DISS. ETH NO. 28939

POLITICAL OPINIONS AND OFFLINE SOCIAL NETWORKS IN A SWISS STUDENT COMMUNITY

A dissertation submitted to attain the degree of

DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

KIERAN DOUGLAS MEPHAM
Master of Science, Utrecht University

born 22.09.1991
Swindon, United Kingdom

accepted on the recommendation of

Prof. Dr. Christoph Stadtfeld
Prof. Dr. Delia Baldassarri
Prof. Dr. Károly Takács

2022

ABSTRACT

Political attitudes and social networks are inherently linked, and societal interest in their connections has grown. Particularly, growing concerns about the fragmentation of society into groups opposed by ideology and affect towards one another have been grouped under the term of polarization. This macro-level outcome is often studied independently from the micro-level processes which may cause it. I study these processes in an offline context, examining primarily similarity-biased selection and similarity-inducing social influence on multiple political attitudes and behaviours in longitudinally-observed cohorts of Swiss students. I aim to link these micro-level processes of selection and influence to their macro-level outcomes, and to understand complicated sets of multiple opinions in parsimonious ways. In doing so, I bridge simulation and empirical models, and additionally examine behavioural outcomes. In the Introduction (Chapter 1) I introduce aims, concepts, and the current state of research.

In Chapter 2, an empirical, dynamic network model is developed for students' friendships and political attitudes over time and across a variety of topics, disentangling selection from influence effects in the stochastic actor-oriented model framework. Conceptualizing attitudes as a valenced, two-mode network between individuals and issues, endogenous structural tendencies in the friendship network and attitude network are accounted for, as are latent effects driving convergence and divergence of similar and dissimilar individuals occurring beyond the confines of their friendships. A metric is developed to capture ideological and relational aspects of polarization, and a test to capture whether it differs from an expectation in the observed network. Selection effects are found, while influence finds more modest support. On the other hand, convergent and divergent forces on attitudes are clearly found. The co-

horts examined are found to be slightly polarized, but this does not clearly change over time.

In Chapter 3, the model produced in Chapter 2 is re-applied as a generative model to examine how processes of selection and influence, and latent divergence and convergence-inducing forces affect the network and its level of polarization. Manipulating the strength of effects and model estimation methods and simulating forward from observed data, this chapter demonstrates a method that may be useful to bridge gaps in the link between agent-based models and empirical calibration and validation. In particular, it incorporates explicit notions of time and tie type, models tie change, and builds in an explicit empirical calibration. Effects of selection and influence on polarization are found to be generally weak, while latent forces are somewhat stronger. Overall there is a tendency towards a drop in polarization over time.

In Chapter 4, political discussion networks, friendships, and voting behaviour are examined, focusing on the choice to vote and the choice of vote. Effects of one's discussants and friends are examined in the understudied context of Swiss direct democratic referenda. Assimilative, normative effects of choosing to vote are expected, while a hypothesis of being isolated in one's opinion could not be tested due to its infrequency. Regarding choice of vote, assimilative influence is hypothesized, with a positive moderating effect of intending to vote but being undecided, as well as a reduced influence amongst the more knowledgeable, and increase amongst the especially interested. Apparent similarities between discussants disappear when controlling for an individual's prior vote intentions, though in friendship an association remains.

Contributions of this thesis are bridging between micro process and macro outcome, while applying a multi-issue frame to understanding selection and influence effects. In terms of results, the effects of selection and influence are not strong, but particularly selection is regularly present. These forces appear to result in little polarization – nonetheless having a potential impact on societal outcomes via voting. Further research should be conducted on identity under complex political systems, as well as integration of other cultural preferences and distinctions between types of influence. To bridge and understand the consequences of offline and online processes, more work is needed explicitly incorporating both modes. More can and should be done to understand how and when socio-political fractures may occur and how they impact us.

ZUSAMMENFASSUNG

Politische Einstellungen und soziale Netzwerke sind von Natur aus miteinander verknüpft, und das gesellschaftliche Interesse an ihrem Zusammenwirken ist gewachsen. Die hierbei wachsende Besorgnis über die Zersplitterung der Gesellschaft in Gruppen, die sich in ihrer Ideologie und in ihren Affekten gegenüberstehen, wurde unter dem Begriff der Polarisierung zusammengefasst. Die Polarisierung auf Makroebene wird oft unabhängig von den Prozessen auf der Mikroebene untersucht, die es möglicherweise verursachen. Ich untersuche diese Prozesse in einem Offline-Kontext, indem ich in erster Linie die auf Ähnlichkeit beruhende Selektion und den sozialen Einfluss, der Ähnlichkeit erzeugt, im Kontext von verschiedenen politischen Einstellungen und Verhaltensweisen durch eine längsschnittlich Beobachtung von Kohorten von Schweizer Studierenden untersuche. Mein Ziel ist es, diese Selektions- und Beeinflussungsprozesse auf der Mikroebene mit den Ergebnissen auf der Makroebene zu verknüpfen und die komplizierten Ansammlungen von Meinungen in ihrer Interaktion zu verstehen. Dabei verbinde ich simulationsbasierte sowie empirische Modelle und untersuche zusätzlich gezeigtes Verhalten. In der Einleitung (Kapitel 1) stelle ich die Ziele, Konzepte und den aktuellen Stand der Forschung vor.

In Kapitel 2 wird ein empirisches und dynamisches Netzwerkmodell entwickelt, welches die Entwicklung von Freundschaften und politischen Einstellungen von Studierenden im Zeitverlauf und über eine Vielzahl von Themen hinweg untersucht. Hierbei werden im Rahmen eines stochastischen akteursorientierten Modells Selektionseffekte von sozialen Einflusseffekten entkoppelt. Durch die Konzeptualisierung von Einstellungen als ein valenzbasiertes und bipartites Netzwerk zwischen Individuen und Themen, werden endogene strukturelle Tendenzen im Freundschaftsnetzwerk und im Einstellungsnetzwerk berücksichtigt, ebenso wie latente Effekte, die die Konvergenz und Divergenz in Bezug auf ihre Einstellungen ähnlicher und ungleicher Individuen

über die Grenzen ihrer Freundschaften hinaus beeinflussen. Es wird eine Metrik entwickelt, um ideologische und relationale Aspekte der Polarisierung zu erfassen, und ein Test, um festzustellen, ob diese im beobachteten Netzwerk von den Zufallserwartungen abweichen. Einerseits werden in der Studie Selektionseffekte festgestellt, während der Effekt von sozialem Einfluss eher gering ausfällt. Andererseits sind konvergierende und divergierende Kräfte auf die politischen Einstellungen eindeutig festzustellen. Die untersuchten Kohorten sind leicht polarisiert, was sich jedoch im Laufe der Zeit nicht signifikant verändert.

In Kapitel 3 wird das in Kapitel 2 erstellte Modell erneut angewendet, jedoch als generatives Modell, um zu untersuchen, wie sich Selektions- und Einflussprozesse sowie latente Divergenz- und Konvergenzkräfte auf das Netzwerk und seinen Polarisierungsgrad auswirken. Durch die Manipulation der Stärke von Effekten, der Modellschätzungsmethoden und der Anzahl an Vorwärtssimulationen von beobachteten Daten, zeigt dieses Kapitel eine Methode, die nützlich sein kann, um die Lücken in der Verbindung zwischen agentenbasierten Modellen und empirischer Kalibrierung und Validierung zu schliessen. Insbesondere enthält diese Methode explizite Operationalisierungen von Zeit und Beziehungsart, modelliert Beziehungsveränderungen und baut eine explizite empirische Kalibrierung ein. Die Auswirkungen von Selektion und Einfluss auf die Polarisierung sind im Allgemeinen schwach, während die latenten Kräfte etwas stärker sind. Insgesamt gibt es eine Tendenz zu einem Rückgang der Polarisierung im Laufe der Zeit.

In Kapitel 4 werden politische Diskussionsnetzwerke, Freundschaften und das Wahlverhalten untersucht, wobei der Schwerpunkt auf der Wahlentscheidung und der Entscheidung zur Abgabe der Stimme an sich liegt. Die Auswirkungen der eigenen Diskussionspartner und Freunde werden im wenig untersuchten Kontext der Schweizer direktdemokratischen Volksabstimmungen untersucht. Es werden assimilative und normative Effekte der Stimmabgabe erwartet, während die Hypothese der Isolation der eigenen Meinung aufgrund ihrer Seltenheit in den Daten nicht getestet werden konnte. Hinsichtlich der Wahlentscheidung wird ein assimilativer Einfluss angenommen, mit einem positiven moderierenden Effekt, wenn man die Absicht hat zu wählen jedoch unsicher ist wen man wählt, sowie mit einem geringeren Einfluss bei den besser Informierten und einem höheren Einfluss bei den besonders Interessierten. Die offenbarten Ähnlichkeiten zwischen den Diskutanten verschwinden, wenn man für die früheren Wahlabsichten einer Person kontrolliert, obwohl in dem Freundschaftsnetzwerk ein Zusammenhang bestehen bleibt.

Der Beitrag dieser Dissertation besteht darin, eine Brücke zwischen dem Mikroprozess und dem Makroergebnis zu schlagen und gleichzeitig einen themenübergreifenden Rahmen zum Verständnis von Selektions- und sozialen Beeinflussungseffekten anzuwenden. Was die Ergebnisse betrifft, so sind die Effekte von Selektion und sozialem Einfluss nicht stark, aber regelmässig vorhanden. Insgesamt scheinen beide Effekte zu einer geringen Polarisierung zu führen – nichtsdestotrotz haben sie durch die Wahl einen potenziellen Einfluss auf die gesellschaftlichen Ereignisse. Weitere Forschung sollte zur Identität in komplexen politischen Systemen sowie zur Integration anderer kultureller Präferenzen und zur Unterscheidung zwischen verschiedenen Arten von Einfluss durchgeführt werden. Um die Unterschiede und Folgen von Offline- und Online-Prozessen zu verstehen, ist mehr Arbeit erforderlich, die ausdrücklich beide Modi einbezieht. Es kann und sollte mehr getan werden, um zu verstehen, wie und wann gesellschaftspolitische Brüche auftreten können und wie sie sich auf uns auswirken.

ACKNOWLEDGEMENTS

Thanks to the many people who have supported me along the very long way. Special thanks to (current and former) members of the Social Networks group, especially Timon Elmer, Marion Hoffman, Isabel Raabe, and Alvaro Uzaheta for sharing difficulties and huge amounts of instrumental, informational, and emotional support throughout. To András Vörös for thoughtful feedback on every. single. chapter.; often with little notice and always with speedy turnaround. Thank you to Christoph Stadtfeld for substantial patience, re-motivation and re-framing, as well as the opportunity to take on this project. And a big thank you to the various participants of the Duisterbelt for thoughtful and detailed feedback on early versions of the ideas used in this dissertation.

Thanks to friends and family, both in the Netherlands and Switzerland, and housemates who have tolerated my nearly complete absence (especially over the last months). Thanks to my family, all of whom I can look up to for their perseverance, and on whom I can always rely.

On a more practical level, I'd like to give thanks to Zoran Kovacevic for help with the German-form abstract of this dissertation, and to the High Performance Computing team for their work on the Euler cluster which was used for model estimation and simulation in Chapters 2 and 3.

Finally, thanks to Sinéad for calming me down, lifting my mood up, for convincing me to accept that sometimes projects are done, and providing robust evidence that one day we will all evolve into crabs.

CONTENTS

1

INTRODUCTION

Political attitudes and their links to social networks have seen growth in interest in recent years. Mass political polarization in particular has been a dominant theme of the 21st century, with some debate as to its definition(s), its causes, and consequences. While some argue a level of disagreement is beneficial to democracy, others are concerned about the fragmentation of societies it may have as a result.

Mass political polarization is typically defined by reference to ideological clefts which outline opposed groups (such as conservatives versus liberals or adherents of political parties against one another) in the public, in some cases focusing on negative relations between such groups (e.g. DiMaggio, Evans, & Bryson, 1996; Iyengar et al., 2019). In understanding how political polarization may come about, fingers have often been pointed at the internet and its affordances, permitting or encouraging filter bubbles of personalized information congenial to pre-existing views or access to echo chambers to spaces of individuals with little ideological diversity (Pariser, 2011; Sunstein, 2007). As a consequence of these processes, people may be influenced to take on new, potentially extreme or overly aligned views (Turner, 1991), and/or develop negative views about the other side (Bougher, 2017; Iyengar et al., 2019; Webster & Abramowitz, 2017). The extent of the online-specific affordances causing polarization appears to have been exaggerated, although not entirely misplaced (Dubois & Blank, 2018; Guess, 2021). However, while a focus has recently been placed on the internet and social media sites, these tendencies to choose interactions with similar others (homophilous selection) and influence one another to similarity (assimilative influence) correspond to common features of social behaviour (McPherson, Smith-Lovin & Cook, 2001; Turner, 1991) also occur offline. The selectively chosen personal relationships in which these interactions can be embedded convey more trust and closer connection when offline than in purely online relationships (Antheunis, Valkenburg, & Peter, 2012; Mesch &

Talmud, 2007), and so may be more consequential for our opinions, and reflect them in a different way[1]. These patterns of selection and influence can have important outcomes for democratic societies.

Studies of political opinion dynamics have typically approached polarization and related problems in one of three ways; the first two are based on analysis of empirical observation, and the third is based on simulation. A common focus amongst empirical researchers has been to examine the macro, societal political opinion landscape and understand it through the lens of macro-level determinants such as media frames or elite opinion (Hetherington, 2009). Alternatively, micro-processes are studied, i.e. individual dispositions and dyadic processes hypothesized to contribute to polarization, and changes in individuals' political opinions or intergroup perceptions within individuals are examined (e.g. Levendusky, 2009; Simas, Clifford, & Kirkland, 2020; Whitt et al., 2021). However, these are rarely able to directly link micro-mechanisms with emergent macro outcomes. With a focus on the classic sociological problem of getting at the micro-macro link (Coleman, 1990), researchers using agent-based models (ABMs) have leveraged an increase in available computing power to build intensive and diverse simulations of societies to understand the dynamics of opinions between individuals and their possible consequences. This line of work has yielded interesting insights, such that we now know stylized sufficient conditions to cause consensus, bipolarization, and fragmentation (Flache et al., 2017). ABMs are somewhat lacking in empirical integration (Baldassarri & Page, 2021; Flache et al., 2017; Mäs, 2019; Sobkowicz, 2009). Typically, omitted from these models are features such as a notion of real-world time, and critically, the dynamics of the networks underlying dyadic influence processes. Simply put, empiricists are rarely able to directly link micro-processes to macro outcomes but are able to provide information on reality, while theory-based ABMs are able to show how these micro-to-macro problems might operate in stylized scenarios, but are hindered by the infrequent provision of empirical grounding.

In this dissertation, I attempt to integrate both micro process and macro outcomes using empirical analysis and simulation of social networks – arguing

1 Though online components may also aid the spread of opinions within these in-person relations (Jost, Baldassarri, & Druckman, 2022).

that our social relations are intrinsically linked to political opinions. Aiming firstly to bridge between micro-level mechanistic tests and macro-level studies of political polarization, I focus on the effects on and of people's chosen ties in interpersonal networks, i.e. micro-processes, which result in possible polarization and democratic vote choices, i.e. macro-level outcomes. As a second aim, I try to understand processes and outcomes in contexts where ideological and partisan affiliations are unlikely to meet commonly assumed bipolar understandings of polarization; in a multiparty system. To do this, I fit the processes and outcomes into a multivariable attitude and behaviour framework.

In tackling the above, I apply and extend concepts and statistical models for social networks to understand polarization, disentangling assimilative selection from homogenizing assimilative social influence effects on political attitudes. I consider both relational and ideological aspects. I also zoom in to explicitly political relationships and direct measures of democratic choices, examining the effects of political discussants on multiple direct democratic votes.

The second section of this introductory chapter will clarify theoretical concepts relevant to this dissertation. The third section will offer an overview of current research and its gaps in more detail, while the fourth will present the empirical context on which I have focused. The fifth section will introduce the chapters, one by one.

1.1 CONCEPTS

1.1.1 Micro mechanisms

Micro mechanisms are taken to mean effects on and effects of individuals, rather than effects on or of a large-scale context. I focus on localized features from actors' social connections, which I take to be approximately micro, or potentially meso-level features. This is essentially somewhere between what Hedström and Swedberg (1998, p.23) term action-formation mechanisms (i.e. those primarily within an individual; micro-to-micro) and situational mechanisms (i.e. those from the environment to the individual; macro-to-micro). In particular, I test social effects occurring primarily between (aggregated) pairs of individuals, examining how these affect individuals. Focal mechanisms are *social influence*, the process of change of an attribute of an individual depending on other individuals around them, and *social selection*, the process of changing an individual's connections to others according to their preferences. In chapters 2 and 3, these are formulated as tendencies towards local structures occurring between pairs of individuals and objects of their issue attitudes, while in Chapter 4, I focus on influence as similar or congenial voting behaviour.

1.1.1.1 Social influence

Throughout this dissertation, the term "social influence" is used. I follow the basic definition given by Rashotte (2007): social influence is "change in an individual's thoughts, feelings, attitudes, or behaviours that results from interaction with another individual or a group". Social influence is not based on the application of power from a sender to a receiver but on acceptance of information by a receiver (Turner, 1991). This dissertation, then, primarily concerns influence on attitudes which may or may not be publicly expressed. Considering two types of social influence, informational and normative (Deutsch & Gerard, 1955), different kinds of outcomes are expected. Normative social influence tells us about what is socially appropriate, and therefore can affect observable behaviours without affecting private beliefs. On the other hand, informational influence gives us evidence about what is true. A middle ground proposes we may be influenced in our perception of 'social reality' (Festinger, 1950), which can be to a greater or lesser extent based in physical reality. These perceptions of social reality are dependent on the knowledge we gain from those around us.

An assimilative, dyad-based, and primarily informational social influence is typically assumed in the studies presented in this dissertation. It is assimilative in that it is social influence which has the effect of making individuals similar. It is dyadic in that it is based on (aggregations of) two-person differences or similarities of attitudes or behaviours. Consistent with common conceptions of social influence in political attitudes, I do not assume that this influence relates strictly to direct information about the political attitudes and behaviour at hand, but can also include heuristics that allow individuals to make 'cheaper' informational inferences about preferences that would be consistent with other beliefs under a complex set of possible considerations such as (dis)agreement between individuals on other beliefs (similar to arguments presented in e.g. Downs, 1957; Sokhey & Djupe, 2011).

1.1.1.2 *Social selection*

Social selection as a process here refers to how individuals decide to whom they connect or choose to remain connected. In analogy to the assimilative influence presented in the previous subsection, I use selection as shorthand to refer to *homophilous* social selection. This is a similarity-based social selection, which may be circumscribed by both an avoidance of particularly dissimilar others (perhaps more properly called 'heterophobic' social selection) or a preference to associate with similar others. This follows what is described by McPherson and Smith-Lovin (1987) as 'choice homophily': a tendency to affiliate with others above the frequency expected given the opportunity set. Notably, in the sociological literature, the term 'homophily' is often used to refer to a static observation of above-chance affiliation between similar individuals (McPherson, Smith-Lovin, & Cook, 2001). This is not reflective of the process by which this observation occurs, which may involve the deliberate social selection[2] to which I refer throughout this dissertation.

1.1.2 Macro outcomes: political polarization

Political polarization is a complex topic of research; definitions abound (Bramson et al., 2016; Iyengar, Sood, & Lelkes, 2012), yielding varied conclusions on extent and change over time. In this dissertation, I focus on definitions aimed at mass publics, i.e. those who are not part of the political elite such as "politi-

2 Observed static homophily may be based on selection on the attribute in question (e.g. friendship between two people due to a preference for the same hobby) or could also involve influence (inducing a friend to take up the same hobby) or some other confounding factor (e.g. both attending the same hobby-related club and therefore becoming friends) (Shalizi & Thomas, 2011).

cians, higher-level government officials, journalists, some activists, and many kinds of experts and policy analysts" (Zaller, 1992, pp.6).

Polarization can be taken to mean either a dynamic process or a state (e.g. Bramson et al., 2016; DiMaggio, Evans, & Bryson, 1996). Focal divides in definitions separate the concept of political polarization into two main streams. Firstly, an ideological perspective on polarization centred on individuals' attitudes towards political issues (e.g. Abramowitz & Saunders, 2008; DiMaggio, Evans, & Bryson, 1996), and secondly, a social view based on attitudes towards other individuals of similar and dissimilar ideological alignment (Druckman et al., 2020; Iyengar et al., 2019). When focusing on issue-centred attitudes as the object of study, I follow the definition given by Eagly and Chaiken (1998): "a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour" when considering polarization.

In *ideological* variants of polarization, operationalizations of its static forms include either a measure of the extremity of single-issue attitudes, the alignment between attitudes and political group affiliation, or alignment of multiple attitudes also known as constraint (Baldassarri & Gelman, 2008; Converse, 1964; DiMaggio, Evans, & Bryson, 1996; Kozlowski & Murphy, 2021; Lelkes, 2016). DiMaggio, Evans, and Bryson (1996) proposed four features evidencing ideological, process-based polarization: increasing dispersion of single issues attitudes, increasing bimodality of single issue attitudes, increasing constraint between attitudes on various issues, or increased association between salient group characteristics (such as partisanship) and issue attitudes. While much polarization research has focused on the two-party political system of the United States, a multiparty context is studied in this dissertation. This means that various combinations of political attitudes can be reflected by various parties which can be substantially more complex than many of the bipolar features considered above. Due to this, I focus on constraint-based definitions of polarization both as a process and outcome.

Affective polarization is defined in considerably simpler terms, with a focus on the difference between one's (dis)like of the political group to which one belongs (the ingroup), and (dis)like of the group to which one does not belong (the outgroup). Feeling more warmly, perceiving less social distance towards the ingroup than the outgroup, or having more negative stereotypes about the outgroup than the ingroup can all be signs of affective polarization (Iyengar, Sood, & Lelkes, 2012; Lelkes, 2016; Reiljan, 2020; Wagner, 2021), and constitute attitudes about both the outgroup elite as well as its adherents (Druckman & Levendusky, 2019). This may be reflected in affiliative outcomes – i.e., with whom one has a positive social tie (Iyengar et al., 2019) – an assumption made in chapters 2 and 3.

