has an explanation been offered.

Thus, observed opinion gaps can be driven at least partially by purely technical artifacts that underlie the social web of different groups within society. However, we emphasize that the interpretation of an observed opinion gap requires far more attention if the aim is to identify the channels that lead to the opinion gap and in particular to understand how severe the opinion gap really is due to *fundamental* differences in ideas held by different groups. In a world where social connections reveal significant traces of homophily, and assuming that MPs actually act as opinion leaders and not as representative agents that passively conform to the views of the median voter, MPs need to reach out to the public and promote their opinions and to establish relations with influential individuals in distant communities as well if they want to create a significant change in people's opinions and lead the way.

References

ACEMOGLU, D., COMO, G., FAGNANI, F., OZDAGLAR, A. 2013. Opinion Fluctuations and Disagreement in Social Networks. *Mathematics of Operations Research.* 38(1): 1-27.

AI, C., NORTON E. C. T.S., SHVETS, J. 2003. Interaction Terms in Logit and Probit Models. *Economics Letters.* 80(1): 123-129.

ÅGREN, H., DAHLBERG, M., MÖRK, E. 2006. Do Politicians' Preferences Correspond to those of the Voters? An Investigation of Political Representation. *Public Choice.* 130: 137-162.

AXELROD, R. 1997. The Dissemination of Culture: A Model with Local Convergence and Global Polarization. *Journal of Conflict Resolution.* 41(2): 203-226.

BAKSHY, E., ROSENN, I., MARLOW, C., ADAMIC, L. 2012. The Role of Social Networks in Information Diffusion. In *Proceedings of the 21st international conference on World Wide Web.* pp. 519-528.

BALA, V., GOYAL, S. 2000. A Noncooperative Model of Network Formation. *Econometrica.* 68(5): 1181-1229.

BAUMGAERTNER, B.O., TYSON, R.C., KRONE, S.M. 2016. Opinion Strength Influences the Spatial Dynamics of Opinion Formation. *Journal of Mathematical Sociology.* 40(4): 207-218.

CAREY, J. M., HIX, S. 2013. District Magnitude and Representation of the Majority's Preferences: A Comment and Reinterpretation. *Public Choice.* 154(1-2): 139-148.

CHRISTAKIS, N., FOWLER, J.H. 2007. The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine.* 357: 370-379.

CURRARINI, S., JACKSON, M. O., PIN, P. 2009. An Economic Model of Friendship: Homophily, Minorities, and Segregation. *Econometrica.* 77(4): 1003-1045.

DANEKAR, P., GOEL, A., LEE, D.T. 2013. Biased Assimilation, Homophily, and the Dynamics of Polarization. *Proceedings of the National Academy of Science.* 110(15): 5791-5796.

DEGROOT, M.H. 1974. Reaching a Consensus. *Journal of the American Statistical Association.* 69(345): 118-121.

DUGGINS, P. 2017. A Psychologically-Motivated Model of Opinion Change with Applications to American Politics. *Journal of Artificial Societies and Social Simulation* 20(1).

FLACHE, A., MACY, M.W. 2011. Small Worlds and Cultural Polarization. *Journal of Mathematical Sociology.* 35: 146-176.

FREY, B.S. 1994. Direct Democracy: Polito-Economic Lessons from Swiss Experience. *American Economic Review*. 84: 338-342.

FRIEDKIN, N.E. 2015. The Problem of Social Control and Coordination of Complex Systems in Sociology: A Look at the Community Cleavage Problem. *Control Systems, IEEE*. 35(3): 40-51.

FRIEDKIN, N.E., JOHNSEN, E.C. 1990. Social Influence and Opinions. *Journal of Mathematical Sociology*. 15: 193-206.

FRIEDKIN, N.E., JOHNSEN, E.C. 1999. Social Influence Networks and Opinion Change. *Advances in Group Processes*. 16: 1-29.

FRIEDKIN, N.E., JOHNSEN, E.C. 2011. *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*. Cambridge University Press.

GERBER, E.R., LEWIS, J.B. 2004. Beyond the Median: Voter Preferences, District Heterogeneity, and political Representation. *Journal of Political Economy*. 112(6): 1364-83.

GIGER, N., KLÜVER, H. 2015. Voting Against Your Constituents? How Lobbying Affects Representation. *American Journal of Political Science*. forthcoming.

GOLUB, B., JACKSON, M.O. 2010. Naïve Learning in Social Networks and the Wisdom of Crowds. *American Economic Journal: Microeconomics*. 2(1): 112-149.

GOLUB, B., JACKSON, M.O. 2012. How Homophilly Affects the Speed of Learning and Best-Response Dynamics. *Quarterly Journal of Economics*. 127(3): 1287-1338.

GROFMAN, B. 2004. Downs and Two-Party Convergence. *Annual Review of Political Science*. 7: 25-46.