1.1.3 Network concepts

1.1.3.1 Social ties

A 'social tie' here is used to mean the relationships occurring between two individuals, represented by a connection between their representations, regarding some perception of and/or interaction with one another. I focus in chapters 2 and 3 on ties of friendship. Friendship is largely choice-based (as opposed to ties of e.g. kinship), and often conveys frequent interaction, intimacy and emotional intensity and as such is a relatively strong tie (Fischer, 1982; Granovetter, 1973; Krackhardt, Nohria, & Eccles, 2003). Friendship can also be seen as a host for different types of interactions, including those of relevance to politics (Marsden, 1987). Friendships have been shown to channel influence and be selected across a multitude of behavioural and attitudinal variables such as cultural tastes (Lewis & Kaufman, 2018; Lizardo, 2006; Lomi & Stadtfeld, 2014), mental health (Schaefer, Kornienko, & Fox, 2011), physical health (Salvy et al., 2012) and delinquent behaviours (Brechwald & Prinstein, 2011) and in the political realm may be an explanation for similarity in political attributes in college students (Lazer et al., 2010; Levitan & Visser, 2009). In the empirical samples used in this dissertation, friendships are relatively common and are formed from the very start of the sampled cohorts' formation, and therefore have the potential to show selection and influence processes well.

Friendships may well be channels for political information but are not inherently strongly political. I therefore examine political discussion ties in addition to friendship in Chapter 4, since these discussions explicitly convey political information and are highly embedded in friendships (Marsden, 1987; Minozzi et al., 2020; Song, 2015), helping to clarify the roles of friendships in political outcomes. Nonetheless, they are not the only way political information may be transmitted across friendships: this may be inferred from other attributes which signify the political positions of one's friends, and repeated observation of co-occurring attributes may help channel further influence (Goldberg & Stein, 2018).

Notably, while we may see a tie such as friendship as a persistent, perceived relationship between two individuals, a political discussion is a relational event that occurs with a more precise, shorter-term beginning and end (though what constitutes a *political* discussion can be subjective, Hopmann et al., 2015), and can occur repeatedly between the same two individuals. In this dissertation, I treat this statistically in much the same way as friendship. This can enable an understanding of the consequences of the diversity of points of view available, but comes at the cost of ignoring qualities of repeated political discussions between two individuals.

1.1.3.2 Two-mode networks

The relationship between individuals and the subjects of their political attitudes can also be treated as a type of network known as a two-mode network. Two-mode, or multilevel networks are networks consisting of two sets of nodes, representing entities such as or individuals and subjects of an attitude (Raabe, Boda, & Stadtfeld, 2019), individuals and social events which they attend (A. Davis, Gardner, & Gardner, 1941), or organizations and their competences (Hollway et al., 2017). These sets of nodes can be interconnected between the sets, but the same tie type does not occur within sets (Lazega & Snijders, 2015). This stands in contrast to a one-mode social network of ties such as friendship occurring within a set of individuals. These can be combined, similarly to cognitive balance theory (e.g. Heider, 1946) which formulated networks of pairs of individuals and their attitudes towards other objects. In the example in Figure 1.1 attitudes about policy issues are depicted as a two-mode network. Positive and negative attitudes are treated as green or red ties respectively, from individuals holding them (circles) to the policy issues at hand (squares).

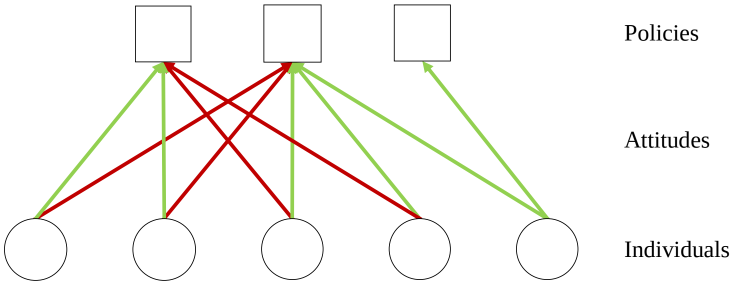


Figure 1.1: An illustrative multilevel network of individuals and attitudes. Squares represent policies, red and green ties represent positive and negative attitudes that circle individuals hold about the policies.

This two-mode representation of policies, individuals, and attitudes, especially in combination with a one-mode network between individuals enables new representations of the macro-level outcomes such as the level of polarization. Mixing these two kinds of networks aids us in treating social and ideological aspects of polarization in combination. To illustrate, the upper two structures in Figure 1.2 show two configurations of attitudes that might be overrepresented under stronger ideological polarization – individuals either holding identical or opposed attitudes on the same set of issues, representing constraint between issues. The lower two structures show two configurations likely to be overrepresented under relational polarization, a potential consequence of affective polarization – individuals are more likely to be connected to those who share an opinion, and less likely to connect to someone with an opposed opinion.

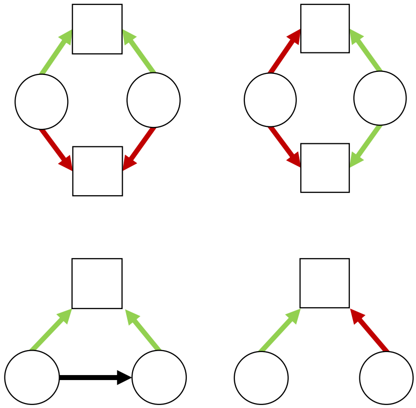


FIGURE 1.2: Four structures expected to be overrepresented in mixed two-mode and one-mode networks under ideological and affective polarization. Squares represent policies, red and green ties represent positive and negative attitudes that circle individuals hold about the policies.

The mixed two-mode and one-mode network representational approach meshes well with the strengths of the stochastic actor-oriented model, a statistical framework for analyzing longitudinal network data (Snijders, van de Bunt, & Steglich, 2010). This statistical framework, also suited for mixed one-mode and two-mode, data allows the estimation of the effects of hypothesized tendencies towards assimilative and repulsive influence, as well as homophilous or heterophobic selection, in a parsimonious way. The specific qualities of individual policy issues are ignored to make simplifying assumptions of uniformly-sized effects of e.g. influence or selection, providing the opportunity to construct a simple general model of micro-processes and link these directly to novel conceptualizations of macro outcomes.

1.2 RESEARCH OVERVIEW

In this section, I present an overview of relevant research and its current state on these topics of networked social influence and selection on political attitudes and behaviors, political polarization; primarily with a view to highlight several gaps.

1.2.1 Political polarization

Most prominent research on political polarization has been conducted in the United States, where the two-party system makes the study of the concept simpler than in many other nations. Often this results in defining polarization via reference to the two parties or two ends of an ideological spectrum: adherents of either party displaying attitudes or affection more positively to their co-party adherents, compared to the opposed party adherents, are considered evidence of polarization (Lelkes, 2016). However, in multiparty systems common in other parts of the world, definitions become more complicated (Reiljan, 2020; Wagner, 2021).

Early work demonstrated that in general the mass public is not particularly constrained by coherence in their views (Converse, 1964), suggesting a low level of multiattitudinal ideological polarization. In examining polarization of multiple issues considered separately, DiMaggio, Evans, and Bryson (1996) similarly find limited evidence of ideological polarization on single issues, with an exception of US citizens having become more polarized on the issue of abortion, and evidence of Democrats and Republicans becoming more different in their views.

Later work by Abramowitz and Saunders (2008) suggested that polarization had increased by means of increased ideological consistency between partisans when considering party adherence and ideological polarization. However, Fiorina and Abrams (2008) argued that any apparent increase in consistency was due to 'sorting'; due to change in the views espoused by party politicians to become more consistent, individuals were better able to choose one that represented their views, while overall views had not changed in distribution. Similarly, Baldassarri and Gelman (2008) showed that apparent issue alignment of individuals' views could be well explained by an individual's partisanship, a finding echoed in a report by Pew Research Center (2014). An analysis based on data that became available later showed that polarization as issue alignment had nonetheless recently increased in addition to the growth in partisanship-based polarization in the United States (Kozlowski & Murphy, 2021). This may be explained be a temporary reflection of a different speed of uptake of progressive moral values between groups (Baldassarri & Park, 2020).

In Europe, where the studies in this dissertation are situated, less systematic analysis of ideological polarization of mass publics has been published. In the United Kingdom there has been some suggestion of depolarization up until the year 2001 (Adams, Green, & Milazzo, 2012), and some mixed evidence until 2010 (Perrett, 2021). A recent preprint proposing new conceptualizations of ideological polarization while studying left-right orientations and two key political issues – immigration and European unification – shows variation in both the levels of, and change in, ideological polarization over time between countries[3] (Gestefeld et al., 2021). In Germany, polarization as constraint or alignment had been suggested to have decreased up until at least 2010 (Munzert & Bauer, 2013), while in the Netherlands, mass polarization appeared to have increased in terms of left-right self-reporting (Silva, 2018).

In the context of Switzerland, where the current studies are based, there are some suggestions of ideological polarization. At the level of the Swiss elite, there does appear to be increasing polarization amongst politicians and parties, and to some extent amongst the voting public, based on voting and media coverage (Bochsler, Gerber, & Zumbach, 2016; Bornschier, 2015). This is linked with unclear direction to Switzerland's high proportion of ideologues – those holding aligned values – in the mass public (Gonthier & Guerra, 2022).

Considering affective polarization, the tendency to have more positive feelings and attributions towards the political ingroup than the outgroup, results are somewhat clearer. In the United States, affective polarization has grown

3 Switzerland, the host of the studies in this dissertation, is actually relatively low in polarization on both issues compared to the other countries analyzed but does appear to be on the higher end of the increase of left-right polarization by a Gini coefficient. These comparisons are not statistically tested.

(Iyengar et al., 2019). Similarly, across nine of eleven European countries[4] by Reiljan (2020) there are increases in mass affective polarization between party adherents over an 8-year period, and is relatively high in an analysis of multiparty affective polarization across various nations (Wagner, 2021). In some single important issues, such as the United Kingdom's secession from the European Union (Brexit), substantial affective polarization occurred (Hobolt, Leeper, & Tilley, 2021), highlighting that the ingroups and outgroups defining affective polarization are not limited to political party affiliations, similarly to correlational evidence of issue alignment and individual-level affective polarization (Bougher, 2017). This aside, there is correlational evidence that party polarization within a country may bring about greater partisan identification (Lupu, 2015), which may widen gaps between preferences and perceptions for the ingroup and outgroup.

1.2.2 Political selection and influence

At a population level, individuals are sorted homophilously with regard to many attributes (McPherson, Smith-Lovin, & Cook, 2001), meaning that our social connections' attributes tend to correlate with our own. Political attributes are not exempt from this relationship, with political similarity correlating with romantic ties (Klofstad, McDermott, & Hatemi, 2013), online connections (Garimella & Weber, 2017) and friendships (Kandel, 1978). However, while we might observe this similarity between individuals, this does not yet explain them. These similarities can come about by homophilous selection, assimilative influence, or second-order effects of social context (Shalizi & Thomas, 2011). These processes can be hard to disentangle in observational settings[5] (Shalizi & Thomas, 2011), but longitudinal data improves our ability to do so.

Why might we select to affiliate on the basis of similarity? One core argument comes from theories of cognitive dissonance and balance (Festinger, 1957; Heider, 1946). Cognitive dissonance theory essentially argues that people strive for consistency; incongruent relations between an individual and the objects of the relations would cause stress and discomfort. To relieve these, one

4 Switzerland is only measured at one time point so change cannot be established.
5 Or even impossible outside of an experiment, depending on the required burden of evidence.

might change their attitude towards one of the objects. Extending this to a network formulation, we can imagine a pair of two individuals between whom a tie could appear, and their attitudes as a tie they hold towards some political issue. As theorized by Heider (1946) and extended by Newcomb (1961) , imagining some object *k* (such as a political issue) to which an individual *i* holds a positive attitude, and a friendship towards another individual *j*, while *j* holds a negative attitude towards *k*, might lead individual *i* to reconsider their friendship with *j*. Alternatively, when choosing between two individuals of which one holds a positive and one a negative attitude towards *k*, *i* might be more inclined to choose the individual who holds the attitude consistent with their own[6].

Furthermore, in the case of voluntary, affective relationships, other motivations may hold beyond a desire for one's own consistency. Huston and Levinger (1978, p.125) argue that amongst other factors related to behaviour and physical attractiveness, impressions of cognitive compatibility are of importance. In considering attitudinal similarity, it is summarized that similarity is 'directly reinforcing as in classical conditioning', i.e. one receives positive feedback when expressing similarity, it 'help[s] confirm one's own feeling of rightness or goodness, and that the other is also good', and 'gives promise of future liking and favorableness [*sic*] toward us', which via reciprocity might also induce one's own liking of the other.

What does the empirical evidence have to say about homophilous social selection on political attributes? I found relatively few studies attempting to directly answer this question. Signs of homophilous selection appear in e.g. dating app messaging preferences (Huber & Malhotra, 2016), but amongst students of public policy, where one might assume higher salience of political attributes, no such effect is found on socializing and co-working ties (Lazer et al., 2010), nor were selection effects found in political ideology and phone-based contact amongst two cohorts of undergraduate students (Wang, Lizardo, & Hachen, 2020). Adolescents have, however, been suggested to select their

6 Or, in the case of two issues, one might hold a negative attitude towards the first if they perceive that it is negatively related to the second, to which the individual holds a negative attitude, for instance, because another individual with whom they agree on the second issue also holds a negative attitude towards the first.

friends in part on political orientations (Kandel, 1978). Partisans select their political discussants on the basis of shared partisanship (Bello & Rolfe, 2014; Huckfeldt & Sprague, 1995). Notably, all studies except Bello and Rolfe (2014) are based in the United States.

Why does influence occur on our (political) attitudes? As mentioned in the Concepts section, on political attitudes primarily informative influence would occur in affecting private beliefs. Hearing a convincing argument on the consequences of a policy, an argument on the values to which it effectively corresponds, or simply hearing what the other's position is may all affect one's belief.

Downs' economic theory of democracy (1957) argues that gathering and considering all information on complex political topics is cognitively expensive, and the shortcut of gathering the opinions of other individuals of a similar alignment would be a heuristic shortcut. Lau and Redlawsk (2001) review the various heuristics that may be applied to voters' decision-making, highlighting the role of observation of pre-existing ideologies. Cultural and political sociologists similarly emphasize the role of such heuristics in understanding which sets of attitudes or behaviours go together (DellaPosta, Shi, & Macy, 2015; Goldberg & Stein, 2018), with DellaPosta, Shi, and Macy (2015) and Converse (1964) highlighting that these associations need not always be beliefs or behaviours internally consistent with a political ideology. How one is influenced is also likely dependent on motivational factors; tendencies towards maintaining a positive ingroup status and confirming existing beliefs may affect how one processes political information (Jost, Baldassarri, & Druckman, 2022).

Individuals frequently receive their information through intermediaries, as found by Katz and Lazarsfeld (1955) and Lazarsfeld, Berelson, and Gaudet (1944). In these early studies, it was found that many individuals constructed their attitudes in part by political information received in a two-step flow; they did not consume political media themselves but received the messaging via personal contacts in the 'primary group', such as friends who had consumed these media. One's alters may directly affect behaviour, such as voting, through their effects on attitudes towards political issues in question. Most directly, an attitude about a given policy issue will affect whether one would vote in favour of or against the policy. However, more complicated effects on an individual's behaviour may also occur. Voting, as an essential tool of democracy, may additionally be affected in other ways such as by making one uncertain about one's choice rather than directly changing one's attitude, potentially reducing the propensity to vote (Bello, 2012; Mutz, 2002; Nir, 2011; Pattie & Johnston, 2009).

In examining political influence, various studies have found links to alters' attributes, typically taking an assimilative view of influence. Amongst the simplest conceptualizations, Lazer et al. (2010) demonstrated a difference in college students' self-identification on a liberal-conservative scale which depended on one's social ties, finding no homophily on this network, nor any effect of work-based rather than social ties. This aligns with work on two US college cohorts which found that roommates had a modest effect on self-reported ideology on a scale ranging from far left to far right (Strother et al., 2020). In a French sample of incoming political science students, Algan et al. (2020) similarly finds that friendship leads to more similar left-right identification. Levitan and Visser (2009), used two studies, one based on college dormitories as a network boundary, and one based on ego networks, and found that less attitudinal congruence in one's ego network predicted a greater likelihood of persuadability and attitude change on affirmative action and George Bush's presidency. Campos, Hargreaves Heap, and Leite Lopez de Leon (2017), on the other hand, used a classroom-based (rather than tie-based) design and found no effect of Brazilian university students' peers' political identification, but found that peers' political engagement resulted in more centrist self-identification. Wang, Lizardo, and Hachen (2020) found no evidence of influence on partisan or ideological variables in college students, in two cohorts. Bello and Rolfe (2014) and Huckfeldt and Sprague (1995) found that discussants influenced one another in their vote choices. Bello and Rolfe (2014) note that this effect is apparent particularly if partisanship is weak (and note also that a deselecting effect of discussion occurs for stronger partisans with disagreeing discussants). In a study of political attitudes on multiple political issues, namely racial integration, marijuana legalization, and foreign policy towards China, Tedin (1980) conducted research amongst adolescents in the United States. He found some evidence of assimilative peer influence net of parental influence, particularly on marijuana legalization. Here, the salience and peer versus parental orientations of the individuals were found to be moderating effects.

Both for theoretical reason and empirical precedent, influence on attitudes in observational contexts is an interesting research avenue. While the handful of dependent variables used in these studies is of interest due to their *potential* impact, more concrete behaviour that has the potential to impact society more directly is of additional interest, revealing at least one consequence of changes in attitudes brought about by one's conversation partners. Primarily elections have been studied. Here, too, assimilative effects are commonly found, as comprehensively summarized in Rolfe and Chan (2017) and Santoro and Beck (2016). When examining such person-to-person effects on voting, a distinction should be made between turnout (i.e. choosing to vote) and the choice ultimately made by a voter. In terms of turnout, political discussion may have an effect via increasing information about how to vote; or increasing the salience of the vote; or it may have normative effects, increasing or decreasing turnout in line with discussants' planned behaviour (Klofstad, 2007, 2015; McClurg, 2003, 2006). Exposure to disagreement had previously been suggested to either decrease (Mutz, 2002) or increase (Huckfeldt & Mendez, 2008) turnout, but a recent meta-analysis suggested an overall null effect when examining a linear relationship between disagreement in one's network and turnout (Matthes et al., 2019). However, being fully isolated in one's opinion may yet suppress turnout (Bello, 2012; Nir, 2011).

On the other hand, studies of the choice of vote frequently (although not always) find effects of various peers and discussion partners on individuals' behaviours. For instance, in the United States, Huckfeldt and Sprague (1995, p. 140) find assimilative effects of discussant votes in the 1984 election. In Brazil, Baker, Ames, and Renno (2006) found frequent disagreeing discussion, which had a powerful influence on vote choices. Schmitt-Beck (2004) examined data from East and West Germany, Spain, the United Kingdom, and the United States at a single time point, generally finding assimilative effects of discussants' political preferences. Interestingly, in multiparty systems it has been suggested that a complete change of intended election vote is unlikely, but that there is essentially a segment of possible parties which would be acceptable to vote for (or one party for which one strictly does not want to vote), and discussants may help to sway their choice amongst them (Zuckerman & Kroh, 2006). Examining sets of issues, where subtler differences may make the difference in choosing one party over another in a more complicated system, may yield new insights into these smaller yet consequential attitude changes amongst the mass publics.

Political attitudes and behaviours are clearly embedded in social networks, and to understand them we, therefore, need methods that reflect this and appropriately account for it.

1.2.3 From selection and influence to polarization

What are the macro-level outcomes of the interpersonal processes of selection and influence? Potentially, they can lead to polarization.

Sociologists and computational social scientists examining culture have considered effects of micro-level processes of interpersonal interactions on producing macro-level outcomes (Baldassarri & Bearman, 2007; DellaPosta, Shi, & Macy, 2015; Flache et al., 2017; Goldberg & Stein, 2018)[7], frequently taking a multiattribute approach.

Users of agent-based models, a social simulation technique assuming autonomous yet interdependent agents who follow simple adaptive rules (Macy & Willer, 2002), have also approached issues relating to opinion dynamics amongst individuals. An advantage of the ABM approach is in its flexibility, allowing the user to specify various functional forms and complexities of an influence process. Nonetheless, results are relatively uncomplicated: Modelling opinion interdependence as assimilative social influence flowing through networks, consensus is guaranteed in the long run, while including a variant of repulsive influence (i.e. influence towards disagreeing with one's alters, typi-

7 Notably, some would consider the size of networks in such studies a meso-level, rather than macro-level unit of observation (Manzo, 2007) While larger scale networks such as entire countries might then be more clearly macro-level observations, the conceptual and statistical approaches taken in these studies should also be appropriate to help understand this higher level (Stadtfeld & Amati, 2021, p.448-450).

cally when they are initially too opposed) (bi)polarization is guaranteed, and under bounded-confidence models (i.e. those with limits on influenceability between connected individuals dependent on their similarity) clustering may occur (Flache et al., 2017).