HALBERSTAM, Y., KNIGHT, B. 2014. Homophily, Group Size, and the Diffusion of Political Information in Social Networks: Evidence from Twitter. *NBER Working Papers*. no.20681.

HEGSELMANN, R., KRAUSE, U. 2002. Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation. *Journal of Artificial Socities and Social Simulations*. 5(3): 1-33.

HERMANN, M., LEUTHOLD, H. 2007. Zwischen Konsens und Polarisierung Volksabstimmungen und Parlamentsentscheide im direkten Vergleich. in Freitag, M., Wagschal, U. (Eds.) *Direkte Demokratie Bestandesaufnahmen und Wirkungen im internationalen Vergleich*, Münster: Lit Verlag, pp.277-302.

HESSAMI Z. 2016. How Do Voters React to Complex Choices in a Direct Democracy? Evidence from Switzerland. *Kyklos*. 69(2): 263-293.

JACKSON, M.O. 2008. *Social and Economic Networks*. Princeton University Press.

JACKSON, M.O., YARIV, L. 2007. Diffusion of Behavior and Equilibrium Properties in Network Games. *American Economic Review*. 97(2): 92-98.

JACKSON, M.O., ROGERS, B.W. 2007. Meeting Strangers and Friends of Friends: How Random are Social Networks? *American Economic Review*. 97(3): 890-915.

KATZ, E., LAZARSFELD, P.F. 1955. *Personal Influence: Part Played by People in the Flow of Mass Communication*. Glencoe, IL: Free Press.

LANCHIER, N. 2010. Opinion Dynamics with Confidence Thresholds: An Alternative to the Axelrod Model. *ALEA*. 7: 1-18.

LAZARSFELD, P.F., MERTON, R.K. 1954. Friendship as a Social Process: a Substantive and Methodological Analysis. in Berger, M. (Ed.) *Freedom and Control in Modern Society*, New York: Van Nostrand, pp.18-66.

LERMAN, K., GOSH, R. 2010. Information Contagion: An Empirical Study of the Spread of News on Digg and Twitter Social Networks. *ICWSM.* 10: 90-97.

LU, J., HO, D.W.C., KURTHS, J. 2009. Consensus over Directed Static Networks with Arbitrary Finite Communication Delays. *Physical Review E.* 80, 066121.

MATSUSAKA, J.G. 2010. Popular Control of Public Policy: A Quantitative Approach. *Quarterly Journal of Political Science.* 5(2): 133-167.

MCPHERSON, M., SMITH-LOVIN, L., COOK, J.M. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology.* 27: 415-444.

MERTON, R.K. 1968. Patterns of Influence: Local and Cosmopolitan Influentials. in *Social Theory and Social Structure.* New York: Free Press.

NEWMAN, M.E.J. 2010. *Networks: An Introduction.* Oxford University Press.

PADOVANO, F. 2013. Are We Witnessing a Paradigm Shift in the Analysis of Political Competition? *Public Choice.* 156: 631-651.

PORTMANN, M., STADELMANN, D., EICHENBERGER, R. 2012. District Magnitude and Representation of the Majority's Preferences: Evidence from Popular and Parliamentary Votes. *Public Choice.* 151: 585-610.

SCHNEIDER, F., POMMEREHNE, W. W., FREY, B. S. 1981. Politico-Economic Interdependence in a Direct Democracy: The Case of Switzerland. In D. A. Hibbs and H. Fassbender (Eds.), Contemporary Political Economy, Studies on the Interdependence of Politics and Economics, pp. 231-248. North Holland, Amsterdam.

SLIKKER, M., VAN DEN NOUWELAND, A. 2001. A One-Stage Model of Link Formation and Payoff Division. *Games and Economic Behavior.* 34: 153-175.

STADELMANN, D., PORTMANN, M., EICHENBERGER, R. 2013. Quantifying Parliamentary Representation of Constituents' Preferences with Quasi-Experimental Data. *Journal of Comparative Economics.* 41(1): 170-180.

WATTS, D.J., DODDS, P.S. 2007. Influentials, Networks, and Public Opinion Formation. *Journal of Consumer Research.* 34(4): 441-458.

WU, C.W. 2005. Synchronization in Networks of Nonlinear Dynamical Systems Coupled via a Directed Graph. *Nonlinearity.* 18: 1057-1064.

A Appendix

A.1 Opinion Diffusion: Population Shares and Opinions of Individuals One or Two Degrees Away from Influentials

Influentials, individuals that are directly linked to influentials (one degree away from influentials) and individuals that are linked to those (two degrees away from influentials) make up the three layers of the society within each community that actually update their opinions. We show in this section how we calculate the opinions of these groups and their shares within each community, because we use these information to calculate the expected opinion in either community and the opinion gap.

Share of influentials in residence (neighbor) community is $S_r^I = \pi_r^{OL}$ ($S_n^I = \pi_n^{OL}$), and we assume they reach an opinion level of α_r^I (α_n^I) upon their contact with the opinion leader.

Share of individuals in residence (neighbor) community that are one degree away from influentials is denoted by S_r^a (S_n^a) and calculated in the following way:

$$S_r^a = (1 - S_r^I)\left(1 - (1 - \pi_{rr}^I)^{\pi_r^{OL} N_r}(1 - \pi_{nr}^I)^{\pi_n^{OL} N_n}\right) \quad \text{(A1)}$$

$$S_n^a = (1 - S_n^I)(1 - (1 - \pi_{nn}^I)^{\pi_n^{OL} N_n}(1 - \pi_{rn}^I)^{\pi_r^{OL} N_r}) \quad \text{(A2)}$$

Expected opinion of an individual in this group is obtained in the following way:

$$\alpha_r^a = \theta_r[N_r S_r^I \pi_{rr}^I \alpha_r^I + N_n S_n^I \pi_{nr}^I \alpha_n^I + N_r(1 - S_r^I - S_r^a)\pi_{rr}\alpha + N_n(1 - S_n^I - S_n^a)\pi_{nr}\alpha] \quad \text{(A3)}$$

$$\alpha_n^a = \theta_n[N_n S_n^I \pi_{nn}^I \alpha_n^I + N_r S_r^I \pi_{rn}^I \alpha_r^I + N_n(1 - S_n^I - S_n^a)\pi_{nn}\alpha + N_r(1 - S_r^I - S_r^a)\pi_{rn}\alpha] \quad \text{(A4)}$$

Share of individuals in residence (neighbor) community that are two degrees away from influentials is denoted by S_r^b (S_n^b) and calculated in the following way:

$$S_r^b = (1 - S_r^I - S_r^a)(1 - (1 - \pi_{rr})^{S_r^a N_r}(1 - \pi_{nr})^{S_n^a N_n}) \quad \text{(A5)}$$

$$S_n^b = (1 - S_n^I - S_n^a)(1 - (1 - \pi_{nn})^{S_n^a N_n}(1 - \pi_{rn})^{S_r^a N_r}) \quad \text{(A6)}$$

Expected opinion of an individual in this group is obtained in the following way:

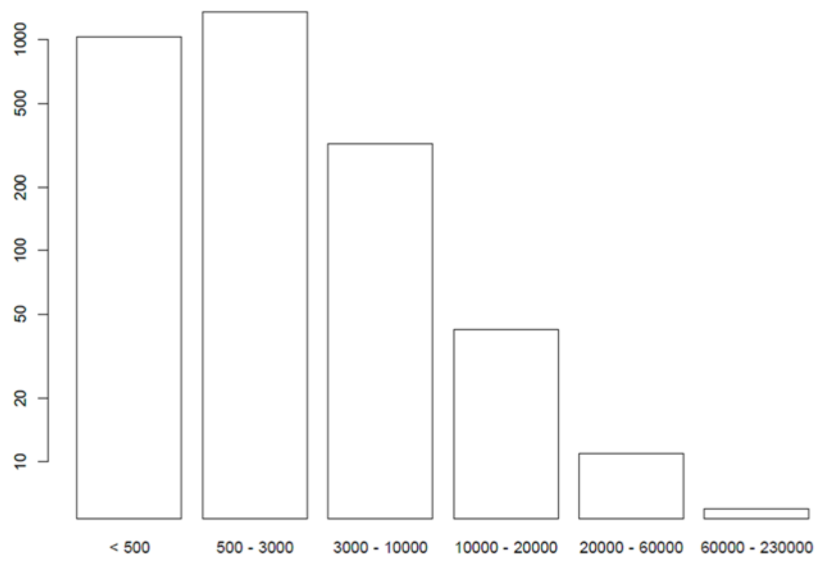
$$\alpha_r^b = \theta_r[N_r S_r^a \pi_{rr}^a \alpha_r^a + N_n S_n^a \pi_{nr} \alpha_n^a + N_r\left(1 - S_r^I - S_r^a - S_r^b\right)\pi_{rr}\alpha + N_n\left(1 - S_n^I - S_n^a - S_n^b\right)\pi_{nr}\alpha] \quad \text{(A7)}$$

$$\alpha_n^b = \theta_n[N_n S_n^a \pi_{nn} \alpha_n^a + N_r S_r^a \pi_{rn} \alpha_r^a + N_n\left(1 - S_n^I - S_n^a - S_n^b\right)\pi_{nn}\alpha + N_r(1 - S_r^I - S_r^a - S_r^b)\pi_{rn}\alpha] \quad \text{(A8)}$$