The steps made in these studies help us understand how micro processes would lead to macro outcomes. However, the relative lack of empiricism in such studies limits their ability to be certain of their conclusions, and for this reason, scholars in the field have called for better anchoring of ABMs to empirical data (e.g. Flache et al., 2017; Mäs, 2019; Sobkowicz, 2009). Various studies are building towards tests of mechanisms and outcomes (see e.g. Chu et al., 2021; Takács, Flache, & Mäs, 2016, for exemplary cases of validation of an outcome and mechanistic (calibration) testing, respectively). Several key problems are common, though not omnipresent, and should be highlighted: absolute notions of time are ignored, relying instead on model time, giving us little understanding of how quickly we may reach certain outcomes. The exact ties constituting the stylized networks are often not made explicit, making tests of the hypothetical model outcomes hard to verify. Additionally, agent-based models typically do not assume dynamic networks, instead opting for highly stylized and static networks (but see Baldassarri & Bearman, 2007, for a counterexample featuring networks with a specified tie (discussion) and dynamic changes) such as grids and other regular lattices, or networks generated from simple mechanistic models such as preferential attachment (Albert & Barabási, 2002; Price, 1976) or small-world graphs (Watts & Strogatz, 1998). This firstly ignores reciprocal effects between attitudes and ties, and secondly does not consider other endogenous features that may be present in real-world networks. These could affect both the speed at which an outcome is reached, and its features.

1.2.4 Remaining gaps

While research thus far has been extremely useful in understanding the links between social ties and political attitudes, they leave gaps to be filled. These gaps cover both conceptual as well as methodological issues.

Firstly, single issues and partisan-related variables, such as preferences for affirmative action, voting, or partisan affiliations, are typically considered in isolation in studies of selection and influence (Lazer et al., 2010; Levitan & Visser, 2009). The consequences for multiple issues simultaneously examined are less well known outside of simulation models (such as DellaPosta, Shi, & Macy, 2015; Flache et al., 2017; Goldberg & Stein, 2018). These have the potential to reinforce any societal clefts where they may be forming (Bougher, 2017; DellaPosta, Shi, & Macy, 2015); the aggregated effects of small issues may be of similar effect to these more obvious issues. This focus on single-issue bipolarity is likely in part due to the focus on the United States, with its two-party system which facilitates bipolarization over other types of fragmentation (e.g. Reiljan, 2020; Wagner, 2021), and in part due to the focus of studies on specific voting behaviour. Particularly in the more complex political systems common throughout the rest of the world, examining a multitude of issues and their social dependencies simultaneously could yield further insights into how we come to shape our beliefs and networks.

Secondly, the issues considered are often selected on their extreme salience and contentiousness – topics such as same-sex marriage, abortion, and gun control in the United States (DiMaggio, Evans, & Bryson, 1996; Shi, 2016), or the United Kingdom's secession from the European Union (Hobolt, Leeper, & Tilley, 2021). Such issues and foci of behaviour are good and obvious candidates to examine, as these are perhaps the most likely grounds on which social influence, selection and polarization may take place. However, to an extent this is selecting on the dependent variable – focusing on variables known to attract much interest and be subject to more frequent discussion likely results in greater measured polarization (Baldassarri & Bearman, 2007). More research is needed in less obvious ways in which influence and selection take place on less salient issues, potentially creating subtler grounds which nonetheless result in distinct ideological profiles constituent of polarization (Baldassarri & Gelman, 2008; DellaPosta, Shi, & Macy, 2015; Goldberg & Stein, 2018; Kozlowski & Murphy, 2021; Parker, Parker, & McCann, 2008).

The third and fourth limitations in many of the empirical studies concern methodology. Thirdly, few studies contain both ego and alter reports of opinions or behaviour. For instance, a single item asking individuals about perceived disagreement may be used (Pattie & Johnston, 2009), which can reflect perceptions about the network, but is not getting at the ground truth of an individuals' attitudinal environment. Asking egos about alters' individual opinions or partisanship, as is done in a number of studies (Bello, 2012; Eveland Jr & Hively, 2009; Morey, Eveland, & Hutchens, 2012; Nir, 2011; Rawlings, 2022), may draw closer to getting at the nominal aim of social influence and selection studies. However, both methods generate results likely to be biased towards homogeneity of opinions between individuals studied due to tendencies to both misrepresent and misperceive our political attitudes (Cowan & Baldassarri, 2018; Goel, Mason, & Watts, 2010; Huckfeldt & Sprague, 1987, 1995; Kitts, 2003; Laumann, 1969). In this way, it is hard to estimate network influence effects in the long term. Repeated measurement of both ego's and alter's opinions, as well as their local networks, permits better inference about changes in both opinions and ties, and the causes of each. Overall, however, relatively few datasets exist appropriate for these purposes (Rawlings, 2020).

Fourth, few studies are able to convincingly disentangle selection from influence due to the fact that these processes can produce outcomes which are similar, if not identical (Shalizi & Thomas, 2011). To disentangle them in observational settings requires longitudinal data on networks and attitudes which are unsurprisingly often not available, given the difficulty with which they are collected (Jussim & Osgood, 1989; Robins, 2015).

A fifth and final limitation of connecting studies looking at micro-level influence and selection, and macro-level outcomes via agent-based models, is that while they note their own relevance to the other topic, they typically do not bridge to one another directly via demonstration of either the plausibility of their processes in the case of simulation models, or the outcomes in the case of studies of selection and influence. Extending extant empirically-grounded models can help us to understand the macro-level consequences of the assumed micro-level processes in a more rigorous way.

1.2.5 Network endogeneity

Many factors may affect the emergence and change of one-mode and two-mode networks. Subject to various confounding variables as in any observational data analysis, social ties in particular also have the interesting property of being inherently self-dependent: ties between individuals depend on other ties occurring at the same time. Methods applied to understand the link between social network ties and political attributes must reflect this.

In social networks, there are at least three important endogenous structural features that consistently appear (e.g. Snijders & Steglich, 2015, p. 228; Stadtfeld & Amati, 2021, p.434-435; Stadtfeld, Vörös & Takács, 2020, p. 130). Firstly, a tendency towards *reciprocity*: if individual *a* nominates individual *b* as a friend or political discussant, then it may be expected that *b* also nominates *a* with a higher probability. , while for a political discussant an occurring event of political discussion may be remembered by both *a* and *b*. Secondly, a tendency towards *transitivity*: if *a* is friends with both *b* and *c*, then *b* and *c* are also more likely to be friends. This could be for various reasons; either due to unobserved variables causing ties simultaneously such as exposure to a shared environment such as cohabitation or homophilous but unconsidered features working in the network, or the effect could flow more directly between ties, for instance, that *a* introduces *b* to *c*. Thirdly, degree-related effects may occur: for instance, if *a* is nominated as a friend by both *b* and *c*, individual *d* is also more likely to nominate *a*. *a* might be a particularly likable person, or the fact that many others act friendly towards them might provide some information on social status (but see Vörös, Block, & Boda, 2019, for the limits of interpreting degree-related effects). As illustrated in the examples, some of these are mechanisms that more or less directly imply endogenous causes of ties in a network (structures causing ties), whereas some are structures that are more likely to occur due to variables outside of the dyadic interrelations (non-endogenous variables causing ties and thereby structure).

Where political attributes reciprocally affect ties, then, we should model the ties' interdependence. Why? Firstly, to improve our inferences about effects of particular interest: assuming dependence may affect the estimates of processes such as selection and influence (e.g. Krackhardt, 1988; Steglich, Snijders, &

Pearson, 2010). Secondly, these network-endogenous effects can moderate the consequences of selection and influence, potentially changing the inferences for a given model in the long run. Making these inferences more realistic over longer simulation periods is key in making the micro-macro linkage, extrapolating from a micro model.

Consider the example in Figure 1.3. Three individuals are connected in a two-path, i.e. at time 1 *a* and *b* are connected, *b* and *c* are connected, but *a* and *c* are not. *a* and *b* have a shared positive attitude about policy item *I*. Assume that individuals adopt the attitudes of their alters, and that transitivity operates in the process depicted in the lower half of the figure, but not in the upper half. At time 2, *c* has adopted *b*'s attitude towards *I*, but also adjusted their own attitude about *II*. If the endogenous mechanism of transitivity is assumed, *a* and *c* will also be directly interpersonally connected by time 2. By time 3, in the network model with transitivity both *b* and *a* will have directly adopted *c*'s attitude, while without transitivity, first *b* must have adopted *c*'s new attitude about *II* before *a* can adopt it too, slowing down the process of homogenizing in this cluster.

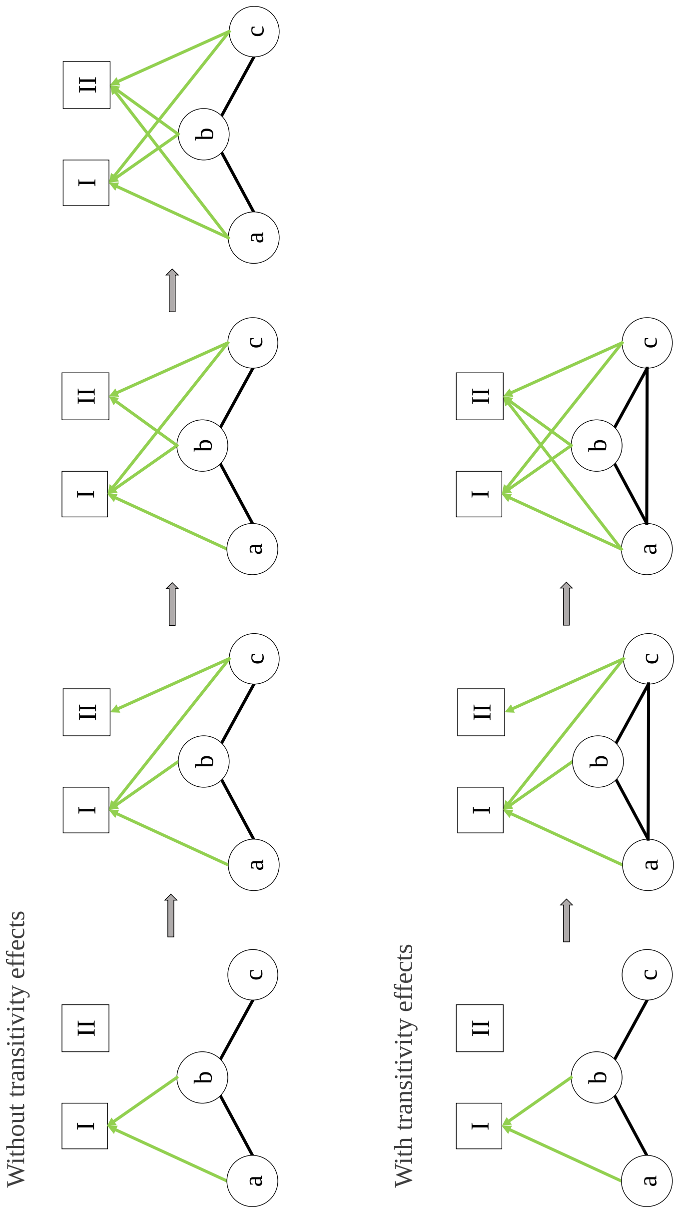


Figure 1.3: An illustrative example of a network model with influence, with and without the endogenous mechanism of transitivity in the interpersonal network, and the consequences for how quickly consensus is reached over time.

Where we can assume that the endogeneity of ties is irrelevant to the direct outcome of interest, for instance due to their stability over the period of interest, we should apply methods that control for them including permutation-based tests such as the Quadratic Assignment Procedure (Dekker, Krackhardt, & Snijders, 2007; Krackhardt, 1988).

1.3 EMPIRICAL CONTEXT

All chapters of this dissertation incorporate data from the Swiss StudentLife Study (Vörös et al., 2021). This panel-based, whole network study followed three cohorts of students at the same technical higher education institution over the course of their entire three-year bachelor's degree, starting in the first week of studying. Multimodal data collection techniques were applied, including the use of surveys, to understand both how the students' multiplex social networks developed within the bounds of their cohort, as well as the consequences of these networks for their well-being and academic outcomes. Data on political attitudes and voting behaviours were repeatedly collected for each of the three cohorts, though not all items were asked at all waves. Chapters in this dissertation use data from long-format surveys administered approximately every three months. Chapter 2 uses data from cohorts 2 and 3 during their first year. Chapter 3 uses data from cohort 2 from this same time period. Chapter 4 uses data from all three cohorts, at varying time periods where direct democratic referendums occurred in Switzerland.

1.4 CHAPTER OVERVIEW

In this dissertation, a localized community is examined to see how these offline social networks influence political attitudes and vice versa, exploring the plausibility and extent of simple mechanisms at this scale in effecting societal outcomes. In doing so a view is taken of social networks as dynamic systems featuring endogeneity: ties between individuals are nonindependent.

The methods applied reflect this point of view and leverage and/or them in judging their outcomes. Sets of political issues are examined and generalized mechanisms are explored that may affect their relation to one another, and to the social networks in which they occur in chapters 2 and 3. Short-term effects that may have a broader societal impact via voting are considered in Chapter 4. Existing methodologies for examining dynamic networks are extended and applied to purpose-collected data.

1.4.1 Chapter 2

In Chapter 2, polarizing processes and outcomes are examined using the Swiss StudentLife Study data. In these panel-based data, students' friendships are tested for reciprocal influence with their attitudes towards various political topics, in processes hypothesized to contribute to polarization. A dynamic two-mode network view is taken of individuals and political issues as nodes, with the positive and negative attitudes of individuals as ties. This leads to a formulation of network polarization as a combined measure of multidimensional ideological polarization and relational polarization, the latter of which is theorized to be a downstream consequence of affective polarization. The stochastic actor-oriented model is applied here. New effects and goodness-of-fit statistics are added to the RSiena framework, and a new statistic for network polarization and its subcomponents is presented, alongside a test of their significance given the distribution of attitudes in the network. In this paper, the construction of new macro-level metrics and the estimation of linked micro-level processes provide new contributions in bridging micro-macro gaps of selection and influence, while acknowledging a multidimensional, constraint-based view of ideological polarization.

1.4.2 Chapter 3

In Chapter 3, the consequences of the findings of Chapter 2 are examined over a hypothetical two-year period. The stochastic actor-oriented model is applied as an empirically-calibrated agent-based model for opinion dynamics, and several approaches used in previous work are tested and discussed. Experiments are conducted with altered parameter sizes to understand the sensitivity of the network to different strengths of processes implicated in polarization, while accounting for endogenous friendship network effects that potentially alter the level of polarization in the outcome. In this way, a contribution is made to the development of agent-based models via the inclusion of explicit time and ties, dynamic networks of influence, and empirical calibration of micro-level processes.

1.4.3 Chapter 4

In Chapter 4, a shift is made from attitudes and friendship as the objects of study towards the relationship between discussion partners, friends and voting behaviour in direct democratic referenda commonly occurring in Switzerland. In doing so, a dyadic behaviour which tends to occur within friendships is zoomed in on, and a behaviour that has the potential to impact society outside of the community's boundaries via formal political institutions is examined. In this chapter, assimilative influence of choosing to vote is tested, and amongst voters whether the attitudes of one's alters affect one's choice of vote. In the context of Switzerland's unique direct democratic system, I capture potential behavioural impacts of social influence across various political topics, using ego and alter reports both pre- and post-referendum.

2

CHAPTER 2 – NETWORK POLARIZATION: THE STUDY OF POLITICAL ATTITUDES AND SOCIAL TIES AS DYNAMIC MULTILEVEL NETWORKS

Ideological and relational polarization are two increasingly salient political divisions in Western societies. We integrate the study of these phenomena by describing society as a multilevel network of social ties between people and their attitudinal ties to political topics. We then define 'network polarization': the extent to which a community is ideologically and socially divided. Using longitudinal network modelling, we examine whether network polarization can be explained by three processes: social selection, social influence, and latent-cause reinforcement. Applied to new longitudinal friendship and political attitude network data from two Swiss university cohorts, we find mild polarization. The models explain this outcome and suggest friendships and political attitudes are reciprocally formed and sustained. (Dis)similar attitudes are more likely to be formed or maintained between like-minded (oppositionally-minded) individuals. Applied across other cultural contexts, our approach may help to understand the degree and mechanisms of divisions in society.

This article is currently in preparation as:

Mepham, K., Vörös, A., & Stadtfeld, C. (TBD). *Network polarization: The study of political attitudes and social ties as dynamic multilevel networks*

Recent research suggests that political polarization is a present and increasing problem globally, fragmenting societies into partisan groups (Bochsler, Gerber, & Zumbach, 2016; Hetherington, 2009; Iyengar, Sood, & Lelkes, 2012; Reiljan, 2020). In the US, some authors conclude that this has come to the point of a 'culture war' (as popularized by e.g. Bishop & Cushing, 2008; Hunter, 1991). Though the terminology is probably hyperbolic (Praet et al., 2021), partisan groups do appear somewhat incongruous not only in their political views but also in cultural consumption and general lifestyle (DellaPosta, Shi, & Macy, 2015). This is accompanied by growing antagonism and the social rejection of political identity-based outgroups in favour of the ingroup (i.e. ingroup bias, Hewstone, Rubin, & Willis, 2002; Hobolt, Leeper, & Tilley, 2021; Iyengar, Sood, & Lelkes, 2012; Reiljan, 2020). Combined with increasing inequalities, polarization threatens to amplify social divisions and lead to large-scale conflicts and disarray in society.

Currently, there is a lack of consensus in the scientific community about the extents, causes, and conceptualizations of divisions in different societies. Dominant perspectives in political science, social psychology, and sociology emphasize ideological polarization (e.g., Abramowitz & Saunders, 2008; Fiorina & Abrams, 2008; Hetherington, 2009) and various facets of (intergroup) relational preferences (e.g. Conover et al., 2011; Iyengar et al., 2019; Iyengar, Sood, & Lelkes, 2012; Reiljan, 2020) as key aspects of political division. Recent studies point out that ideological and social processes may work in conjunction to amplify initial differences between people, extending political divides to the realms of cultural and social life (DellaPosta, Shi, & Macy, 2015). However, an integrative approach to the study of these two processes is missing to date.

We propose a joint examination of ideological polarization and relational segregation, by conceptualizing societies, and smaller communities within them, as multilevel networks[1]. In these networks, people may be connected to each other by social ties, such as friendships. At the same time, people are also linked to various political topics by attitude ties, such as supporting or oppos-

1 Other notable network-based approaches for understanding attitudes (e.g. Boutyline & Vaisey, 2017; Dalege et al., 2016) focus primarily on the within-person interrelations of attitudes, where we focus on the social structuring of these connections.

ing certain issues. While we focus on political topics, the concept of attitudinal ties can be naturally extended to other social objects to represent cultural consumption and lifestyle choices.

The multilevel network approach enables us to understand the extent and causes of social divisions. First, we define and quantify 'network polarization', which captures the extent to which the network of a community is characterized by ideological and relational polarization. A metric is proposed that is based on weighted counts of triads and four-cycle in the multilevel network. Second, by further developing existing statistical models for dynamic multilevel networks, we explore the role of three network processes, social selection, social influence, and latent-cause reinforcement, in creating and maintaining ideological and social divides.

We apply these techniques to a unique longitudinal network dataset of two Swiss undergraduate student cohorts. We measured friendship and positive and negative political attitude networks of students in these cohorts over three years. We calculate the network polarization metric and assess how well it is reproduced by an empirically-calibrated stochastic actor-oriented model that considers the co-evolution of friendships and attitudes.

THEORY AND RESEARCH QUESTIONS

Two perspectives on polarization

Polarization is a frequently discussed topic in the popular debate about growing divisions in Western societies. What polarization means is often left unclear in the public realm. Academic works have offered more clarity, but they use a variety of concepts and measures. Dominant perspectives focus on either the ideological or the social facets of polarization. Below, we summarize the main concepts and empirical findings linked to these two key approaches. We argue that ideological and social aspects should both be part of a generalized definition of polarization.

The first perspective on polarization we describe here focuses on political ideology. In this approach, a community is considered polarized if its members have divergent political attitudes and opinions. Studies differ in how they conceptualize and measure 'attitudes' and 'divergence'. In examining single political issues, the variance and kurtosis of the distribution of people's attitudes have been used as key indicators of polarization. By such metrics, Western societies appear polarized by only a few key issues, such as the US public on the issue of legalizing abortion (DiMaggio, Evans, & Bryson, 1996).

Other researchers have considered multiple political issues simultaneously. Here, an increase in correlations between people's attitudes towards different issues has been considered as the indicator of polarization (Abramowitz & Saunders, 2008; Baldassarri & Gelman, 2008). This is sometimes described as the level of 'alignment' or 'constraint' (Converse, 1964; DiMaggio, Evans, & Bryson, 1996). Empirical evidence for polarization as political attitude alignment is controversial. To our knowledge, it has only been examined in recent years in the United States using General Social Survey and American National Election Survey (ANES) data. Although earlier results suggested that the alignment between political attitudes had not increased in the previous decades (Baldassarri & Gelman, 2008; Fiorina & Abrams, 2008), research extending these analyses to include more recent data from the ANES find a rapid rise in correlations between various political attitudes within multiple domains (Kozlowski & Murphy, 2021).

The second main perspective on polarization focuses on interpersonal relations. Relational polarization is defined here as the clustering of social ties so that people preferentially associate with their political ingroups. A measure of this form of polarization has been proposed in the area of computational social science: the extent to which individuals are socially tied within and between their respective political groups (Conover et al., 2011), or within and between sets of individuals sharing an attitude towards an issue (Guerra et al., 2013). The central assumption of such operationalizations of polarization is that the pattern of social ties should reflect underlying opinion similarity.

Iyengar et al. (2019) and Iyengar, Sood, and Lelkes (2012) explore the psychological roots of affective polarization through the study of changes in affect towards the perceived political in- and outgroup. They find increasing differences in warmth of feelings, willingness to attribute positive and negative traits, and social distance between political in- and outgroups in the US. Similar patterns have been found between partisans in Europe (Gidron, Adams, & Horne, 2020; Reiljan, 2020), and groups opposed on single issues such as Brexit (Hobolt, Leeper, & Tilley, 2021). Studying affect offers insight into how much people from different political factions may be willing to constructively interact.

For this reason, we consider relational polarization as a behavioural outcome of affective processes. We also argue that it is closely linked to political attitudes and self-identification. In our view, social ties often capture the affective dimension of polarization. Individuals' opinions of one another are impacted by (perceived) similarities in political attitudes and identity. This, in turn, has consequences for their affect towards one another, partially determining their social ties.