A.2 Background Information: Swiss Cantons, Municipalities, and Referenda

Figure A1. Distribution of Community Size in Switzerland

A. Number of Swiss Communities, grouped according to their Population



B. Number of Observations on Communities (grouped according to their Population) with at least one resident MP

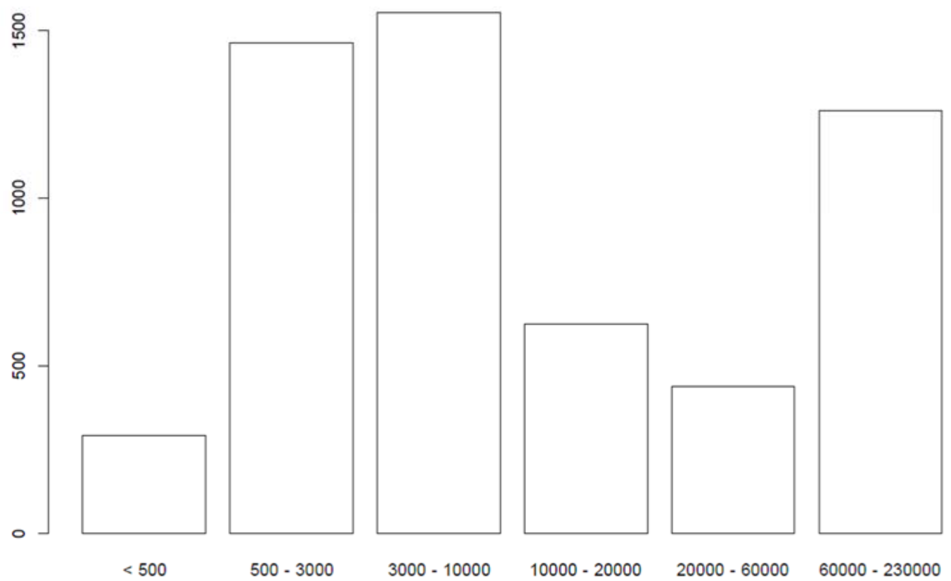


Table A1. Swiss Cantons and Municipalities

Canton	*Number of municipalities*	*Canton*	*Number of municipalities*
Aargau	225	*Nidwalden*	11
Appenzell Innerrhoden	6	*Obwalden*	7
Appenzell Ausserrhoden	20	*St. Gallen*	86
Bern	378	*Schaffhausen*	28
Basel-Landschaft	86	*Solothurn*	123
Basel-Stadt	3	*Schwyz*	30
Fribourg	167	*Thurgau*	80
Geneva	45	*Ticino*	168
Glarus	19	*Uri*	20
Graubünden	186	*Vaud*	364
Jura	67	*Valais*	144
Lucerne	89	*Zug*	11
Neuchâtel	54	*Zurich*	171
		Total	2588

Note: The number of Swiss municipalities has strongly decreased during the last decades due to municipal mergers. The table indicates the average number of municipalities in the respective canton within the analyzed sample of referenda from 2008 to 2012.

Table A2. List of Referenda Topics 2008-2012

ID	Date	Proposal / Topic of the Referendum
530	24.02.2008	Initiative "against fighter jet noise in tourist regions"
531	24.02.2008	Federal law on the improvment of the tax environment for business activities and investments
532	01.06.2008	Initiative "for democratic naturalizations"
533	01.06.2008	Initiative "Popular sovereignity instead of authority propaganda"
534	01.06.2008	Constitutional article "For quality and efficiency in health insurance"
535	30.11.2008	Initiative "for the abolition of the statute of limitations regarding pornography offences involving children"
536	30.11.2008	Initiative "for a flexible AHV-age"
537	30.11.2008	Initiative "Right of appeal for associations: Stop the obstruction policy - More growth for Switzerland!"
538	30.11.2008	Initiative "for a reasonable hemp-policy including an effective protection of minors"
539	30.11.2008	Federal law on narcotics and psychotropic substances
540	08.02.2009	Federal enactment on the approval of the continuation of the Agreement on the Free Movement of Persons between Switzerland and the European community as well as on the approval of the protocol on the extension of the Agreement to Rumania and Bulgaria
541	17.05.2009	Constitutional article "Future with complementary medicine" (Counter proposal to the initative "Yes to complementary medicine")
542	17.05.2009	Federal enactment on the approval and the implementation of the exchange of notes between Switzerland and the European community relating to the adoption of regulations for biometrics in passports and travel documents
543	27.09.2009	Federal enactment on temporary supplementary funding of the invalidity insurance through an increase in the value added tax
544	27.09.2009	Federal enactment on the waiving of the introduction of the general people's initiative
545	29.11.2009	Federal enactment to create special funding for tasks related to air traffic
546	29.11.2009	Initiative "for a ban on the export of war material"
547	29.11.2009	Initiative "against the construction of minarets"
548	07.03.2010	Constitutional article on research on humans
549	07.03.2010	Initative "against animal cruelty and for better legal rights for animals"
550	07.03.2010	Federal law on occupational retirement, surviving dependants' and disability pension (BVG) (minimum conversion rate)
551	26.09.2010	Federal law on obligatory unemployment insurance and insolvency compensation (unemployment insurance act (AVIG))
553	28.11.2010	Federal enactment on the initative "In favor of fair taxes. Stop tax competition abuse (initative for tax justice)" of 18 June 2010
5521	28.11.2010	Federal enactment on the initative "for the expulsion of foreign criminals (expulsion initative)"
5522	28.11.2010	Federal enactment on the deportation of foreign criminals within the framework of the Federal Constitution (Counter proposal to initative "For the expulsion of foreign criminals (expulsion initative)" of 10 June 2010
554	13.02.2011	Initative "Protection against firearms violence"
555	11.03.2012	Federal enactment on the initative "Stop the excessive construction of secondary homes"
556	11.03.2012	Initative "In favor of tax-supported building society savings for the purchase of owner-occupied residential property and for the financing of constructional energy-saving and environmental measures (building society savings initative)" of 29 September 2008
557	11.03.2012	Federal enactment on the initative "6 weeks vacation for everybody"
558	11.03.2012	Federal enactment on the regulation of betting games in favor of charitable purposes (Counter proposal to the initative "For betting games serving the common good" of 29 September 2011
559	11.03.2012	Fixed Book Price Law (BuPG) of 18 March 2011
560	17.06.2012	Initative "Owning a home thanks to building society savings"
561	17.06.2012	Initative "Reinforcing popular rights in foreign policy (Let the people decide on treaties!)
562	17.06.2012	Amendment to the federal law on health insurance (KVG) (Managed Care)
563	23.09.2012	Federal enactment on the promotion of youth music (Counter proposal to the initative "youth + music")
564	23.09.2012	Initative "Safe housing for the elderly"
565	23.09.2012	Initative "Protection from exposure to tobacco smoke"
566	25.11.2012	Amendment of 16. March 2012 to the Law on Epizootic Diseases

A.3 Further Robustness Tests

We provide estimation results using the percentage match between the MP’s vote and the referanda outcome as dependent variable instead of binary match we used in the analysis shown in tables 1, 2, and 3 to insure that our results are robust to such alternative definitions of the opinion match. The dependent variable in tables A3, A4, and A5 is the actual share of votes in residence and neighbor communities that match MP's vote, e.g. if the MP voted "no" on a proposal and 55% of votes in a community is "no" in the corresponding referendum, then the *percentage match* between the MP and that community is 55%. Similarly, if 48% of a community voted in the same way as the MP then the *percentage match* is 48%. It is important to note that the *opinion match* (the dependent variable used in Table 1 in the main text) in the former case would be 1 whereas it would be 0 in the latter case. Although point estimates for the coefficient of *Residence Community* are smaller in Table A3 than those in Table 1, they remain positive and highly significant. Similarly, results shown in tables A4, and A5 yield qualitatively and quantitatively comparable results as results based on binary matching results presented in the main text (tables 2 and 3).

Moreover, we present additional evidence for robustness using additional subsamples in tables A6 and A7, which complement tables 2 and 3 in the main text, respectively. First, we briefly describe the content of subsample restrictions in tables A6 and A7. The subsample labeled *tight referendum* is restricted to those referenda that are decided by only a small margin nationwide. Other subsamples are restricted to MPs' residence communities which:
- are birthplace of the resident MP, *residence is birthplace* (to ensure exogeneity of residency);
- talk a different language than some of their neighbor communities, *other language* (to exploit another potential source of homophily);
- are rural, *rural community* (to ensure results are not due to cities);
- contain a resident MP that is a farmer, *MP is farmer* (to ensure exogeneity of residency);
- are also the official correspondance adress of the resident MP, *residence=correspond.* (to ensure that MPs do not only indicate residency in a community);
- have several resident MPs with the same opinion AND there is no MP residing in the neighbor communities, *resident MPs w/same view and no MP in neighbor* (to isolate potential disturbance from MPs in neighbor communities);
- have several resident MPs with the same opinion AND there are MPs residing in the neighbor communities that have the opposite view as the MPs in the residence community, *resident MPs w/same view and MPs in neighbor w/opposite view* (to analyze a situation where strong resident MP bias should be present).