We argue that both the ideological and relational perspectives presented capture important aspects of the polarization of society. Ideological polarization – the alignment, or constraint, of political attitudes – indicates the absence of a common ground for consensus between individuals. Relational polarization – the occurrence of social ties between individuals with aligned political attitudes and the absence of ties between those with opposed attitudes – reflects the absence of opportunities for finding or creating such common ground between people of different political views. We believe that the extent and causes of social divisions can only be fully understood by taking into account both aspects. We propose an approach rooted in social network analysis to integrate the two strands of research on polarization.

Political attitudes and social ties as a multilevel network

We propose an integrative framework for the joint study of ideological and relational polarization as a multilevel social network. In the social network literature, multilevel networks are graphs in which multiple types of nodes (e.g. people, topics) are connected by multiple types of edges (e.g. social ties, attitudes). The study of multilevel networks has been receiving increasing attention within the area of social networks. They can now be analyzed by a large number of theoretical and statistical tools (Lazega & Snijders, 2015).

In the present context, political attitudes and social ties are represented as a multilevel network. The nodes of the network are people, on the one hand, who are connected to each other by social ties, specifically friendships. On the other hand, political issues form the second set of nodes. People are linked to issues by positive or negative attitudes towards the issue at hand, or have no tie when their opinion is in the neutral range. Representing actor-issue attributes as a two-mode network allows for the modelling and representation of sets of opinions and their changes in one coherent network model, rather than treating single issues separately. Figure 2.1 presents a hypothetical ideologically and relationally polarized network of four individuals (circles), their friendships (black arrows), and attitudes (darker red and light green arrows) to two issues (squares). Polarization is evident from the fact that socially connected individuals consistently agree while unconnected individuals consistently disagree in their attitudes.

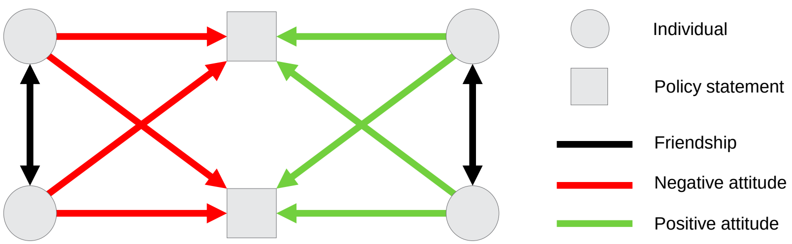


FIGURE 2.1: An illustrative multilevel network of individuals and attitudes. This stylized network is maximally polarized relationally and ideologically: people are only friends with others with whom they agree, and if they agree on a randomly selected issue, then it is guaranteed that they agree on others.

Our choice to focus on friendships as social ties is motivated by an extensive literature demonstrating their often reciprocal relationship with a variety of individual outcomes, such as cultural tastes (Lewis & Kaufman, 2018; Lizardo, 2006; Lomi & Stadtfeld, 2014), mental health (Schaefer, Kornienko, & Fox, 2011), physical health (Salvy et al., 2012) and delinquent behaviours (Brechwald & Prinstein, 2011). We note that the empirical content of our approach using social and attitude ties can be flexibly chosen, and thus can be generalized beyond friendships and political attitudes.

Our novel conceptualization and network approach allow us to contribute to the literature on political polarization in two ways. First, we can define and measure the extent of both the ideological and social aspects of the phenomenon in a given community. Second, we can test hypotheses about generative network processes that explain the observed levels of polarization – that is, we can study how communities become polarized.

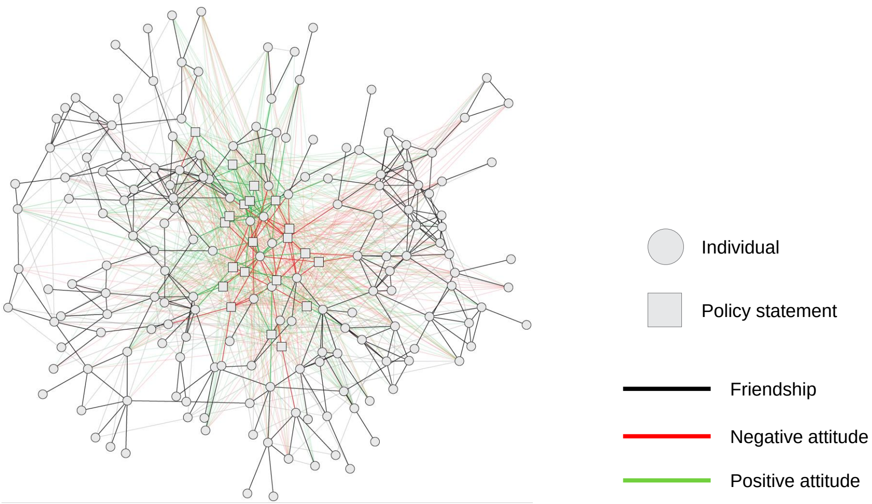


FIGURE 2.2: Multilevel network of individuals and attitudes at Wave 5, in Cohort 1. Layout by backbone algorithm (Nocaj, Ortmann, & Brandes, 2015). Arrowheads representing tie direction omitted for clarity.

Network polarization

In this paper, we introduce the term *network polarization*, which expresses how much the multilevel network of political attitudes and friendships described above is polarized. Network polarization has two dimensions. Firstly, the extent to which pairs of individuals consistently agree or disagree in their attitudes, in the sense of 'constraint' or 'alignment' (Baldassarri & Gelman, 2008; Converse, 1964). Secondly, the extent to which friendship relations are present between agreeing, but not disagreeing, individuals, i.e. a clustering of balanced triads of people and attitudes (Cartwright & Harary, 1956; Heider, 1946). The former represents the extent of *ideological polarization*, the latter the extent of *relational polarization* in the network.

We define two metrics to measure the dimensions of network polarization in empirical data. These metrics are based on assessing the frequencies of specific structural configurations in the network: certain four-cycles (two people and two issues connected by specific patterns of attitudes) and certain triads (two people and an issue connected by specific patterns of friendship and attitudes). See Appendix A.1 for graphical depictions of the considered configurations. The four-cycle structures in the network relate to the degree to which the network is ideologically polarized, and the triad structures to the degree to which it is relationally polarized. The exact mathematical definitions of the statistics are presented in Appendix A.2 and relate to what is known as a subgraph census in the social networks literature (Wasserman & Faust, 1994). The symbolic representations of the mathematical definitions are given in Table 2.1.

TABLE 2.1: Icons used in formalizing network polarization

Icon	Description
◇A	Agreeing four-cycles: two individuals share an opinion on a pair of political issues, either in favor of or in opposition to a given statement.
◇D	Disagreeing four-cycles: two individuals have opposed opinions on a pair of political issues, with one individual having a positive and one individual a negative opinion on each issue.
◇I	Inconsistent four-cycles: two individuals share an opinion on one issue, and have opposed opinions on another.
△A	Agreeing opinion closed triads: Two individuals both oppose or both support a given issue and are friends.
/A\	Agreeing opinion open triads: Two individuals both oppose or both support a given issue, but are not friends.
△D	Disagreeing opinion closed triads: Two individuals have opposed opinions on an issue, but are friends.
/D\	Disagreeing opinion closed triads: Two individuals have opposed opinions on an issue, and are not friends.

Ideological polarization. The metric of ideological polarization has two components, each based on unordered sets of complete four-cycles (see Table 2.1). The first two structures, ◇A and ◇D may be seen as indicators for consistency of agreement, or ideological attraction, and consistency of disagreement, or ideological repulsion. ◇A represents the case where two people are in agreement across two issues: both of them support or oppose both issues or they both support one issue and oppose the other. In ◇D, the two individuals disagree on both issues: if one of them supports either, the other opposes it. The third structure, ◇I refers to inconsistency in agreement: the two individuals have the same attitude regarding one issue but not on the other.

The first component of ideological polarization is thus *ideological attraction*. This is defined as the proportion of cases where two people consistently agree on two issues rather than having views that are inconsistent with ideological polarization (measured by inconsistent four-cycles $\Diamond_{\mathrm{I}}$):

$$\text{Ideological attraction} = \frac{\#\Diamond_{\mathrm{A}}}{\#\Diamond_{\mathrm{A}} + \#\Diamond_{\mathrm{I}}}. \tag{2.1}$$

The second component of ideological polarization is *ideological repulsion*. This is similarly defined as the proportion of cases where two people consistently disagree on two issues rather than having inconsistent views on those issues:

$$\text{Ideological repulsion} = \frac{\#\Diamond_{\mathrm{D}}}{\#\Diamond_{\mathrm{D}} + \#\Diamond_{\mathrm{I}}}. \tag{2.2}$$

The two components can be interpreted as analogous to the probability of two individuals to agree (disagree) on one issue if they agree (disagree) on another. We consider the mean of the two component metrics to measure the level of ideological polarization as the probability of observing consistent attitudes (either disagreeing or agreeing) in pairs of individuals. It is noteworthy at this point that the baseline probability of two individuals to agree is affected by various factors, such as the general prevalence of positive and negative attitudes (the attitude network densities) and the different tendencies of items to attract positive or negative attitudes (the item degrees). The final measure proposed corrects for these baseline probabilities.

Relational polarization. The metric of relational polarization also has two components that relate to within-group attraction and between-group repulsion. Both are based on ordered triad counts: each triad is counted twice, once considering the potential friendship from individual *i* to individual *j* and once the other way round. In the following, we refer to these structures by the remaining symbols given in Table 2.1. The structure $\triangle_{\mathrm{A}}$ represents agreement among friends: they both support or oppose an issue. $\wedge_{\mathrm{A}}$ represents agreement in the absence of friendship. Together, these can be used to represent the extent to which people tend to be friends depending on their political agreement. Similarly, $\triangle_{\mathrm{D}}$ stands for disagreement between friends, and $\wedge_{\mathrm{D}}$ for disagreement in the absence of friendship. Counting these two structures in the network can help to understand the extent to which people's friendships are structured in a way that avoids encountering disagreement.

The first component of relational polarization is *relational attraction*. We define the corresponding statistic as the ratio of ordered triads in which two individuals who agree on a given issue are also friends:

$$\text{Relational attraction} = \frac{\#\triangle_{\mathsf{A}}}{\#\triangle_{\mathsf{A}} + \#\wedge_{\mathsf{A}}} \tag{2.3}$$

The second component is *relational repulsion*. This is defined as the ratio of ordered triads in which two individuals who *dis*agree on an issue are *not* friends:

$$\text{Relational repulsion} = \frac{\#\wedge_{\mathsf{D}}}{\#\wedge_{\mathsf{D}} + \#\triangle_{\mathsf{D}}} \tag{2.4}$$

The two components can be interpreted as the empirical probabilities that a) i considers j a friend given they agree on an issue and that b) i does not consider j a friend given they disagree on an issue. The expected value of these probabilities will be affected by the general prevalence of friendship relations in the sample (equation 2.3 is expected to equal the friendship network density if no relational attraction exists).

Network polarization. Finally, we take the arithmetic mean of the two pairs of attraction and repulsion statistics to define a metric of ideological polarization and another of relational polarization. The maximum value for each metric is 1, which represents perfect polarization. Their minimum values are zero. In the extreme case where there is complete consensus on all issues that individuals have a non-neutral attitude about in the network, the maximum value of the ideological polarization metric is 0.5 (the repulsion statistic would be assumed to be zero as it is mathematically undefined, but the attraction statistic is 1). Under an empty social network or full social network, the mean of both relational polarization metrics is 0.5 again, as in the former case the attraction metric is assumed to be zero, and in the latter case, the repulsion metric is assumed to be zero. Together, this highlights the relevance of considering both attraction and repulsion patterns in the formation of ideologically and socially divided subgroups (Stadtfeld, Takács, & Vörös, 2020).

The two defined metrics place a community at a given point in time in a two-dimensional space of *network polarization*. The networks of some communities may show low levels of ideological and relational polarization; some may be high in one of the dimensions but low in the other; others may be highly polarized in both respects.

Figure 2.3 presents four hypothetical networks of four individuals and two political issues. The figures are placed in a coordinate system in which the horizontal axis measures the extent of relational polarization and the vertical the extent of ideological polarization. It can be seen in the figure that the four example networks occupy different quadrants in the system.

Due to the above-discussed biases induced by the network densities and degree distributions, we normalize each structure count in each of the four metrics by their expected value as described in the methods section. This aids the interpretation of the metrics: The 0.5 mid-point on each dimension is then the expected level of ideological/social repulsion/attraction conditioning on the observed network densities and degree distributions. If friendship were independent of attitude similarity in a given context (for example, when no attitude homophily and social influence on attitudes exists) one would expect the level of relational repulsion and attraction to be 0.5. The level of ideological repulsion and attraction, however, could still have a value larger than 0.5, indicating the presence of ideological polarization.

The metrics defined here may be useful for the comparison of polarization levels in different communities. Further, the approach is applicable to following temporal changes in polarization in a single community[2]. Beyond that, the multilevel network representation allows us to explore the social processes that may explain observed levels of network polarization.

---

2 It should be considered what this implies over time, however. Conditioning on the degree distribution means that as density changes (i.e. more non-neutral attitudes are added or some are dropped), the magnitude of polarization is still relative to the random expectation at that observation of the network. For instance, we might see the same values of polarization where people only have attitudes on two issues as a case where people have attitudes on 13 issues, provided they are organized in a way that produces the same proportions of the different structures considered in the metrics.

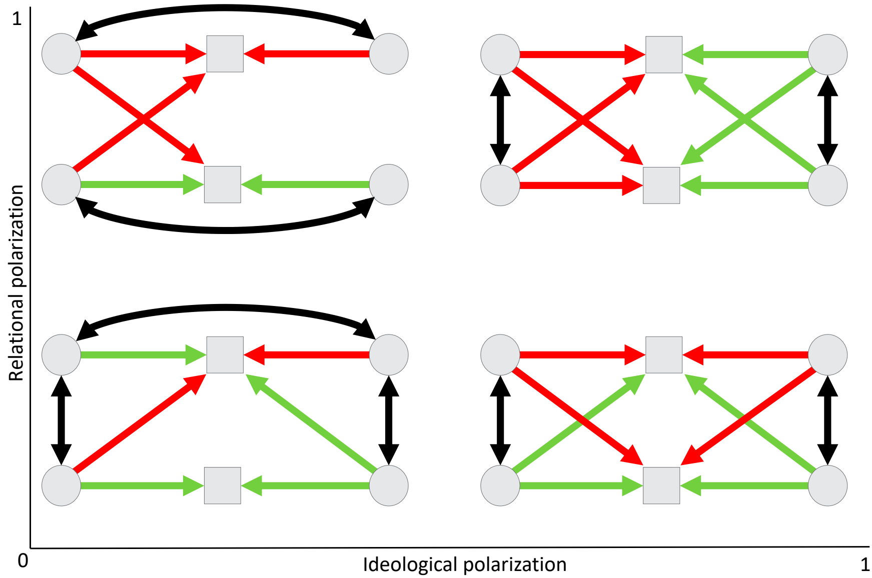


FIGURE 2.3: Stylized representation of the two-dimensional space of network polarization. Upper left represents a maximally relationally polarized, but minimally ideologically polarized configuration, and lower right represents the opposite. Squares represent issues, circles represent individuals. Black arrows represent friendships, while light green and dark red arrows represent positive and negative attitudes, respectively.

Becoming polarized

We have introduced a framework that allows us to study the level of polarization in society. Our network approach naturally offers a number of potential social mechanisms that may generate different levels of ideological and relational polarization, thus moving from the view of polarization as a state, to the dynamic, process-defined view.

Three general processes have been proposed in the network literature to explain the sorting of individuals by certain characteristics (such as their attitudes): homophilous selection, social influence, and latent similarity or exposure to similar contextual effects (Shalizi & Thomas, 2011). These classes of mechanisms may provide plausible explanations for network polarization. Here we review empirical evidence suggesting that each may be at play in ideological and relational polarization.

Homophilous selection: the dynamics of social ties

There is evidence in the political domain of homophily, i.e. a tendency to be connected to politically similar others. Social ties are suggested to be subject to homophilous selection on political attributes in the offline (e.g. romantic, marriage, informational discussion, and friendship ties, respectively; Alford et al., 2011; Huber & Malhotra, 2016; Huckfeldt & Sprague, 1987; Kandel, 1978) and online (primarily researched on Twitter networks, e.g. Boutyline & Willer, 2017; Colleoni, Rozza, & Arvidsson, 2014; Conover et al., 2011) realms, although offline analysis does not always support this proposition (for null findings on offline political selection, see Lazer et al., 2010; Wang, Lizardo, & Hachen, 2020). Furthermore, some studies suggest that individuals are not exposed to the opposing opinions of their social connections (Cowan & Baldassarri, 2018; Goel, Mason, & Watts, 2010; Kitts, 2003), which would reduce their ability to avoid social ties with disagreeing others.

Generically, however, people tend to be connected to similar others (McPherson, Smith-Lovin, & Cook, 2001). In a review of primarily social psychological literature, Huston and Levinger (1978) suggest that interpersonal attraction, leading to ties such as friendship, may be related to attitudinal similarity for several reasons. Two of these are relevant to the matter at hand: "Another's similarity may help confirm one's own feeling of rightness or goodness and indicate that the other is also good ... Another's similarity also gives promise of [their] future liking and favorableness toward us, and it can be argued that such inferences are responsible for its impact on our liking of [them]" (Huston & Levinger, 1978, p.125). In sum, both empirical and theoretical support is lent to the notion that (political) attitudinal similarity may be important for the structuring of one's social ties.

Social influence: the dynamics of political attitudes

While homophilous selection might explain similarity between connected individuals, another causal phenomenon examined more closely in the political and cultural domains is (assimilative) social influence. Several theories predict assimilative influence on political attitudes between connected individuals. The complexities of political issues may inhibit our ability to form consistent systems of beliefs ourselves (as observed by Converse, 1964) and lead to the use of actively shared information (Downs, 1957). Individuals may be convinced by arguments given by alters as to why one opinion may be the correct one to hold (Parker, Parker, & McCann, 2008). Furthermore, adopting similar opinions to one's friends may help fulfill relational goals such as a need for a shared reality or to enjoy smoother interactions with others (Jost, Federico, & Napier, 2009; Jost, Ledgerwood, & Hardin, 2008). Cognitive balance theory (Festinger, 1957) would have a similar prediction: the positive affect implied in friendship together with opposed attitudes may cause psychological discomfort; the individual may resolve this by adapting their attitudes to be more harmonious with those of others. All of the above gives us reason to expect influence on political attitudes.

There is also empirical support for social influence on political attitudes stemming from one's social ties. Experimental evidence of networked communication of strangers on various issues suggests that individuals have an assimilative influence on one another, albeit not with consistent effects (eg. Friedkin, 1999). Evidence has been found for influence of persistent social ties on college students' self-placement on a left-right political spectrum (Lazer et al., 2020; although Wang, Lizardo & Hachen, 2020 find no such effect) and stability of their issue attitudes (Levitan & Visser, 2009). Further, the vote choices of citizens have been shown to be affected by the choices of connected others in both the U.S. (Huckfeldt & Sprague, 1987) and the U.K. (Bello & Rolfe, 2014)[3].

Do social influence effects contribute to polarization? Simulation studies suggest so. Much work applying agent-based modelling has examined the theoretical macro-level consequences of individual attitudes changing through

3 These empirical studies used a single individual attribute as their dependent variable. Our approach considers attitudes across multiple issues simultaneously.

interpersonal mechanisms. Typically, such models are based on theoretically grounded assumptions about how people influence one another (Flache et al., 2017), but they less often consider how people connect to one another (Sobkowicz, 2009). These models consider various types of influence, often in combination. Flache et al. (2017) categorize the ideal types into assimilative influence, by which connected individuals always become similar; similarity-biased influence, in which connected individuals only become more similar given some pre-existing similarity; and repulsive influence, in which dissimilarity breeds further dissimilarity.

In agent-based models of opinion dynamics, equilibria (i.e. stable end states) of the simulated systems fall into one of three categories: consensus, in which all individuals agree; bipolarization, where two camps form within which individuals hold identical attitudes, and between which individuals hold perfectly opposed attitudes; and clustering, in which multiple camps of densely connected individuals exist with shared attitudes within, but not between camps (Flache et al., 2017). Which outcome is reached depends on the combination and strength of the aforementioned styles of influence incorporated in the model.

An exemplary agent-based model given for *cultural* polarization-as-clustering (of which we may consider political polarization a type) explicitly using a network approach is offered by DellaPosta, Shi, and Macy (2015). Inspired by empirical correlations between ideologically unrelated variables, this model takes a stylized interpersonal network of densely clustered 'caves', between which there are fewer links. Individuals have some set of binary attributes (called 'tastes'). Similarity between connected individuals increases their chances of influencing each other's opposed attitudes (an assumption made based on research into biases in information assimilation; Lord, Ross, & Lepper, 1979) and thereby makes them more similar. In the long run, this creates clusters within which there are highly similar individuals, but between which individuals are almost entirely dissimilar.

Micro-level processes of inter-individual influence thus offer potential explanations for the macro-level outcome of polarization. Nonetheless, while existing models in the area use a theoretically-grounded static network structure, the dynamics of interpersonal networks are typically ignored. These networks, such as friendship networks, are known to change (Snijders, van de Bunt, & Steglich, 2010), for example, due to homophilous tie formation as discussed above. Therefore, we argue that both selection and influence need to be incorporated into models to explain network polarization.

Latent-cause reinforcement: individuals, contexts and the unobserved

While we have thus far focused our investigation on interpersonal processes, we also consider auxiliary hypotheses related to unobserved factors that may contribute to the reinforcement of similar or opposed attitudes between individuals, regardless of their (lack of) social connection. Accounting for such mechanisms would counter confounding in testing hypotheses on similarity between socially tied individuals, a problem prominently noted by Shalizi and Thomas (2011). We refer to this process as latent-cause reinforcement, and it could be explained by at least three possible mechanisms.