Table A3. Representation of the Resident Community - -using Percentage Match

A. Residence Community vs. All Neighbor Communities

	(1)	*(2)*	*(3)*	*(4)*
Residence Community	0.0124***	0.0124***	0.0124***	0.0119***
	(0.004)	(0.004)	(0.003)	(0.003)
Intercept	0.542***	0.561 ***	0.547 ***	0.577***
	(0.005)	(0.026)	(0.029)	(0.026)
Referendum FE	no	yes	yes	yes
Canton FE	no	no	yes	no
MP FE	no	no	no	yes
(Pseudo) R-squared	0.0005	0.165	0.174	0.217
Observations	53403	53403	53403	53403

B. Residence Community vs. a Single "Synthetic" Neighbor Community

	(1)	*(2)*	*(3)*	*(4)*
Residence Community	0.0126***	0.0126***	0.0126***	0.0126***
	(0.003)	(0.003)	(0.003)	(0.003)
Intercept	0.542***	0.585***	0.571***	0.628***
	(0.003)	(0.019)	(0.022)	(0.02)
Referendum FE	no	yes	yes	yes
Canton FE	no	no	yes	no
MP FE	no	no	no	yes
(Pseudo) R-squared	0.0013	0.189	0.195	0.248
Observations	11908	11908	11908	11908

Note: Dependent var. is the percentage match between MP's vote and ref. outcome.

Standard errors (in parenthesis) are clustered by residence-neighbor-pair.

****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

Table A4. Representation of the Resident Community – Robustness – using Percentage Match

A. Residence Community vs. All Neighbor Communities

	Residence is Birthplace	*No Border*	*Local Office*	*Long Time before Voting*	*Resident MPs w/same View*	*Neighbor MPs w/same View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Residence Community	0.012**	0.011***	0.011***	0.013***	0.048***	0.015
	(4.8e-03)	(3.5e-03)	(3.3e-03)	(4.2e-03)	(7.5e-03)	(0.011)
Intercept	0.574***	0.574***	0.556***	0.532***	0.652***	0.624***
	(0.035)	(0.027)	(0.029)	(0.022)	(3.7e-04)	(5.4e-04)
Referendum FE	yes	yes	yes	yes	no	no
Canton FE	no	no	no	no	no	no
MP FE	yes	yes	yes	yes	yes	yes
R-squared	0.246	0.223	0.213	0.112	0.228	0.107
Observations	18202	48148	31317	18028	6265	12445

B. Residence Community vs. a Single "Synthetic" Neighbor Community

	Residence is Birthplace	*No Border*	*Local Office*	*Long Time before Voting*	*Resident MPs w/same View*	*Neighbor MPs w/same View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Residence Community	0.015***	0.012***	0.012***	0.014***	0.047***	0.015
	(4.2e-03)	(3.1e-03)	(3.2e-03)	(3.7e-03)	(6.3e-03)	(9.8e-03)
Intercept	0.627***	0.627***	0.610***	0.589***	0.659***	0.648***
	(0.029)	(0.020)	(0.023)	(0.018)	(3.1e-03)	(4.9e-03)
Referendum FE	yes	yes	yes	yes	yes	yes
Canton FE	no	no	no	no	no	yes
MP FE	yes	yes	yes	yes	yes	no
R-squared	0.282	0.246	0.238	0.151	0.295	0.152
Observations	3528	11970	6840	4014	1082	2344

Note: Dependent var. is the percentage match between MP's vote and ref. outcome.

Standard errors (in parenthesis) are clustered by residence-neighbor-pair.
****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

Table A5. Resident MP Bias and Size Difference of Resident and Neighbor Communities –using Percentage Match

A. Residence Community vs. All Neighbor Communities

	All	*All*	*Local Office*	*Long Time before Voting*	*Resident MPs w/same View*	*Neighbor MPs w/same View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Residence Community	0.017***	0.024***	0.023***	0.029***	0.055***	0.025
	(6.3e-03)	(4.9e-03)	(4.9e-03)	(5.9e-03)	(0.016)	(0.016)
Size Difference (Residence-Neighbor)	-2.3e-03	5.4e-03	5.5e-03	8.1e-03*	4.2e-03	2.4e-03
	(5.2e-03)	(3.8e-03)	(5.1e-03)	(4.2e-03)	(0.022)	(6.3e-03)
*Residence*SizeDifference*	0.019**	0.021***	0.021***	0.026***	0.035*	0.029
	(7.4e-03)	(7.0e-03)	(6.4e-03)	(8.9e-03)	(0.019)	(0.025)
Intercept	0.549***	0.572***	0.550***	0.524***	0.648***	0.622***
	(0.028)	(0.025)	(0.027)	(0.021)	(0.020)	(5.7e-03)
Referendum FE	yes	yes	yes	yes	no	no
Canton FE	yes	no	no	no	no	no
MP FE	no	yes	yes	yes	yes	yes
R-squared	0.174	0.217	0.213	0.113	0.228	0.108
Observations	53403	53403	31317	18028	6265	12445

B. Residence Comm. vs. a Single "Synthetic" Neighbor Comm.

	All	*All*	*Local Office*	*Long Time before Voting*	*Resident MPs w/same View*	*Neighbor MPs w/same View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Residence Community	0.019***	0.019***	0.018***	0.021***	0.049***	0.025*
	(4.5e-03)	(4.5e-03)	(3.8e-03)	(5.7e-03)	(6.2e-03)	(0.014)
Size Difference (Residence-Neighbor)	6.1e-04	-0.084***	-0.114***	-0.078	-0.067	-0.120***
	(7.9e-03)	(0.032)	(0.031)	(0.056)	(0.441)	(0.026)
*Residence*SizeDifference*	0.016**	0.016**	0.016**	0.020**	0.024	0.034
	(7.4e-03)	(7.4e-03)	(7.7e-03)	(9.3e-03)	(0.015)	(0.023)
Intercept	0.573***	0.628***	0.614***	0.589***	0.545***	0.648***
	(0.022)	(0.020)	(0.023)	(0.018)	(0.089)	(7.2e-03)
Referendum FE	yes	yes	yes	yes	yes	yes
Canton FE	yes	no	no	no	no	yes
MP FE	no	yes	yes	yes	yes	no
R-squared	0.196	0.249	0.24	0.152	0.737	0.156
Observations	11908	11908	6840	4014	1082	2344

Note: Dependent var. is the percentage match between MP's vote and ref. outcome.

Standard errors (in parenthesis) are clustered by residence-neighbor-pair.

****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

Table A6. Representation of the Resident Community - Further Robustness Tests

A. Residence Community vs. All Neighbor Communities

	Other Language	*Rural Community*	*MP is Farmer*	*Residence = Correspond.*	*Tight Referendum*	*Resident MPs w/same View and NO MP in Neighbor*	*Resident MPs w/same View and MPs in Neighbor w/opposite View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Residence Community	0.117	0.105**	0.085	0.102*	0.206*	1.548**	0.944***
	(0.086)	(0.048)	(0.088)	(0.055)	(0.113)	(0.621)	(0.301)
Intercept	-0.645	1.905***	1.872***	0.345	2.059***	9.874***	-1.433***
	(0.575)	(0.431)	(0.527)	(0.347)	(0.168)	(2.198)	(0.036)
Referendum FE	yes	yes	yes	yes	yes	yes	no
Canton FE	no	no	no	no	no	no	no
MP FE	yes	yes	yes	yes	yes	yes	yes
(Pseudo) R-squared	0.25	0.31	0.38	0.19	0.271	0.74	0.26
Observations	3280	11968	6669	18669	6552	1628	1881

B. Residence Community vs. a Single "Synthetic" Neighbor Community

	Other Language	*Rural Community*	*MP is Farmer*	*Residence = Correspond.*	*Tight Referendum*	*Resident MPs w/same View and NO MP in Neighbor*	*Resident MPs w/same View and MPs in Neighbor w/opposite View*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
Residence Community	0.155	0.146**	0.166	0.113*	0.374**	2.829**	1.168**
	(0.112)	(0.065)	(0.110)	(0.063)	(0.155)	(1.227)	(0.460)
Intercept	-0.211	1.309***	1.719***	1.521***	12.802***	9.090***	-1.768***
	(0.640)	(0.472)	(0.616)	(0.361)	(1.030)	(2.418)	(0.298)
Referendum FE	yes	yes	yes	yes	yes	yes	no
Canton FE	no	no	no	no	no	no	no
MP FE	yes	yes	yes	yes	yes	yes	yes
(Pseudo) R-squared	0.30	0.36	0.42	0.24	0.466	0.85	0.42
Observations	780	3174	1666	4138	1466	316	362

Note: Dependent var. is the matching between MP's vote and ref. outcome. Standard errors (in parenthesis) are clustered by residence-neighbor-pair.
****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

Table A7. Resident MP Bias and Size Difference of Resident and Neighbor Communities- Further Subsamples

A. Residence Community vs. All Neighbor Communities	*Residence is Birthplace* (1)	*Other Language* (2)	*Rural Community* (3)	*MP is Farmer* (4)	*Residence = Correspond.* (5)	*Tight Referendum* (6)	*Resident MPs w/same View and NO MP in Neighbor* (7)	*Resident MPs w/same View and MPs in Neighbor w/opposite View* (8)
Residence community	0.219***	0.274**	0.121	0.732**	0.275***	0.560***	2.711**	1.393**
	(0.067)	(0.129)	(0.119)	(0.368)	(0.104)	(0.198)	(1.079)	(0.589)
Size Difference (Residence-Neighbor)	0.056	0.027	0.048	0.336**	0.081	0.242	0.788	-0.132
	(0.082)	(0.120)	(0.061)	(0.133)	(0.083)	(0.149)	(0.741)	(0.602)
*Residence*SizeDifference*	0.121	0.377*	-0.037	0.541	0.264*	0.417	2.417	1.927***
	(0.125)	(0.205)	(0.182)	(0.585)	(0.149)	(0.295)	(1.507)	(0.581)
Intercept	0.615*	-0.620	1.921***	1.782***	0.254	1.833***	9.012***	-1.400***
	(0.356)	(0.574)	(0.432)	(0.511)	(0.339)	(0.183)	(2.401)	(0.514)
Referendum fixed effects	yes	yes	yes	yes	yes	yes	yes	no
Canton fixed effects	no	no	no	no	no	no	no	no
MP fixed effects	yes	yes	yes	yes	yes	yes	yes	yes
(Pseudo) R-squared	0.20	0.25	0.31	0.39	0.19	0.273	0.74	0.27
Observations	18202	3280	11968	6669	18669	6552	1628	1881