First, latent individual variables may cause attitudes on some issues to be consistently similar or opposed over time. The most prominent of such latent variables is the consumption of ideologically similar media which may induce attitudes on multiple topics, causing similarity, or in the case of consumption of ideologically opposed media, dissimilarity. The internet has been prominently considered as a cause of polarization, through new media including social networking sites and related phenomena such as filter bubbles and echo chambers (Pariser, 2011; Sunstein, 2007). Furthermore, increasing polarization of the political elites to whom the mass public may look via such media for cues (Druckman, Peterson, & Slothuus, 2013) may cause polarization. If individuals' sources of information, such as news outlets and other opinion leaders thus become more polarized, this may lead the public in the same direction. Those being exposed to the same information contexts could become increasingly similar in their political attitudes over time, without any social relation. Notably, however, some have debated the direct role of echo chambers and filter bubbles brought about by the internet (Boxell, Gentzkow, & Shapiro, 2017; Guess, 2021), and others have argued that online information may not be consequential for attitudes without being channeled through a trusted source (Weeks & Gil de Zúñiga, 2019), such as friends or family.

Second, unobserved relations between individuals could cause the reinforcement of attitude patterns over time between pairs of individuals. While not considering one another friends, some pairs of individuals discuss multiple topics in a classroom context or have mutual friends due to whom they communicate on occasion, and thereby may have an assimilative or repulsive influence on one another's attitudes.

These two explanations also align with cultural sociological theory suggesting that the cultural meaning, i.e. perceptions of what some behaviour or belief may represent and how sets of behaviours and beliefs fit together (i.e., how they co-occur), is affected by repeated observations of their co-occurrence (Goldberg & Stein, 2018)[4]. Individuals may learn the associations between attitudes on multiple issues both in the context of media consumption and direct interpersonal communication, and adjust their own beliefs to these patterns.

Third, the environment in which the empirical study is conducted may induce similarity in some attitudes. For example, in the sample of university students which we use in this study, many may have recently moved away from culturally homogeneous neighborhoods into a more heterogeneous city, similarly affecting their perceptions on issues related to ethnic attitudes. This could, for example, make their attitudes towards other ethnic groups more favorable through intergroup contact (Pettigrew, 1998). Some correlational evidence also suggests that commitment to environmental sustainability might be affected by university attendance (Cotton & Alcock, 2013).

Given these three considerations, we can expect some latent forces which reinforce similar and dissimilar attitudes to be at play. These should be accounted for in the explanation of observed levels of polarization.

4 Notably, these authors assume fully random selection of other individuals by whom an actor may be influenced, where actors learn from observation which objects fit together. This bears some similarity to the modelling of the latent causes given here. A key difference is that Goldberg and Stein (2018) explicitly intend to avoid social selection and influence processes and focus on shared understandings. These processes are incorporated in the current work via endogenously changing friendship networks, where Goldberg and Stein instead focus on a model of everyday life without deep social ties.

Research questions and hypotheses

Following the multilevel network approach to polarization described above, we aim to answer three research questions:

1. What is the extent and trajectory of network polarization in two empirically observed communities?
2. Do three types of process – social selection, social influence, and latent-cause reinforcement – affect the evolution of the multilevel networks of social ties and political attitudes?
3. Do these three types of process explain the observed levels of network polarization?

To answer the first question, we examine network polarization in our empirical data using the metrics and visual representation introduced above. To answer the second question, we apply statistical models for dynamic networks on our dataset to explore the interpersonal and latent processes shaping political attitudes and social ties. We answer our third question by examining model fit: we assess how much our models from question two are able to explain the observed levels of polarization established in question one.

In answering the second question, we test three hypotheses. First, in line with the aforementioned theory and evidence suggesting homophilous social selection occurs on political variables, we hypothesize that:

<u>H1: Social selection</u>

H1a (attraction of attitude similarity): The more individuals agree on political topics, the more likely they are to be friends over time.

H1b (repulsion of attitude dissimilarity): The more individuals disagree on political topics, the less likely they are to be friends over time.

In terms of our multilevel network approach, this would be represented as shown in Figure 2.4 with the correspondingly-marked hypotheses: The probability of a friendship tie between two individuals is expected to increase with the number of shared, same-valenced connections to any political issue node

(*H1a*), and decrease with the number of shared, opposed valence connections to any political node (*H1b*).

Second, in line with research suggesting influence between individuals on political variables, we hypothesize that

<u>H2: Social influence</u>

H2a (influence of friends): Individuals who are friends are more likely to share attitudes over time.

H2b (non-opposition to friends): Individuals who are friends are less likely to have opposed attitudes over time.

In our multilevel network formulation, this is represented as shown in Figure 2.4 with the correspondingly-marked hypotheses. Here, the probability of an attitude tie of a given valence is expected to increase with the number of friends having the same attitude to the political issue node. For H2a, we would expect individuals sharing a friendship tie to have a higher probability to create or maintain an attitude that is the same as that of their friends. For H2b, we expect a decreased probability of maintaining or forming an attitudinal tie in opposition to that of their friends.

Third, while we are most interested in direct interpersonal processes in the network, we recognize that there may be important unobserved and/or contextual factors that may lead to the reinforcement of (dis)similarity over time. We thus formulate the hypothesis that:

<u>H3: Latent-cause reinforcement</u>

H3a (latent-cause convergence): The more similar individuals are in some political attitudes, the more likely they are to be similar in others.

H3b (latent-cause divergence): The more dissimilar individuals are in some political attitudes, the more likely they are to be dissimilar in others.

In our multilevel network formulation, this is represented as shown in Figure 2.4 with the correspondingly-marked hypotheses. Here, the probability of an attitude tie of a given valence is expected to increase with the number of others having the same attitude to the issue and the number of further issues on which they agree with the focal individual, whose attitude tie is considered (*H3a*). In the diverging formulation, the attitudinal tie is expected to be more likely if more others who have opposite attitudes on different issues have opposite attitudes on the issue at hand as well (*H3b*) [5].

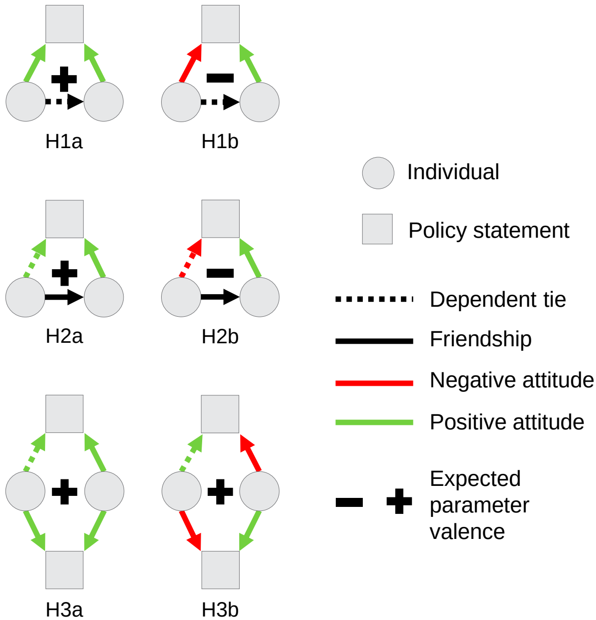


FIGURE 2.4: Depiction of multilevel network hypotheses. H1: Selection. H2: Influence. H3: Latent-cause reinforcement.

5 It is important to note here that while these structural dynamic hypotheses explicitly refer to generating the structures depicted, as a consequence of being endogenous they will also generate other structures relating to network polarization. For instance, a pair of friends who come to agree on two topics have not only added two triangles to the numerator and denominator of social attraction (see Formula 2.4), but will also have created at least one agreeing four-cycle, impacting the level of ideological attraction (see Formula 2.2).

DATA

To answer our research questions, we collected longitudinal survey data on friendships and political attitudes in two student communities in the scope of the Swiss StudentLife study. The data come from two undergraduate student cohorts, majoring in two different STEM subjects, at a technical university in Switzerland. We use data from the first five survey waves of the study (denoted W1-W5), which were collected between September 2017, the first week of the degree program, and May 2018, the end of the first academic year. We will refer to the two cohorts as Cohort 1 and Cohort 2 from here onward. The cohorts comprise respectively a total of $N_1 = 261$ and $N_2 = 660$ registered students, of whom 72% and 75% participated in any of the five surveys. Response rates amongst students belonging to the cohorts at the time of each survey varied between 77% and 60% for Cohort 1, and 76% and 48% for Cohort 2[6]. For demographic details including year of birth, gender proportion, political orientation, and nationality see Table 2.2.

The surveys were online, and conducted in German using the Qualtrics platform. The data collection plan was examined and approved by the institutional ethics committee of ETH Zürich. Details of the data collection and ethical considerations are discussed extensively in Vörös et al. (2021). A visualization of the final observation used in Cohort 1 is presented in Figure 2.2.

6 See Vörös et al. (2021) for an extensive discussion of how missingness was determined. This is likely a slightly downward-biased estimate of response rates due to the over-inclusion of individuals in the network.

TABLE 2.2: Individual descriptives

	Cohort 1		Cohort 2	
	M	*SD*	*M*	*SD*
Proportion female	.34		.13	
Proportion Swiss national	.88		.90	
Birth year	1997	1.96	1997	1.60
W1 political orientation	4.00	2.16	4.58	2.07
W5 political orientation	4.29	2.09	4.76	1.94

Note. Political orientation ranged from 0 to 10, and was assessed by a standard left-right self-placement scale adapted from the European Social Survey: "In politics people sometimes talk of 'left' and 'right'. Where would you place yourself on this scale, where 0 means the left and 10 means the right?" (European Social Survey ERIC, 2020).

Friendship networks

The friendship networks were collected in every survey wave by a hybrid roster-name generator approach (see e.g. Robins, 2015, chapter 5). Individuals were asked the question: "Whom do you consider a friend?". They could name up to 20 people in their answers. The names were limited to those of fellow students and could be chosen from an auto-completing list that appeared as participants began to type a name. Each reported name was coded as a directed tie in the friendship network. The friendship data from each wave is an adjacency matrix in which an entry '1' represents that a student nominated one of their peers as a friend; the value '0' means the absence of a friendship nomination.

Political attitude networks

The political attitude networks were constructed from answers to a 22-item battery of political statements asked in every survey wave. We provide a summary of the topics covered by these statements here and report the full list of items in Appendix A.3.

The attitude statements were taken from two sources. First, from an online tool created by the German Federal Agency for Civic Education, which was designed to give individuals advice on the German political parties best representing their views – the Wahl-O-Mat[7]. We selected a subset of the items that did not solely apply to the German political system, but were face-valid issues for much of Western Europe at the time. For example, items covered whether there should be mandatory child vaccination against contagious diseases, government-regulated cannabis sales, increased state support for social housing, and expansion of video surveillance in public spaces.

The second source of our attitude statements was the list of upcoming national referenda in Switzerland, sourced from the Swiss Federal Chancellery[8]. The referendum topics at the time of our study included, for instance, whether mandatory radio and television fees should be replaced by commercial financing, more bike paths should be added to public roads, and imported foods should be held to standards set for Swiss-cultivated products.

Each statement taken from these two sources comprised a position that a policy should attempt to fulfill. Participants indicated their (dis)agreement with each statement on a 7-point Likert scale: "strongly disagree" (= 1), "disagree", "somewhat disagree" through the midpoint, "neither agree nor disagree", to "somewhat agree", "agree" and "strongly agree" (= 7). An additional option was provided for "no opinion", which we recoded to the same value as the midpoint of the scale for the present analysis. In a robustness check, it was

7 Bundeszentrale fur politische Bildung (2017)

8 Referenda on proposed changes to the law and/or constitution, initiated either by popular groups or the government, are held up to four times to per year. All Swiss citizens resident in the country are eligible to vote in these referenda (Swiss Federal Chancellery, 2017a, 2017b).

coded as missing[9]. Appendix A.4 shows means and standard deviations of these variables by cohort at the first and last observations.

Based on these items, we defined two two-mode political attitude networks for each survey wave. In both of these networks, ties connect students to statements depending on their reported (dis)agreement with them in the given wave, or ties may be absent in case of a neutral opinion. One type of network represents agreement with the items – we label this as the "positive" attitude network. The other type represents disagreements and is labelled as the "negative" attitude network[10].

We used the following thresholds to recode the original 7-point attitude scales to the two binary attitude networks. The three middle options ("somewhat agree", "neither agree nor disagree", and "somewhat disagree") represented a neutral attitude and were coded as '0' in both networks. The two outer options on either end of the scale ("agree", "strongly agree" vs. "disagree", "strongly disagree") were treated as a tie and were coded as '1' in the positive and negative two-mode network, respectively. We checked the robustness of our models by using two alternative thresholds for defining the attitude networks, one using only the extreme values, one using values adjacent to the midpoint[11]. Thus, towards each of the 22 political statements per wave, each individual could have either a '1' in the positive network and a '0' in the negative network, a '0' in the positive network and a '1' in the negative network, or a '0' in both networks.

9 See Appendix A.5. In summary, weaker support is found for H1a in Cohort 1, but stronger in Cohort 2, than in the main analysis. Consistent support for H3a and H3b is found in both cohorts.

10 We note that the labels "positive" and "negative" are arbitrary, since statements could have been phrased either way around. However, this does not affect our analyses, which focus on network patterns involving a single item and the *consistency* of patterns across multiple items. Our key metrics and model terms are invariant to item reversal.

11 Here, we found that the model using values adjacent to the midpoint did not converge to a stable solution; this is likely due to the minimal change here (i.e., participants' attitudes tended to shift within the same side of the scale). See Appendix A.5. H3a and H3b are supported in both cohorts, while H1a and H2a are supported only in Cohort 1 under the higher-threshold model.

Individual background variables

In the dynamic network model, we accounted for the impact of various individual attributes on the changes in the friendship network. These included gender, age, and language spoken with one's family. Further, we used a categorical variable to represent the four different sub-majors of Cohort 1 (Cohort 2 consisted of a single major). Gender and age were included as they are fundamentally important in social networks research (McPherson, Smith-Lovin, & Cook, 2001, p. 417). In the empirical context, where a variety of languages were spoken both by Swiss students and those from abroad, language had a clear potential to help or hinder friendships. Finally, sub-major was included in our analyses since students in the same sub-major they shared more classes in Cohort I, which affected their opportunities to become friends[12].

METHODS

Assessing the level of network polarization

The two metrics of network polarization in Formulas 2.1 to 2.4 can be used to assess the extent of polarization in communities. However, the number of ties in each network may affect the theoretical maximum level of polarization. For this reason, we divide each term in each metric by its expectation given the distribution of ties in the network. We generate an expectation for the ideological aspect of polarization by rewiring the attitude network randomly, holding the in- and outdegrees of all nodes in the positive and negative attitude networks constant and leaving missing tie data in place. Similarly, we generate an

12 We asked participants' left-right political orientation using the German version of a question from the International Social Survey Programme (International Social Survey Program Research Group, 2016). The question is phrased as follows: "In politics people sometimes talk of left and right. Where would you place yourself on a scale from 0 to 10 where 0 means the left and 10 means the right?". In a robustness analysis, we controlled for similarity on this variable, which has been suggested to be a causal factor in connected people's political attitudes (Lazer et al., 2010) or be a source of tie homophily (Kandel, 1978). The inclusion of the parameter did not change results substantively.

expectation for the relational aspect of polarization by permuting the friendship network and recalculating the relevant terms, holding the attitude network constant. We take the average of each formula term from 1000 networks, generated separately for each observed network and not counting structures involving missing tie data. This way of normalizing the network polarization sub-metrics, as discussed in the section explaining network polarization, yields an expected midpoint on each dimension of 0.5.

The resulting metrics have a range from zero to one, with values above the midpoint being more polarized, and values below less polarized, than the expectation. To test whether the observed level of polarization is significantly different from the midpoint, we compare the observed values of the metrics to the distribution of random outcomes generated by the procedure above, thereby generating non-parametric p-values.

Modelling friendship-attitude dynamics

Stochastic actor-oriented models for multilevel networks

We apply stochastic actor-oriented models (SAOMs), which are suited to examine the co-evolution of the multilevel network (Snijders, Lomi, & Torló, 2013) of friendships between individuals, i.e. a one-mode network, and positive and negative attitudes from individuals to political statements, i.e. two two-mode networks. We used the implementation of this model in RSiena (Ripley et al., 2019), version 1.2-16, modified with new structural effects that were implemented for this article in Appendix A.6. The model was estimated by the Method of Moments (Snijders, 2001)[13]. Missing data in all networks were treated by the default method in RSiena: if available, the last observation is carried forward, otherwise, the tie is imputed as zero. Notably, these imputed

13 For each model specification, the default algorithm settings were used, until a reasonable level of convergence was reached (overall maximum t-ratio $< .30$, parameter-wise maximum t-ratio $< .15$). After this, the resulting model was used for a longer single additional phase 2 iteration, and 5000 phase 3 iterations (Option 1, described in Section 6.4 of the manual, Ripley et al., 2019) to ensure convergence diagnostics reached the recommended cutoffs of $< .25$ and $< .10$ respectively. Furthermore, we treated the attitude networks as disjoint; an individual could not hold two opposed attitudes on a single topic at the same time.

(non-)ties are not used directly in the calculation of structural statistics and therefore standard errors, but are used in the simulations via which the models are estimated.

The SAOM is a statistical model for tie changes in dynamic networks. It aims to estimate the change in relative probability of actors' choice of ties to create and maintain, given the relevant local network structures specified by the modeller. It is assumed that the network observation at time T_n transitions to the network observation at time T_{n+1} by a process in which individuals take turns to update their outgoing ties to other entities (in our case, other individuals and political items) based on the current state of their local network. Thus, in this model, changes in one network can influence changes made in another – for example, if individuals Iain (i) and Julia (j) are tied by friendship, and j notes that many in her friendship network have a negative attitude tie towards arms exports, she may decide she has a negative attitude tie towards arms exports, thereby closing many possible triangles of friendship and negative arms export attitude ties. In a later phase, i, who has a positive attitude tie to arms trade, may choose to drop his friendship to j or to adopt a neutral attitude about the issue. Note that, as an example, a positive tendency towards sharing an attitude with friends not only means that individuals are more likely to gain attitudes in line with their friends', but are also less likely to change attitudes that are already consistent with their friends'.

Model specification

In our main model, we include all parameters that are necessary to test our hypotheses. In addition, we include standard structural terms for friendship networks: reciprocity, a three-parameter specification to capture clustering, outdegree popularity, indegree popularity, and outdegree activity to capture mechanisms generating degree distributions, as well as effects of gender on sending and receiving ties. Finally, we included terms for homophily on the background variables (gender, age and language spoken with family). Similarly to the friendship network, we include terms for indegree popularity on the attitude networks, accounting for the fact that individuals may be more likely to have stronger attitudes on some topics than others.

Due to the high level of change in friendships between waves 1 and 2 (wave 1 was collected during the first university week when few friendship relations had been established, and wave 2 after one month), we estimate the models for change in the networks from waves 2-5 separately from wave 1-2. As the negative and positive attitude networks are treated as separate networks, all hypotheses are represented by duplicate parameters: once for each pairing of the two attitudinal tie valences. From Figure 2.4, all hypothesized structures pictured are thus also included with each dark red and each light green tie swapped for the alternate color. As a concrete example, a friendship is dependent on both a shared positive attitude of some political statement (homophily from positive opinion), and on a shared negative attitude of another political statement (homophily from negative opinion). See Appendix A.7 for two tables with a graphical depiction of the complete main parameters, and Appendix A.8 for standard RSiena notation for all parameters.

As the conceptual approach to the model is novel, several effects were newly implemented in RSiena. The first effect represents the tendency against having friendships in which disagreement occurs (corresponding to *H1b*), and the second represented the expected tendency against having contrary attitudes to one's friends (corresponding to *H2b*). Furthermore, two effects were implemented to capture the tendency to become less similar to others with whom one disagrees (corresponding to *H3b*). These effects are also indicated in the aforementioned tables in Appendix A.7.

Linking polarization to micro-processes and macro outcomes

We now assess through simulations in how far the process-based micro-level model is able to reproduce the observed static macro-level outcome of network polarization (Snijders & Steglich, 2015; Stadtfeld, 2018). To study the micro-to-macro link, we follow the goodness-of-fit method proposed for SAOMs in Lospinoso and Snijders (2019). This method first populates distributions of selected structural statistics from networks simulated from the model over a fixed period, starting from the observed data at the earlier time point. The final simulated distributions are then compared with the same statistics in the observed network at the following time point.

In our case, we calculate the level of polarization in simulated networks used for goodness-of-fit testing, following the exact same procedure applied to the observed data. We can then compare the distribution of levels of polarization in these simulated networks to the level of polarization in our observations. Since the simulations are generated from the estimates of micro-level processes, they help us to understand if the estimated micro-level *model* produces, at the macro-level, polarization similar to the macro-level *observation*.

RESULTS

Descriptives

Table 2.2 shows that there is a slight leftward bias in political orientation in both cohorts. The average orientation is to the left of the midpoint of the scale throughout the observation period. This difference was usually statistically significant, with $p < 0.05$ in all cases for Cohort 1, and all observations except three and four, for Cohort 2. In these two exceptions, means were 4.79, ($p = .059$), and 4.87, ($p = .197$), respectively, using one-sample t-tests. The size of the difference is thus not substantial: the average is never more than one point below the midpoint of the 0-10 scale.

Since data were collected from the beginning of the first week of the participants' study program, the students had not yet developed many friendships to report at the start: The average number of friends chosen (mean outdegree) grows substantially from Wave 1 to Wave 2 in both cohorts: from 0.92 ($SD = 1.73$) to 2.39 ($SD = 3.17$) in Cohort 1, and from 1.54 ($SD = 1.98$) to 1.93 ($SD = 2.63$) in Cohort 2. Using the Jaccard index, we quantify the stability of friendships between subsequent survey waves. This measure indicates the ratio of ties that are present in both Wave t and Wave t-1 to ties that are present in either one or both. A zero indicates complete instability, while one indicates complete stability. Friendship networks became increasingly stable over time:

starting at .37 and .55 between waves 1 and 2 for Cohort 1 and 2 respectively, they increased to .66–.76 in both cohorts later. Appendix A.9 shows the full descriptive statistics of the friendship networks for the two cohorts.

Turning to the attitudinal networks, the participants were consistently more likely to be in favour of a statement than against it. The fraction of existing ties (densities) ranged from .30 to .37 in the positive network and from .21 to .25 in the negative network, across cohorts and time points. Overall, participants had a total average of 5.05–6.76 attitude ties across both attitude networks in Cohort 1, and 4.44–6.80 in Cohort 2. In both cases, the highest observed average number of ties was at Wave 1, and the lowest was at Wave 5. Attitude ties were largely stable over time, with a Jaccard index ranging from .69 to .78 for all pairs of subsequent waves in both cohorts. Appendix A.9 shows the full descriptive statistics of the political attitude networks for the two cohorts.

How polarized are the networks?

To answer our first research question, which asks about the extent and trajectory of network polarization in our two empirically observed communities, we first look at the development of these metrics over time. We look at the levels of polarization as defined in Formulas 2.1 to 2.4, using the normalization described in the first methods subsection 'Assessing the level network polarization': a relational expectation defined by permutation of the friendship network, and an ideological expectation defined by a rewiring of attitudinal ties while preserving degree distributions at the level of both the person and the object. A graphical overview of its development of the metrics is shown in Figure 2.5.

We see that network polarization is consistently higher than the random expectation of 0.5 on each dimension (relational and ideological) and slightly increased from the first until the last observation. However, intermediate time points don't show a clear trend and all points are quite tightly clustered near the midpoint of the scale. Overall, it seems that Cohort 1 is slightly more polarized than Cohort 2. Ideological polarization appears stronger than relational polarization in both cohorts. This is likely in part due to the fact that relational

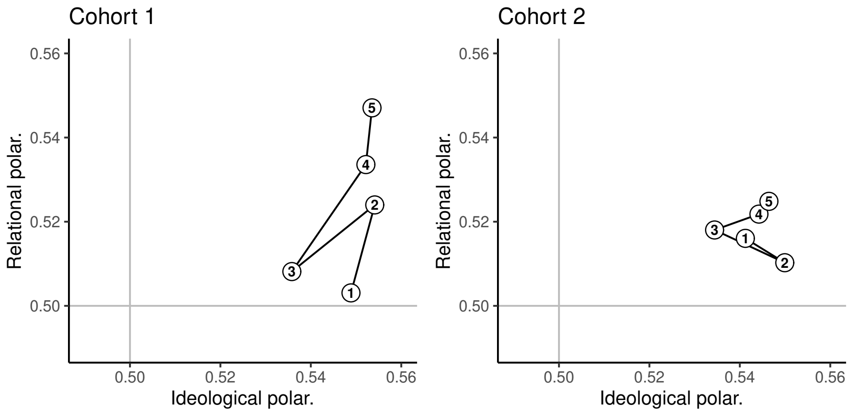


FIGURE 2.5: Two-dimensional network polarization in two cohorts, with each circle representing one observation point, indicated by the digit in the circle. Axes are truncated. There is no clear trend towards polarization over time in our data, and the data are tightly clustered relative to the range of the measure.

polarization can only occur by processes occurring within the cohort, while ideological polarization may also have occurred earlier. Overall, the communities appear to become more polarized over the year, but only to a small extent and through a noisy process.

Figure 2.6 shows the observed values of the polarization metrics alongside a set of values drawn from our null models in the two cohorts in waves 1 and 5. The figure can be read as a set of non-parametric tests for the differences between the observed metrics and their expected values (0.5). These investigate whether the differences from the 0.5 mid-values are significant. The larger red circles indicate the observed values of the metrics. The cloud of smaller, light grey points represents values in random draws from the null distributions, with the blue diamond marking their means. In all cases, observed ideological polarization lies clearly to the right of the vertical dotted line, showing that its value is significantly above expectation ($p < .001$ in all cases). Results for relational polarization are mixed: at wave 1, the metric in Cohort 1 is close to the expected value ($p = .480$), while it is significantly above it in Cohort 2

($p = .011$). By wave 5, however, relational polarization is significantly above expectation in both cohorts (Cohort 1: $p < .001$; Cohort 2: $p = .002$).

Which of the hypothesized processes affect the evolution of the social-attitudinal network?

Having examined the extent of network polarization in our student communities, we now present the dynamic models explaining changes in the multilevel networks. Here we report results from the micro-level model testing hypotheses 1-3, which stated 1) a positive expectation of selection on political similarity or negative on dissimilarity, 2) a positive expectation of influence on shared attitudes or a negative expectation on opposed attitudes, and finally, 3) positive effects of disagreements on further disagreement, and positive effects of agreements on further agreement. Appendix A.11 describes and shows goodness-of-fit on relevant structures which the estimated model should reproduce, following Lospinoso and Snijders (2019). These structures are primarily those that are also used in the calculation of network polarization. Counts of mixed triad structures are based on those presented in Hollway et al. (2017). For the model fit across all included waves, the model does not significantly misfit on either tetradic person-issue-person-issue structures (Cohort 1: $p = .998$, Cohort 2: $p = .323$) nor on triadic person-person-issue structures (Cohort 1: $p = .437$, Cohort 2: $p = .693$)[14][15].

The results of this model are consistent but not identical across cohorts. Each hypothesis yields multiple parameters in the model due to the coding of positive and negative attitudes as two separate networks: Two for each of hypotheses 1a-2b, and four of each for hypotheses 3a and b. For this reason, we simplify interpretation by using Wald tests of the linear combination of

14 Periodwise goodness-of-fit statistics are presented in Appendix A.12. The models tend to generate too few of most triads in the period from waves 2 to 3, and too many from waves 4 to 5, while waves 3 to 4 seem noisier but less biased towards generating too many or too few of all structures. A similar pattern is observed for four-cycles. In both cases, fit tends to improve over time. Given the high number of parameters included in the model, however, we opt to retain the uniform parameter estimates to reduce the risk of overfitting and type II errors.

15 Other fit statistics focused on degree distributions in all networks, and other structures in friendship networks, see Appendix A.13.

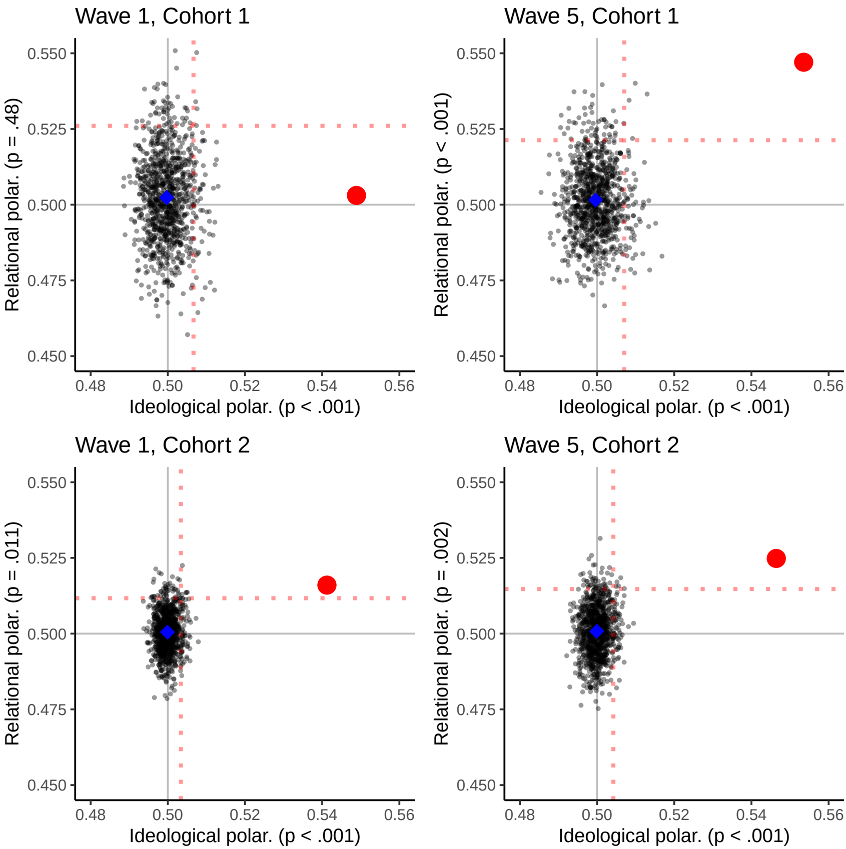
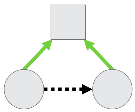
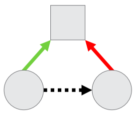
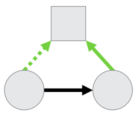
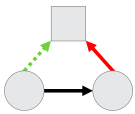
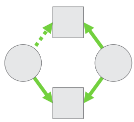
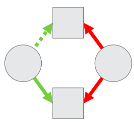


Figure 2.6: Two-dimensional observed network polarization (red circle) relative to expectation in two cohorts (blue diamond centered in the point cloud). Dotted lines indicate the boundary for $p < .05$ (one-tailed). The horizontal line indicates relational polarization, while the vertical line indicates ideological polarization

respective groups of parameters, testing whether the sum of the parameters differs from zero. For these main results, see Table 2.3. For the raw parameter estimates, see Appendix A.7, and see Appendix A.5 for test statistics for this model and robustness checks.

TABLE 2.3: Joint tests of SAOM estimates by hypothesis

Hypothesis	Structure	Summed parameter	
		Cohort 1	Cohort 2
H1a		0.13^{***}	0.08^{*}
H1b		-0.10	0.08[b]
H2a		0.12^{**}	-0.02[b]
H2b		-0.02	-0.09
H3a[a]		0.26^{***}	0.14^{***}
H3b[a]		0.10^{***}	0.03^{***}

p-value from Wald tests of summed parameters.

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Dependent tie is dotted line.

[a] Parameter multiplied by 10. [b] contra-hypothesis parameter direction.

Regarding *H1a*, that individuals select each other for friendship based on political agreement, we found a positive and significant effect for both parameters in Cohort 1, and one in Cohort 2. Using Wald tests of summed parameters, we find evidence of selection on political similarity in both cohorts, supporting H1a: individuals are more likely to form and maintain friendships to alters with whom they share more political attitudes.

Examining *H1b*, that individuals tend to avoid friendships with those with whom they disagree, none of the parameters are significant, and indeed are in the opposite direction to expectation for Cohort 2. The Wald tests also fail to reject the null hypothesis. We thus find no evidence that people are increasingly likely to avoid and discontinue their friendships with people with whom they share more opposed political attitudes.

For *H2a*, that individuals connected by friendship are more likely to adopt and maintain similar attitudes over time, results were mixed. One of our positive influence effects was found to be positive and significant in Cohort 1. In Cohort 2, both parameters were found to be in the opposite direction to expectation, though neither significantly so. The Wald tests support the positive influence hypothesis in Cohort 1, and find the opposite-expectation parameters in Cohort 2 to be non-significant. We thus find evidence of social influence on shared attitudes in Cohort 1.

Examining *H2b*, that individuals will avoid or drop attitudes that make them dissimilar from those with whom they are befriended, we find one significant effect in the expected direction, in Cohort 1. In neither cohort do Wald tests suggest evidence in favour of H2b. Overall, we find minimal evidence for the combined *H2*, i.e. for social influence on political attitudes.

Finally, *H3* stated that existing attitude patterns should reinforce over time: attitudinally similar individuals should tend to gain or maintain agreement over time, while dissimilar individuals should tend to gain or maintain disagreement over time. The final two rows of Table 2.3 show the structural representations of latent-cause similarity and dissimilarity alongside the results of the Wald tests of summed parameters. According to *H3a* shared attitudes between two individuals increase the probability of gaining and maintaining shared attitudes in the future. Generally, this hypothesis is supported by the positive and significant parameter estimates in three out of four agreeing attitude configurations in both cohorts, and in both cohorts the Wald test supports the hypothesis. The effect of disagreeing structures is positive and significant in four out of four tests in both cohorts. This too is supported in both cohorts by a Wald test. This suggests, in line with *H3b*, that the more opposed individuals' sets of attitudes are, the more likely they are to gain and maintain more opposed sets of attitudes over time.

Similarly to the static polarization metrics, we thus see consistent evidence of ideological structuring, but weaker evidence of relational processes supporting polarization. Particularly, H1a (positive social selection), H3a (latent-cause similarity), and H3b (latent-cause dissimilarity) are consistently supported. H2a (positive influence) is supported only in Cohort 1, while H1b and H2b are not supported in either cohort. For results of robustness checks, which largely support the main results, see Appendix A.5[16].

All models controlled for structural processes in the social networks, such as reciprocity, transitivity, and popularity, structural processes in the affiliation networks, such as item popularity, and attribute-related processes such as attribute activity, attraction, and homophily. For a full results table in standard RSiena notation, alongside brief interpretations, see Appendix A.8.

Do the three types of process explain network polarization?

To complete our answer to the question of whether the micro-level processes tested above also produce the macro-level network polarization observed, we calculate the network polarization index on a random sample of 100 networks that were simulated from the estimated model at the end of each of the four observation periods. The simulations were performed with the exact stochastic actor-oriented model presented in the previous section (i.e. at the expected

16 In tests in Appendix A.10, splitting the tests into the components of repulsion and attraction by each wave, we find only a significantly raised level of social repulsion in Cohort 1 at Wave 2, and both cohorts at Wave 4 and 5. Social attraction is only significantly raised at Wave 1 in Cohort 2. This is somewhat counter to the pattern of results from the model, which suggests a significant tendency towards social attraction structures.

outcomes on the second, third, fourth, and fifth observation)[17]. Each of the 400 simulated outcomes' level of network polarization (100 at the end of each period) is normalized with a sample of networks with the same degree distribution, again compared to a .5 expectation. The same normalization procedure was performed earlier to investigate the static level of network polarization in the data. Taken together, we see the range of results on the macro-level outcomes as a result of the estimated micro-level processes. In Figure 2.7 we show the observed values with the larger red diamond indicating the mean across periods, while the simulation polarization values corrected for the degree-preserving expectations are in grey, while the smaller blue diamond in the center of the point cloud represents their mean. Overall, the average observed network polarization metrics tend to lie within the cloud of simulated values, indicating that the model generates the observed macro-level outcome reasonably well. However, some of the more extreme values are found in the periphery of the distribution. In Cohort 1, the micro model appropriately generates levels of macro polarization, although at the last observation there is more relational polarization than expected. There is some suggestion that the process model produces too little ideological polarization in Cohort 2, as three of the four observed values, relative to their expectation, are more polarized than the simulated data. In both cohorts, the estimated model produces more polarized configurations than under the respective random expectations despite the non-significance and reversal of some hypothesized model parameters. In sum, the micro-level model seems to explain the observed macro-level metrics of network polarization well, with the important caveat that in both cohorts at most time points there is a small but consistent bias towards producing less polarization than observed on either dimension.

17 We exclude the first observation, as the simulations are the result of running the model from one observation to the next. We thus never simulate until the first observation.

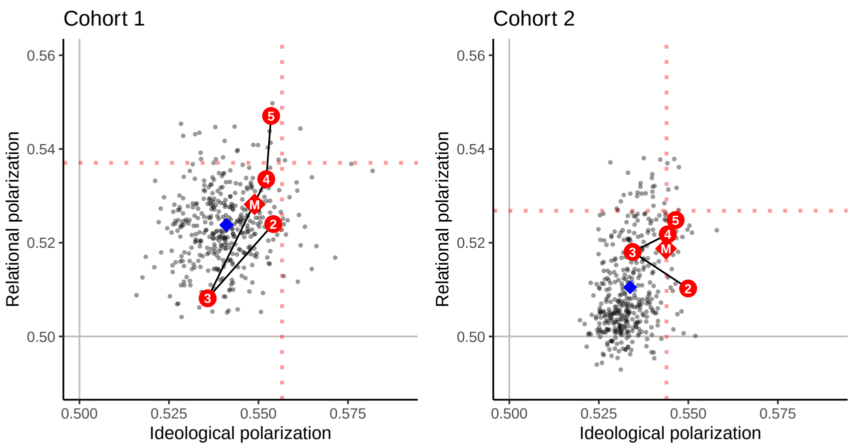


FIGURE 2.7: Polarization in observation and from micro-model simulation. The former is represented by larger connected red circles, while the latter is represented by separate light grey dots. The mean of the simulated values is represented by the blue diamond centered in the point cloud. The larger red diamond (M) presents the mean of the observed statistics. Micro simulations included are a random sample of 100 simulations for each period, for a total of 400 points, horizontal and vertical dotted lines indicate one-tailed p-values from this distribution.

DISCUSSION

We have introduced a dynamic multilevel network framework for the study of political polarization in society. We defined the novel concept of network polarization, encompassing current perspectives on political division as an ideological and relational phenomenon. We proposed a set of measures that allow for a quantification of the extent and significance of network polarization in a multilevel network of political attitudes and social ties at a given moment in time. First, we assessed ideological polarization – as alignment, or constraint – of actors' attitudes. Secondly, we assessed relational polarization – the co-occurrence of agreeing attitudes, and friendships. Additionally, we estimated dynamic network models on a novel longitudinal dataset and assessed which social processes explain the observed levels of network polarization in two cohorts of Swiss undergraduates over time. We find robust evidence for the selection of alters on the basis of overlapping attitudes on political topics and latent-cause reinforcement, alongside weaker evidence for social influence. These processes largely explain the mild level of polarization found in this context.

Interestingly, we do not find significant evidence for selection nor influence of directly opposed attitudes. A potential explanation is that people may try to avoid sharing their opinions when they are expected to be opposed (Cowan & Baldassarri, 2018; Kitts, 2003), and may systematically misperceive the attitudes of alters in the case of disagreement (Goel, Mason, & Watts, 2010). While such an explanation cannot be tested in our data, it may be that mild indicators of opinions are already sufficient to cause an avoidance of further topical discussion, hence preventing this from causing frequent dissolution of friendships or of discussion provoking opinion change.

A notable finding is that there is social influence on the adoption of a position in one of the cohorts, but not on the dropping of a position that is opposed to one's friends. This implies that there is little force back to a midpoint, but some force away from it. While the evidence for this former effect is only found in one cohort, it nonetheless aligns with arguments that political moderates will tend to shift away from the middle due to selection and influence (DellaPosta & Macy, 2015), and the empirical observation of high party polarization in Switzerland (Hänggli & Häusermann, 2015). Nonetheless, the lack of shift towards the midpoint under negative social influence does potentially imply some pluralism of political opinions.

Our work aims to contribute to the scientific study of polarization in societies and communities in real-life settings. The key theoretical strength of our network approach is that it can be used to study a wide variety of interpersonal (micro-level) processes jointly with their societal (macro-level) consequences. Our unique empirical dataset presents a viable example for collecting longitudinal network data about social divisions. Such data are straightforward to analyse using the presented framework. Finally, our extension of stochastic actor-oriented models with specific effects and goodness-of-fit statistics can be readily applied by researchers to the study of polarization.

There are a number of limitations to our work, which highlight productive directions for future research. First, we elaborated our approach focusing on political attitudes, or more precisely: people's attitudes towards a set of specific political topics. The literature on political polarization suggests that social divides are likely to exist in and spread through other domains as well (DellaPosta, Shi, & Macy, 2015). Our multilevel network framework allows the generalization from political topics to alternative social objects about which people may have positive and negative attitudes. This includes *other* political issues not captured in our data, but also other objects of cultural consumption and lifestyle choices. Studying these jointly with political attitudes is now possible and would be crucial for understanding a number of divides and conflicts in modern societies.

Second, we only studied a single type of social relation (friendships). Social network studies highlight that negative ties, such as dislike and conflict between people, jointly evolve with positive ones in communities and impact a variety of outcomes (Harrigan, Labianca, & Agneessens, 2020). Negative ties play a role in the formation of group boundaries (Stadtfeld, Takács, & Vörös, 2020) and some evidence suggests that outgroup differentiation may contribute to political polarization (Bail et al., 2018). However, direct evidence for the effects of negative ties on opinions in offline settings is thus far lacking (Takács, Flache, & Mäs, 2016). Our approach could be generalized to allow both positive and negative ties between individuals in a similar way as it incorporates positive and negative attitudes to topics. Such an extension can enable researchers to explore further processes that may explain polarization, such as the appearance of negative ties as a result of opposing views on political issues.

Third, we only used data about the offline social ties of individuals. We think that this choice is reasonable given our empirical context, where students had a chance to interact with their peers face to face almost daily most of the year. However, as political discussions increasingly occur on online platforms, often with exclusively online contacts, it would be crucial to test our approach in the context of social media. For example, social ties are readily observed and political attitudes can be inferred in online discussion communities, such as Reddit (e.g. An et al., 2019; Himelboim, McCreery, & Smith, 2013). As we argue in this paper, polarization in such datasets could be analysed as a multilevel network.

Fourth, and related to the previous point, we explored polarization in a small-scale empirical setting (and therefore results may not generalize to the broader population). However, our approach could be applied to the study of large communities. It is more feasible to collect large-scale multilevel network data about political attitudes and social ties in online settings, but offline ties could be observed or surveyed in future studies. The presented concepts and measures of polarization should scale, although more computationally efficient normalization strategies may be required for the attitude networks. Stochastic actor-oriented models are not yet suited for the analysis of communities larger than a few hundred individuals, but to our knowledge there are ongoing developments in this direction.

Fifth, a possible criticism of our work is that our empirical setting was not strongly polarized. This was confirmed both by the proposed metrics and by our informal knowledge of the empirical context. The main focus of this paper was to develop and test a multilevel network approach to studying polarization. The introduced metrics and models can be used to study communities that are extremely polarized and which are not polarized at all. We would find it highly valuable if future studies applied our approach in more polarized settings. Comparing results from a number of contexts may reveal important variations in how political polarization evolves in current societies. We should also highlight that our approach does not attempt to examine polarization through extreme attitudes on single topics, deferring instead to measures favouring multiple topics which treat extreme attitudes in the same way as more moderate ones. Incorporating extremity on single topics and smaller differences occurring across multiple may be fruitful in understanding polarization.