Note: Dependent var. is the matching between MP's vote and ref. outcome. Standard errors (in parenthesis) are clustered by residence-neighbor-pair.
****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

Table A7. Resident MP Bias and Size Difference of Resident and Neighbor Communities- Further Subsamples (continued)

B. Residence Comm. vs. a Single "Synthetic" Neighbor Comm.	*Residence is Birthplace* *(1)*	*Other Language* *(2)*	*Rural Community* *(3)*	*MP is Farmer* *(4)*	*Residence = Correspond.* *(5)*	*Tight Referendum* *(6)*	*Resident MPs w/same View and NO MP in Neighbor* *(7)*	*Resident MPs w/same View and MPs in Neighbor w/opposite View* *(8)*
Residence community	0.202***	0.335**	0.220	0.118	0.240**	0.554**	2.471**	1.714***
	(0.072)	(0.156)	(0.145)	(0.551)	(0.104)	(0.237)	(0.995)	(0.475)
Size Difference (Residence-Neighbor)	-1.901	1.877	1.356	-4.822***	1.423	-1.297	0.707	-10.932
	(1.305)	(1.711)	(1.888)	(1.296)	(1.613)	(2.506)	(1.586)	(22.071)
*Residence*SizeDifference*	0.051	0.502*	0.116	-0.073	0.312*	0.458	2.273	1.846**
	(0.118)	(0.297)	(0.246)	(0.828)	(0.168)	(0.406)	(2.229)	(0.865)
Intercept	1.476***	1.629	2.588	-2.071*	0.846	12.782***	1.385	-1.822***
	(0.437)	(1.618)	(1.810)	(1.104)	(0.738)	(1.044)	(2.084)	(0.529)
Referendum fixed effects	yes	yes	yes	yes	yes	yes	yes	no
Canton fixed effects	no	no	no	no	no	no	yes	no
MP fixed effects	yes	yes	yes	yes	yes	yes	no	yes
(Pseudo) R-squared	0.23	0.31	0.36	0.42	0.24	0.468	0.78	0.43
Observations	3528	780	3174	1666	4138	1466	316	362

*Note: Dependent var. is the matching between MP's vote and ref. outcome. Standard errors (in parenthesis) are clustered by residence-neighbor-pair. ***, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*

A.4 MPs' Characteristics and the *Resident MP Bias*

For each MP we create a measure of *agreement with her residence community* (ARC): for each referendum outcome, the MP is assigned 1, if her vote matches with the outcome in her residence community but not that in neighbor communities; assigned -1 if her vote matches neighbor community but not her residence community; and assigned 0 if her vote matches both or neither communities. Taking the average of these values for each MP over all referenda, we create an individual measure for an MP's ARC. Results of regressing ARC on individual characterisctics of MPs are shown in Table A8. Time spent in council has a non-linear effect: up to about eight years it is positively correlated with an MP's ARC. After eight years, more time in the council is correlated with decreasing ARC. MPs associated with left wing parties have higher ARC on average. MPs who did not change their residence since birth have higher ARC on average. This is not a surprising result, since such an individual would be expected to have better established long-term relationships within that community, and this is what we also exploit in our robustness analysis in the paper.

Table A8. Resident MP Bias and the Individual Characteristics of MPs

dependent variable: average yearly ARC of MPs

	(1)	*(2)*	*(3)*
Intercept	-0.008 (0.011)	0.018* (0.01)	-0.04 (0.025)
Gender		0.051*** (0.02)	0.024 (0.02)
Residence is same as BirthPlace			0.028* (0.015)
Graduate Education			-0.021 (0.016)
Children			0.008 (0.02)
Married			-0.003 (0.02)
Army			0.01 (0.02)
Interest Groups			0.003** (0.001)
PartyLeft	0.119*** (0.02)		0.129*** (0.02)
PartyRight	0.007 (0.016)		0.017 (0.017)
R-squared	0,076	0,014	0,091
Observations	1060	1060	1055

Note: Standard errors (in parenthesis) are clustered by residence-neighbor-pair.
****, **, and * denote significance levels p<0.01, p<0.05, and p<0.1, respectively.*