Lastly, we have to note that our approach does not explicitly incorporate emergent social identities, which are increasingly considered important to explaining political polarization (Iyengar et al., 2019). As people interact and develop attitudes to issues, group identities emerge (Tajfel & Turner, 1979). Once they exist, these identities may influence the attitudes and actions of individuals, beyond influence of specific peers. This may especially be important in the case when these are, or come to merge with, political identities. Future work should aim to account for emergent group identities in network models for a more complete understanding of political polarization.

Polarization is a process that, in the extreme, may disintegrate societies. This happens in at least two ways. The social aspect of polarization represents that dissimilar individuals in society avoid one another and "grow apart", while local structure may reinforce existing similarity. The ideological aspect reflects that they may even lose common ground for coming together. As a result, polarization can lead to long-term conflicts and fragmentation, undermining the functioning of societies in a number of domains, such as politics, culture, or the economy.

Polarized politics and the "culture war" have been in the center of public discourse in Western societies recently. However, the interpersonal processes leading to and sustaining polarization, such as ideological differences and the lack of social ties between groups, can arguably be found everywhere. Western societies might experience "polarization crises" at the moment, but the underlying problems are likely to be present in other societies to some extent as well, and a better understanding of these phenomena may bear fruit in the future.

It is tempting to view polarization as defined by extreme "fringe" beliefs, but it may be seen as a subtler yet more pervasive phenomenon. Media often focus on extreme groups, for instance (and with good reason) noting the rising popularity of far-right groups in recent years. Ideological conflicts around science and science education are driven by groups with anti-mainstream beliefs such as "anti-vaxxers" and "flat-earthers". However, societies may also be polarized in a milder sense through smaller differences along a variety of social-political topics, such as the role of the federal state in the US, the future of the EU in Europe, and approaches to immigration in many countries. Our work highlights that political polarization can take many forms, and that its extent and causes should be explored across societies.

We need to understand the processes that create and sustain political polarization in order to be able to tackle it. Our work highlights the importance, and enables the study of selection and influence processes in the context of polarization. If people tend to become more similar to those they talk to and dissimilar from others, ideological polarization is expected to increase. If, in turn, people tend to talk only to others with similar views, social polarization is also expected to increase. We have a long way to go to fully understand the dynamics of polarization. We believe that a dynamic social network approach, such as the one presented in this paper, can be helpful in this endeavor.

3

CHAPTER 3 – PEEKING FORWARD AT POLARIZATION: APPLYING A STOCHASTIC ACTOR-ORIENTED MODEL AS AN EMPIRICALLY-CALIBRATED AGENT-BASED MODEL

Agent-based models (ABMs) for opinion dynamics have provided theoretical evidence for conditions resulting in political polarization. Increasingly, calls are made for stronger links to empirical data. We demonstrate one promising method, using simulations from an empirically-estimated stochastic actor-oriented model (SAOM) of the coevolution of 22 political attitudes and social ties in a complete network of undergraduate students ($N = 261$). We do so to understand the consequences of estimated interindividual and background processes on relational and ideological polarization.

We first consider the benefits of using a SAOM as a tool for calibrating and validating an ABM, including linking to empirical data, incorporating evolving networks, and its time-based interpretation. We then counterfactually strengthen and weaken processes of homophilous selection, positive social influence, and global latent forces capturing convergence and divergence of opinions. Simulating forward from an observation of the social network, we find a tendency towards a reduction of polarization under virtually all manipulations of effects.

We find that the contribution of the manipulated effects is typically quite small, producing only slightly slower decreases in, or a slightly higher level of polarization in the long run as these mechanisms increase in strength. Notably, people tend to hold fewer opinions, but more friendships in the long run – suggesting that socializing forces outweigh the negative effects of polarizing processes.

This article is currently in preparation as:

Mepham, K., & Keijzer, M. (TBD). *Peeking forward at polarization: Applying a stochastic actor-oriented model as an empirically-calibrated agent-based model*

INTRODUCTION

Understanding the nature and potential causes of mass political polarization has been a prominent objective of modern political science research (Fiorina & Abrams, 2008; Iyengar et al., 2019). Concerns have been raised about societal breakdown linked to the phenomenon (Iyengar et al., 2019), and its presence has been documented in various nations (Iyengar & Westwood, 2015; Reiljan, 2020; Wagner, 2021). Polarization has been documented in both an ideological sense—considering opposition in ideologies—and an affective sense—considering a more negative view of the (ideological) outgroup compared to the ingroup (Abramowitz & Saunders, 2008; Fiorina & Abrams, 2008; Iyengar & Westwood, 2015; Kozlowski & Murphy, 2021; Reiljan, 2020; Wagner, 2021). Conceptually, the two phenomena are often clearly interlinked (Webster & Abramowitz, 2017).

Traditional social science explanations for the phenomena rest on psychological effects (e.g. identity-based phenomena and partisanship) or society-level trends (e.g. elite polarization, globalization, or changes in media landscape), which tell only part of the story. In reality, the emergence of polarization is a consequence of the actions and expressions of a large population of autonomous but interdependent agents who create, consume, and exchange information. Researchers working on agent-based models (ABM) have produced an extensive literature examining how opinion landscapes emerge from two processes: *influence* and *selection* in interpersonal contact to capture this dynamic and multi-faceted process (for reviews, see Castellano, Fortunato, & Loreto, 2009; Flache et al., 2017).

Influence, defined by Rashotte (2007) as "change in an individual's thoughts, feelings, attitudes, or behaviors that results from interaction with another individual or a group"[1] typically takes one of two main forms in the ABM literature. Firstly, as an assimilative social influence that makes connected indi-

1 We take this to include the possibility of change of *outcome*, meaning that not only change of an individual's 'thoughts, feelings, attitudes or behaviours' may be a consequence of influence, but also non-change where it would have otherwise occurred. For instance, if one had heard a convincing argument against animal testing that may convince them that it is wrong and started them on the path to changing their attitude on the topic, but a social contact offered

viduals more similar, secondly, a repulsive social influence that makes individuals more dissimilar—this latter form tends to be conditional on having some prior level of disagreement (e.g. DellaPosta, Shi, & Macy, 2015; Flache et al., 2017). The term 'influence' is used as shorthand for assimilative social influence throughout this paper unless otherwise specified.

Selection typically refers to homophilous social selection, i.e. a choice to be connected to similar others (McPherson & Smith-Lovin, 1987; McPherson, Smith-Lovin, & Cook, 2001). While empirically evidenced, this phenomenon is often somewhat indirectly modelled in ABM. The well-known bounded confidence model of Deffuant et al. (2000) and Hegselmann, Krause, et al. (2002) implements a somewhat analogous process: where connected individuals are too dissimilar, they are no longer capable of influencing one another. Note, though, that this does not change the underlying network but the channels of influence. More directly, models by Baldassarri and Bearman (2007) and Carley (1991) include a selection term for a hypothetical discussion network dependent on actors' similarity. Throughout this paper, 'selection' is used synonymously with homophilous social selection unless otherwise specified.

Although selection and influence are two core processes considered in ABMs, reality is naturally more complicated. Beyond the effects, we additionally consider that there may be latent causes reinforcing (dis)agreement. Shared information environments such as ideologically aligned media causing agreement, or other media which intentionally directly oppose one another, unobserved relations outside of the empirically observed friendship including classroom discussions, or the co-occurring change in the university environment are some examples that might induce further agreement or disagreement between pairs of individuals (see Chapter 2 of this dissertation). We thus account separately for these latent tendencies towards increasing agreement and disagreement.

Highly stylized, theoretical models for opinion dynamics are appealing for their parsimony and elegance, but may be subject to critique on external validity, and thus, in turn, on their empirical applicability (Steglich & Snijders, 2022). Many have pondered the question of how certain behavioral rules for sets of autonomous but interdependent agents might generate polarization,

a counter-argument which prevented this change in attitude, this would still be considered social influence by their social contact.

using formal models and computational experiments (for reviews, see e.g. Castellano, Fortunato, & Loreto, 2009; Flache et al., 2017). As this literature developed, the community documented core threats to the validity of such models, arising from challenges in the calibration and validation of opinion ABMs (Chattoe-Brown, 2014; Flache et al., 2017; Mäs, 2019; Sobkowicz, 2009). Two broad streams of calibration and validation of ABMs against real-world quantitative data have been categorized by Chattoe-Brown (2014). Firstly, some use experimental designs with real-life subjects to calibrate models on their micro-level foundations, i.e. they start from a set of theorized mechanisms, test for their presence in an experimental setting, and on the basis of the evidence derived therein, incorporate mechanisms in the conceptual model (e.g. Takács, Flache, & Mäs, 2016). Secondly, observational data can be used in comparison to a simulation, to examine whether a model gives a plausible explanation of a societal phenomenon, i.e. validation (Lorenz, 2021). Typically, this approach is used to compare which of a selection of models provides the most accurate prediction and thus the most plausible to have generated a given distribution or outcome. While this may be a useful approach when discriminating between trusted, plausible models, this approach can, strictly speaking, only establish generative sufficiency of an explanation (Epstein, 1999; Steglich & Snijders, 2022). While the explanation posited by an ABM for a phenomenon at the societal level lies in the interaction between agents, the empirical test rests on the outcome and not on the process that led to this outcome. Both of these two methods are important for linking and testing ABMs against reality. The current study falls primarily into the second category, as will be explained below.

Recent developments in statistical modelling have offered a new method for simulating the process of selection and influence in small social systems as a way to understand the micro-macro link (Snijders & Steglich, 2015; Stadtfeld, Takács, & Vörös, 2020; Steglich & Snijders, 2022). The stochastic actor-oriented model (SAOM; Snijders, van de Bunt, & Steglich, 2010) allows the estimation of the strength of micro-level mechanisms to explain the coevolution of relationships and individual characteristics in observational data of moderately-sized social systems. It is used for statistical inference on mechanisms for network formation and influence, simulating paths between multiple observations of individual traits and ties. To achieve that, it models the time between two observations as a series of micro steps in which autonomous but interdependent agents form, dissolve or maintain ties and adjust their mutable traits to their social surroundings. In its essence, SAOM is agent-based and can therefore form a good starting point for model-based computational experimentation, alongside other presented benefits which may be welcomed by users of ABMs.

In this paper, we use a specification of the stochastic actor-oriented model presented in Chapter 2 of this dissertation to examine the consequences of estimated parameters representing social selection, influence, and external ideologically polarizing factors on a community of university students, under the assumption that they remained together for a period of two years following the final observation at the end of their first year. The model applied considers the co-evolution of both social ties and opinions on multiple issues, thus also allowing for a dynamic social network starting from an empirically observed structure.

In the early days of polarization research, Abelson almost prophetically asked "what on earth one must assume" in order to arrive at polarization (1964, p.15). The vast literature on opinion dynamics tends to be better at explaining consensus formation than polarization. Here, we focus on the unintended consequences of individual, utility-maximizing processes. We show how tendencies towards (dis)agreement-maximization (latent-cause convergence and divergence) can affect relational polarization, and how ideological polarization is affected by homophilous selection and assimilative influence. Under the baseline model, the parameters representing social selection and influence create structures considered signs of relational polarization. Similarly, parameters representing latent (dis)agreement maximizing forces create structures considered signs of ideological polarization. We, therefore, focus on outcomes not directly generated by the structural effects modelled: ideological polarization as a consequence of positive selection and influence, and on outcomes of relational polarization as a consequence of latent polarizing forces.

AIMS AND CONTRIBUTIONS

The aims of this study are twofold. Firstly, to demonstrate and explore the use of a new (type of) method to link agent-based models for opinion dynamics to reality, and secondly, to explore the consequences of selection and influence processes versus latent ideological factors in polarization.

In this work, we aim to deepen our understanding of how selection and influence processes, as well as latent processes maximizing (dis)agreement, individually and jointly contribute to polarization. To do so, we analyze a formal, theoretical model of opinions and friendships, rooted in a model estimated from an empirical network. The approach taken on here is a comparative one, deploying three relatively novel and understudied methods for simulating empirically calibrated agent-based models via Stochastic Actor-Oriented modelling. We apply simplified versions of metrics proposed in Chapter 2 of this dissertation to test for polarization in both a relational and ideological sense.

The study stands out in its ability to disentangle relational and ideological polarization. The former sense of polarization may be a downstream consequence of a generalized affective polarization (Iyengar et al., 2019), which captures the idea that people come to dislike those from the other side (Iyengar et al., 2019; Reiljan, 2020; Wagner, 2021). The latter ideological form relates to constraint or consistency in multiple opinions, as shown in e.g. Baldassarri and Gelman (2008), Converse (1964), DellaPosta (2020), DellaPosta, Shi, and Macy (2015), and Kozlowski and Murphy (2021). While both forms are documented, there is, to the best of our knowledge, no calibrated ABM study that aims to disentangle the two like the SAOM-as-ABM approach used here.

AGENT-BASED MODELS AND THE STOCHASTIC ACTOR-ORIENTED MODEL

Common agent-based model limitations

While conventional agent-based models thrive on their flexibility, limitations in their application can hinder their ability to make effective predictions or explanations of macro-level phenomena. Here, we highlight four key issues which the SAOM helps to resolve via its use of empirical data that directly build validation of a model into the estimation procedure. The first of these limitations is that in ABMs it is often unclear what the operationalization of time and timescales represents (Mäs, 2019; Sobkowicz, 2009). Typically, the length of a simulation run is prespecified, or it may be terminated after reaching some (stochastic) equilibrium, but whether these end-points or equilibria may be reachable in real life is unclear. In opinion dynamics models, for instance, simulation steps might (when specified) reflect interaction events (Baldassarri & Bearman, 2007; Carley, 1991), but how many change-of-mind interaction events will individuals actually experience on a day-to-day basis? ABM does not offer a clear way for determining how and how often changes are expected to occur in real-life situations.

Second, much of the agent-based modelling literature is concerned with theoretical models without empirical calibration or validation of assumptions, mechanisms, and results. Empirical approaches have hitherto often been neglected in the ABM research paradigm. Increasingly, the community is calling for calibration and validation (e.g. Flache et al., 2017; Sobkowicz, 2009) and contributions are appearing that empirically calibrate micro-level behavior (e.g. Takács, Flache, & Mäs, 2016), calibrate to macro-level outcomes (e.g. Gestefeld et al., 2021; Vu et al., 2019) or validate to macro-level outcomes (e.g. Chattoe-Brown, 2014; Duggins, 2017).

Third, many agent-based modelling procedures treat social ties as the fixed topology of individuals who have some opportunity to influence one another. Here, the obscurity observed in the former critique returns, as ties can represent lasting, stable relationships or (opportunities for) interactions in the short term. While behaviourally impactful variations of stylized topologies (Rolfe, 2014) have been applied—including regular lattices (Axelrod, 1997), small-world networks (Watts & Strogatz, 1998), scale-free networks (Giardini, Vilone, et al., 2021), and connected caveman graphs (Flache & Macy, 2011) to name a few—dynamic *and* endogenous networks, in particular, are usually not considered in opinion models (although random dynamic networks and networks dynamically determined by opinions are not unheard of, cf. Albi, Pareschi, & Zanella, 2016; Baldassarri & Bearman, 2007; Durrett et al., 2012; Su et al., 2014) despite empirical research explaining structural, dynamic tendencies (Minozzi et al., 2020).

Fourth and finally, ties are often given without a theoretical explanation of what they represent. These might be taken to represent, for example, communication opportunities, interaction events, or perceived (mutual) social relationships like acquaintanceship, friendship, kinship or antagonism. Ambiguity about the identity of the modelled relationship makes it harder to pin down theoretical expectations specific to it (Butts, 2009)[2].

Strengths of the stochastic actor-oriented model as an agent-based model

Stochastic actor-oriented models are now an established tool for studying the dynamics of social networks over time (Snijders, van de Bunt, & Steglich, 2010), including and accounting for effects of and on individual attributes (Steglich, Snijders, & Pearson, 2010). These models, estimated on whole-network panel data, make the agent-centric assumption that individuals change their outgoing ties to other individuals or objects depending on their current state, which may include network and individual variables specified by the modeller.

Using stochastic actor-oriented models as tools for agent-based modelling (i.e. as a tool for 'generative social science') is a possible way to resolve or avoid some of the aforementioned issues in agent-based models (Steglich & Snijders, 2022).

2 It is worth clarifying our stance that work restricted to empirical methodologies does not necessarily fare better: measurement is hard, and defining the scope of the possible relevant causes of an outcome is similarly tricky.

SAOMs have a clear interpretation of time, are estimated directly from empirical data, use dynamic networks by definition, and due to their use of empirical data, require explicit operationalization of the ties in question. The SAOM describes a process, with variables specified by the researcher, by which an observation of a social network occurring on a set of individuals at some time point evolves to an observation at a second time point. While this is in principle a statistical model based on empirical data, the SAOM estimates can be used as parameters of an agent-based model, with certain mechanisms for selection of ties and influence on those ties specified by the researcher.

First, the SAOM gives a sense of time. It uses panel-based data and estimates the rate at which individuals take opportunities to update their social ties and attitudes. Given that this rate is based on the time window between observations, the real-time between observations can be mapped to this rate. To do so requires the assumption that the rate of change estimated in the observation period is constant over time, and that the actor-level mechanisms of change are of constant sign and size.

Second, a core aspect of the SAOM approach is that it estimates theoretically specified effects against panel empirical data. This means that the user calibrates their model against the development of observed data at timepoint t to $t+1$ in a way that requires that it explains aspects of the structure of the network (Ripley et al., 2019), and is validated against user-specified goodness-of-fit statistics; comparisons between observed and model-generated statistics of interest not directly included in the model (Lospinoso & Snijders, 2019). Calibration is thus built into the estimation procedure, while the user validates the model by goodness-of-fit testing.

Third, the SAOM was specifically designed to estimate plausible dynamic network models for social networks. For instance, friendship networks may be subject to forces causing transitivity (the friend of a friend is more likely to be a friend), reciprocity (if individual i nominates j as a friend, j is more likely to nominate i as a friend), popularity-like effects (individuals considered a friend by many are more likely to be nominated as a friend by others) and homophilous selection on attributes such as political attitudes (similar individuals are more likely to be friends) (Snijders, van de Bunt, & Steglich,

2010) as well as less intuitive tendencies against e.g. transitive reciprocal triads (Block, 2015). Such mechanisms for tie selection are specified by structural tendencies for which parameters are optimized to produce, on average, the same structural statistics as observed when simulating from one observation of the network to the next. These can be treated at the same time as features representing changes in individuals' attitudes such as influence. Dynamic network effects are thus a fundamental part of the SAOM's use.

Finally, the constraints of empirical data forming the input require that the researcher is explicit about the (operationalization of) ties forming the social network, requiring a more concrete specification of the consequences of a tie based on previous research, and thereby allows for better falsification of hypotheses in the empirical setting.

Challenges in using the stochastic actor-oriented model as an agent-based model

While generative SAOMs are empirically calibrated and therefore quantitatively close to observed data, a SAOM-based strategy does not guarantee the realism of the social simulation model by default. It is, in the current case, still an assumption-rich artificial model of selection and influence. For instance, assuming that a single type of social relation—even when comprehensively measured—is the sole explanation for a change in opinions is an obvious oversimplification. There are likely multiple qualities co-occurring or part of a social tie that vary between dyads, or there may be other tie types that are of similar importance. Furthermore, there may be unobserved, exogenous factors that affect the empirical data and bias the statistical estimates.

In sum, the SAOM offers opportunities to improve the realism of opinion models via its empirical constraints, particularly by emphasizing the importance of specifying the objects to be studied.

Approach

We explore the consequences of strengthened, weakened, or absent forces of homophilous selection, assimilative social influence, and latent causes of divergence and convergence of opinions on polarization in a relational and ideological sense. We take three approaches to the use of a stochastic actor-oriented model as a counterfactual tool, stemming from debates in the literature on SAOMs: After adjusting the parameters of theoretical interest, simulating forwards directly, re-estimating only the density prior to simulation, or re-estimating the entire model outside of the parameters of interest. Each method comes with its own assumptions, to be discussed below.

APPROACHES TO CALIBRATION VIA SAOM

Two different methods for inferring the consequences of a process have been more or less explicitly proposed (Snijders & Steglich, 2015; Stadtfeld, Takács, & Vörös, 2020). Snijders and Steglich (2015) propose the re-estimation of a model with specific parameters of interest removed, and examining some outcome which is not explicitly implied by the remaining parameters. This means that other parameters will be able to change. Taking this approach follows a logic of network modelling as a counterfactual, asking the question 'if these were the structural tendencies by which a network developed over time, what is the strength of each tendency?' For some models, re-estimation with or without certain parameters makes model convergence (i.e. stable estimation of the magnitude and precision of the parameters) less likely – the plausibility of the estimated model is reduced in such a scenario. That aside, if the goodness of fit on relevant structural features and convergence on parameters is equal across models, the difference in realism comes from theory. In this sense, we can answer the question of whether we believe that a specific structural tendency is a plausible, or necessary, component of the generative process of a given network with respect to some set of outcomes.

Another method is not to re-estimate the parameters of the model, but to fix the non-focal parameters to their estimated values, and manipulate only the focal parameters (Stadtfeld, Takács, & Vörös, 2020). This is an alternative counterfactual method with stronger causal assumptions on the estimated parameters, since it assumes that we have observed 'true' representations of causal tendencies, which we then manipulate following our theoretical interests. In this case, we are not interested in goodness-of-fit, as we are concerned with extrapolation – we do not aim to answer whether a specific force is a cause of a network observation, but what would happen if this force were different from the observation (i.e. stronger, or weaker). This has consequences for network-level statistics including the density; we have to make the assumption that the number of relationships in the system is not constrained by an approximate limit due to e.g. broadly construed costs and benefits of ties. Instead, it is constrained purely by the opportunities that the network's characteristics provide to allow individuals to satisfy their preferences.

A midway alternative between the two approaches mentioned has been suggested by Block (2018), which could represent this latter belief that ties tend toward some limit in the interval in which the model is estimated (as would also be suggested by e.g. Hill & Dunbar, 2003, albeit at the level of the network rather than the level of the individual). In this approach, the intercept of the objective function (i.e. the density parameter) is re-estimated, while the focal parameters are manipulated and the other model parameters are held constant. This aims to approximately preserve the number of ties in the network, accounting for changes in the focal structural features that may increase the average utility of ties, and assuming other parameters were initially correct.

We opt to simulate in all three ways; firstly, using the method proposed by Snijders and Steglich (2015), simulating from a model re-estimated after manipulation of the focal parameters, next using the method used by Stadtfeld, Takács, and Vörös (2020), simulating from the model with only the focal parameters changed, and finally simulating from a model with a re-estimated density parameter after manipulation of the focal parameters. This allows the addressing of the consequences under all three sets of assumptions.

METHODS

To understand the impact of (various levels of) selection and influence on polarization we simulate counterfactual scenarios using a stochastic actor-oriented model. At its core, the model is an agent-based model starting from an empirically observed configuration at timepoint t, calibrated to best predict the statistics of the observed network at timepoint $t + 1$. Here, we describe first a unique treatment of the attitudinal data as a two-mode network, followed by the data upon which the model is estimated, the mechanics of the agent-based model, then the stochastic actor-oriented model from which it inherits, the manipulations that we wish to study with computational experiments, and the measures by which we quantify polarization in the system.

People, issues and attitudes as two-mode networks

The relationship between individuals and the issues about which they hold attitudes is treated as a type of network known as a two-mode network. Two-mode, or multilevel networks are networks consisting of two sets of nodes, representing entities such as or individuals and subjects of an attitude (Raabe, Boda, & Stadtfeld, 2019), individuals and social events which they attend (A. Davis, Gardner, & Gardner, 1941), or organizations and their competences (Hollway et al., 2017). These sets of nodes can be interconnected between the sets, but the same tie type does not occur within sets (Lazega & Snijders, 2015). This stands in contrast to the more common one-mode social network of ties—in this paper, the network of friendships. These can be combined, similarly to cognitive balance theory (e.g. Heider, 1946) which formulated networks of pairs of individuals and their attitudes towards other objects. In the example in Figure 3.1 attitudes about policy issues are depicted as a two-mode network. Positive and negative attitudes are treated as green or red ties respectively, from individuals holding them (circles) to the policy issues at hand (squares).

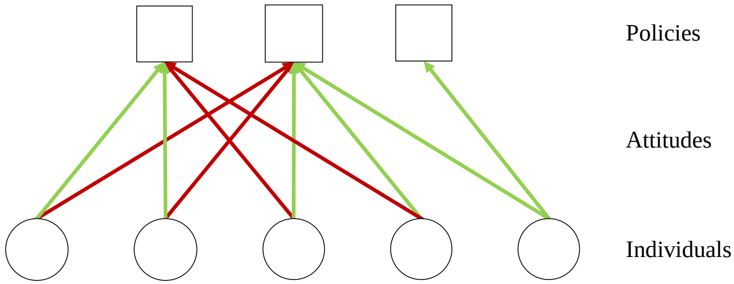


FIGURE 3.1: An illustrative multilevel network of individuals and attitudes. Squares represent policies, red and green ties represent positive and negative attitudes that circle individuals hold about the policies.

The smallest closed unit of this two-mode network is a *four-cycle*; a combination of two individuals and two issues, connect by individuals' attitudes on these issues. Combining this two-mode network with one-mode friendship networks, the smallest closed structure is a *triad* of two individuals and one issue, with friendship connecting the individuals to one another, and attitudes connecting the individuals to the issue. The presence or absence of these closed structures is used in the modelling treatment of selection, influence, and latent causes of opinion divergence and convergence, and the development of the polarization metric described later.

Data

The data on which we estimated the model came from student Cohort 2 of the Swiss Student Life data (Vörös et al., 2021). The data used related to 261 individuals, from approximately two months into the start of their studies until the end of their first year (i.e. omitting the first observation due to the instability of the social network) with information about attitudes on 22 policy items and the friendships of these individuals at four observation points.

The policy topics on which the cohort indicated attitudes came from a battery of items stemming from two sources. First, the 2017 German language Wahl-O-Mat (Bundeszentrale fur politische Bildung, 2017), a vote choice helper, from which we hand-picked policies which we believed would be relevant for the cohort and therefore more likely to be discussed and/or change over time. Second, we used information from upcoming Swiss referenda (in which citizens vote up to four times per year), as these are likely to come up for discussion.

Friendship ties were chosen due to their potential impact on opinions for reasons of trust over more incidental ties (Weeks & Gil de Zúñiga, 2019), a more general tendency for positive social ties to be related to social influence (Lazer et al., 2010; Levitan & Visser, 2009) their potential to be selected based on shared opinions or values (Huston & Levinger, 1978; McPherson, Smith-Lovin, & Cook, 2001), and finally because friendship ties have extensive anchoring in the social networks literature, giving us more certainty in the auxiliary part of the model specifications (e.g. Snijders & Steglich, 2015; Snijders, van de Bunt, & Steglich, 2010).

To allow for estimation of the multivariate opinion part of the model, opinions were reduced from a seven-point Likert scale to three ordered categories, represented as negative, neutral, or a positive attitude respectively. In the variant presented here, we use this reduced form with cutoffs two points away from the extremes (i.e. a score of one or two is treated as negative, three, four or five as neutral, and six or seven as positive).

We use these data as they contain several convenient features: a network with a relatively clear boundary, individuals being students who are enrolled in intensive study programmes and, therefore, spending much time together in the study context. The items used were expected to be among those most likely to be subject to change of opinion, given that they focused on contemporary topics and/or upcoming topics in the political arena.

The stochastic actor-oriented model as an agent-based model

The agent-based model applied here is a continuous-time sequential-updating model. At any point in time, a single actor is sampled at random to update their state of social relations or their attitude to political topics. For a given observation period, rate parameters determine the average frequency with which actors have an opportunity to update their state. The rates are separate for each of the social and the opinion networks, but tie changes in these networks are assumed to occur in continuous time, thus interspersed.

When activated, agents decide to update their relationships with other actors (or topics). Their relationships and political attitudes are represented as three networks: a one-mode friendship network and two disjoint two-mode actor-topic networks (one network for positive and one for negative attitudes). The latter two networks are disjoint, meaning no one can hold both a negative attitude and a positive attitude towards the same object. The choice of (in)action in changing ties in a network follows from a probabilistic function dependent on counts of specified local network structures weighted by the estimated parameters, and assumes that actors could potentially be tied to any other nodes in the network.

In the model specified here, there is a set of general parameters by which all individuals in the opinion network are expected to have opinions related to one another (with dissimilar ties becoming more probable amongst already-dissimilar individuals, and vice versa for similar individuals). Specific parameters for the social network capture the process by which people selectively tie to others with a higher probability given more overlap in opinions (selection), and other parameters specify the tendency to adopt an opinion the same as one's friend (influence).

Baseline stochastic actor-oriented model estimation

We start from the model described in Chapter 2 of this dissertation. In this research to which we made two modifications. First, we excluded parameters removing non-significant hypothesized parameters for negative selection on opposed opinions and assimilative social influence on opposed opinions[3]. Second, we removed one homophilous control (age similarity) in the social network as this did not explain tie formation in the empirical model, and third, following Snijders and Steglich (2015), opt for a geometrically-weighted edgewise shared partners term alongside a parameter for complete triads to avoid problems of degeneracy seen in other models such as exponential random graph models, which run until a stochastic equilibrium is reached (Robins et al., 2007). This re-estimated model maintained a good fit to the data in aggregate but was somewhat poor on a wave-by-wave basis[4]. To establish the time frame, we extrapolate from changes between the last two observations of the community. These observations were spaced 71 days apart. Multiplying the estimated rate parameter between these observations by 11, we get an estimated 781 days, or 2.13 years. For simplicity, we treat this as a two-year period.

Baseline model parameterization

The SAOM relies on two main functions in each network for simulation: an objective function and a rate function. The objective function determines the relative probability of specific tie changes by a chosen actor. In other words, it gives the probability of a specific change (but also the probability of not

3 Since each of these is represented by two parameters (one for each triadic structure of the directed social network and either positive or negative opinion ties), we based this selection on a test of the summed parameters for significance using Wald tests (Ripley et al., 2019). This is furthermore consistent with the weak-to-null effects of negative influence in offline contacts found in e.g. Takács, Flache, and Mäs (2016), though constrasting with online effects in larger samples such as Bail et al. (2018).

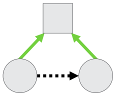
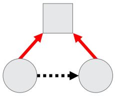
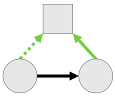
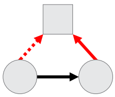
4 Due to this partial poor fit, we also applied the original and the adapted model to waves four and five of the Swiss StudentLife Study dataset (Vörös et al., 2021) used in Chapter 2 of this dissertation. While the original model fit well, the removal of the additional parameters did not. Given the relatively small amount of change (421 tie changes in the positive, and 251 changes in the negative network, with 8 parameters controlling each) and extremely good fit, we suspect that this model was overfitting the data. Manipulating these parameters and re-estimating all the models resulted in several non-converged models, but otherwise frequently fit the data well. Removing the parameters for negative influence and negative selection as well as the age homophily parameter resulted in a model that failed to come close to convergence after two runs with default settings except 5000 phase 3 simulations.

changing) when an actor is activated. Local network structure determines the values put into the objective function; each possible change is evaluated according to the modelled statistics such as the number of agreeing four-cycles, agreeing friendship-attitude-attitude triads or reciprocal friendships. The rate function determines how often actors are activated and thus how many opportunities individuals have (on average) to update their outgoing ties. Rates of change between different networks are independent in the given model. The changes based on the objective function, however, are not, since they rely in part on cross-network effects between the two-mode opinion network and the one-mode friendship network.

The estimated parameters of selection and influence, as well as the background latent processes inducing (dis)agreement from Chapter 2 of this dissertation and a visual representation of each, are given below in Table 3.1 and Table 3.2[5]. Both tables 3.1 and 3.2 show diagrams containing individuals (circles), topics (squares), and social (coloured black) or attitudinal (coloured green for positive and red for negative) relationships. Positive parameters indicate a relative increase in the probability of ties being formed or preserved given that they are part of the displayed structure, negative parameters the opposite. Table 3.1 shows the estimates for selection and influence between individuals. Two parameters represent social selection, for which a positive parameter can be interpreted as homophily (and a negative parameter thus as heterophily). Two parameters represent social influence, where a positive parameter represents an assimilative tendency (and a negative parameter would represent negative influence or distancing). Finally, Table 3.2 shows the estimates for eight parameters that represent tendencies towards convergence and divergence of opinions. These account for latent influences, be it from ties uncorrelated with friendship within the network, or external influences such as a pre-existing shared media consumption (see Chapter 2 of this dissertation).

5 Full model tables are large, and are available as HTML files on request.

Table 3.1: Main Stochastic Actor Oriented Model parameter estimates: Selection and influence

Visual representation	Effect	Estimate
	Homophily from pos. att.	0.06^{*} (0.03)
	Homophily from neg. att.	$0.06^{\cdot}$ (0.03)
	Influence, pos. att.	$0.04^{\cdot}$ (0.02)
	Influence, neg. att.	0.07^{*} (0.03)

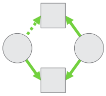
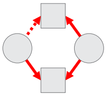
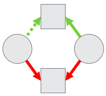
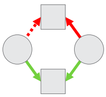
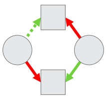
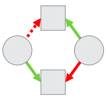
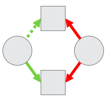
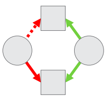
$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$. Dependent tie is dotted line. SEs in brackets below parameter estimates.

Manipulations

We use the four aforementioned groups of parameters, i.e. selection, influence, and the latent confounds affecting opinion convergence and divergence, to assess their unique effects when interacting with the rest of the model under the computational experimental conditions. We apply four types of manipulation: removing, halving, doubling and quintupling all parameters in a group. The manipulation is thus based on intuitively defined multipliers answering questions of what the consequences would be in a similar society where these forces were completely absent, substantially weakened, strengthened, or substantially strengthened. We apply these manipulations to each of the four groups of parameters and treat the other parameters in line with one of three approaches to the generative SAOM: holding constant all other parameters, re-estimating all other parameters, or re-estimating density parameters[6]. For the full compu-

6 In one model, multiplying the strength of latent convergent tendencies by 5 and re-estimating all other parameters, the model did not converge.

TABLE 3.2: Main Stochastic Actor Oriented Model parameter estimates: Convergence and divergence

Visual representation	Effect	Estimate
	Latent convergence type 1, pos. att.[b]	0.28^{***} (0.03)
	Latent convergence type 1, neg. att.[b]	0.49^{***} (0.09)
	Latent convergence type 2, pos. att.[b]	0.07 (0.05)
	Latent convergence type 2, neg. att.[b]	0.18^{***} (0.05)
	Latent divergence type 1, pos. att.[ab]	0.90^{***} (0.13)
	Latent divergence type 1, neg. att.[ab]	0.62^{***} (0.09)
	Latent divergence type 2, pos. att.[ab]	0.61^{***} (0.12)
	Latent divergence type 2, neg. att.[ab]	0.43^{**} (0.15)

$^{***}p < 0.001$, $^{**}p < 0.01$. Dependent tie is dotted line. SEs in brackets below parameter estimates. [b]Estimate and SE multiplied by 100.

tational experiment, this implies a 4x4x3 design. We simulate 900 runs for each condition plus the baseline unmanipulated model[7]. These are denoted by their adjusted parameter set (selection, influence, latent convergence, and latent divergence), their numerical multiplier, and their estimation method (manipulation only, density re-estimated, or all re-estimated).

Polarization

We capture polarization following Chapter 2 of this dissertation. In that work, a two-dimensional approach to polarization was taken, and the following description is modestly adapted from it. Firstly, ideological polarization was measured by the extent to which pairs of individuals consistently agreed or disagreed in their attitudes, in the sense of 'constraint' or 'alignment' (Baldassarri & Gelman, 2008; Converse, 1964; Kozlowski & Murphy, 2021). This may be best related to common understandings of polarization as an ideological phenomenon, such as in DiMaggio, Evans, and Bryson (1996), with the exception that we focus on a multidimensional understanding of polarization, as is increasingly opted for by scholars in the field and in line with one of most likely occurring forms of ideological polarization (Dinkelberg et al., 2021; Kozlowski & Murphy, 2021; Schweighofer, Schweitzer, & Garcia, 2020). Secondly, relational polarization was measured by the extent to which social ties were present between agreeing, but not disagreeing, individuals, i.e. clustering of balanced triads of people and attitudes (Cartwright & Harary, 1956; Heider, 1946). This relates most closely to affective polarization, potentially being a downstream consequence of it (Iyengar et al., 2019), and thereby bears some conceptual similarity to social-structural polarization (Baldassarri & Bearman, 2007) previously shown in online settings (Conover et al., 2011).

The metrics applied use relative frequencies of selected structural configurations in the network: four-cycles of two people and two issues connected by attitudes and triads of two people and an issue connected by friendship and attitudes. The four-cycle structures in the network relate to the degree to which

7 due to space constraints on the high-performance computing cluster this is reduced from the originally intended 1000; a number selected intuitively as a reasonable bound for robust estimation.

the network is ideologically polarized, and the triad structures to the degree to which it is relationally polarized.

Ideological polarization. Proportions of the ordered sets of complete four-cycles define ideological polarization. The count of four-cycles where a pair of individuals holds consistent attitudes towards a pair of issues, i.e. agreeing or disagreeing on either set of issues, is taken as the numerator. These two structures are indicated by $\langle A \rangle$ and $\langle D \rangle$, respectively. The denominator is the count of these structures, and additionally, structures in which a pair of individuals do not hold consistent attitudes towards pairs of issues, i.e. where the individuals agree on one issue but not on another. These structures are represented by $\langle I \rangle$

The first component of ideological polarization is *ideological attraction*; defined as the proportion of cases where two people consistently agree on two issues rather than having inconsistencies in their attitudes (measured by inconsistent four-cycles $\langle I \rangle$):

$$\text{Ideological attraction} = \frac{\#\langle A \rangle}{\#\langle A \rangle + \#\langle I \rangle} \tag{3.1}$$

The second component of ideological polarization is *ideological repulsion*. This is defined as the counterpart to ideological attraction, being the proportion of cases where two people have consistently disagree on two issues rather than having inconsistent attitudes on those issues:

$$\text{Ideological repulsion} = \frac{\#\langle D \rangle}{\#\langle D \rangle + \#\langle I \rangle} \tag{3.2}$$

The two components can be interpreted as analogous to the probability of two individuals agreeing on one issue if they agree (disagree) on another, given that they hold non-neutral attitudes on both. The mean of the two component metrics measures the level of ideological polarization: the probability of observing consistently agreeing or disagreeing attitudes relative to inconsistent ones in pairs of individuals and issues. Formally this is:

$$\text{Ideological polarization} = \frac{\text{ideological repulsion} + \text{ideological attraction}}{2} \tag{3.3}$$

It is noteworthy at this point that the probability of two individuals agreeing is affected by various factors, which need not necessarily correspond to polarization-specific explanations, such as the general prevalence of positive and negative attitudes (the attitude network densities) and the different tendencies of items to attract positive or negative attitudes (the item degrees).

The procedure for generating an expectation on these statistics given the distribution of opinions and social ties in the network presented in Chapter 2 of this dissertation is computationally intensive, and not amenable to the volume of data generated in terms of the numbers of simulations and over time. For this reason, we opt not to normalize this statistic.

Relational polarization. The metric of relational polarization is defined by two components that relate to homophilous attraction and heterophilous repulsion. Both are based on ordered triad counts: each triad is counted twice, once considering the potential friendship from individual *i* to individual *j* and once the other way round, as in empirical data these are not guaranteed to co-occur. We refer to these structures by the following symbols. The structure $\triangle_{A}$ represents agreement among friends: they both support or oppose an issue. $\wedge_{A}$ represents agreement in the absence of friendship. Together, these can be used to represent the extent to which people tend to be friends depending on their political agreement. Similarly, $\triangle_{D}$ stands for disagreement between friends, and $\wedge_{D}$ for disagreement in the absence of friendship. Counting these two structures in the network can help to understand the extent to which people's friendships are structured in a way that avoids encountering disagreement.

The first component of relational polarization is *relational attraction*. An appropriate statistic is calculated as the proportion of ordered triads in which two individuals who agree on an issue are friends, out of all agreeing pairs of individuals and an issue:

$$\text{Relational attraction} = \frac{\#\triangle_{A}}{\#\triangle_{A} + \#\wedge_{A}} \tag{3.4}$$

The second component is *relational repulsion*. This is defined as the ratio of ordered triads in which two individuals who *dis*agree on an issue are *not* friends, out of all disagreeing pairs of individuals and an issue:

$$\text{Relational repulsion} = \frac{\#\wedge_{\mathrm{D}}}{\#\wedge_{\mathrm{D}} + \#\triangle_{\mathrm{D}}} \tag{3.5}$$

The two components can be interpreted as the observed probabilities that a) i considers j a friend given they agree on an issue and that b) i does not consider j a friend given they disagree on an issue. The expected value of these probabilities will be affected by the prevalence of friendship relations in the sample (for example, equation 2.3 has an expectation equal to the friendship network density if there is no relational attraction). For this reason, we normalize by the density of the friendship network, giving an expectation of 0.5 under randomly distributed friendship ties. To do this, we multiply all triads included in the calculation of a statistic by one minus the probability of the social tie if the social tie is present, or by the probability of a social tie if it is not present. With social tie probability p, this is formally:

$$\text{Relational attraction} = \frac{\#\triangle_{\mathrm{A}}(1-p)}{\#\triangle_{\mathrm{A}}(1-p) + \#\wedge_{\mathrm{A}}\,p} \tag{3.6}$$

and for relational repulsion:

$$\text{Relational repulsion} = \frac{\#\wedge_{\mathrm{D}}\,p}{\#\wedge_{\mathrm{D}}\,p + \#\triangle_{\mathrm{D}}(1-p)} \tag{3.7}$$

and finally, we take the mean of these two measures to get the relational polarization statistic:

$$\text{Relational polarization} = \frac{\text{relational repulsion} + \text{relational attraction}}{2} \tag{3.8}$$

RESULTS

In the results, we examine the two-year period of forward simulations starting from the last observation in our model (i.e. at the end of the academic year).

Firstly, we take a look at a simple statistic: the density of each network. This contextualizes our interpretation of the metrics of interest (polarization), as one may judge polarization in a network with few ties quite differently to one with more ties. These give different opportunities to agree or disagree with connected others, and different opportunities for opinions to be more and less opposed. Following the examination of the density, we examine the metrics directly related to the outcomes of interest. Along the way, we also examine plots of the resulting networks to get an idea of whether these networks are intuitively plausible. Before examining the complete results of the computational experiment, however, we first show the behaviour of the model under the baseline specification.

Network density

In Figure 3.2, the results of the forward simulation from the baseline model for the density statistic are presented. Note the scale of the axes, as these are quite small and in later figures are rescaled to the results presented. Quantile intervals are presented rather than e.g. confidence intervals: under simulated experiments, mean differences can arbitrarily made significant by increasing the number of simulation runs. Quantiles instead present a more intuitive understanding of possible outcomes. Here it can be seen that the friendship network density drops slightly, before returning to approximately the baseline level again at the final observation. In addition, as would be expected under path-dependent stochastic processes, the interval of outcomes (in this case, the 50% interval) around the average grows over time. Opinions, however, drop substantially in frequency to near-zero values and remain there. We see from this that the given configuration of opinions is particularly unstable under the model. It is notable that the minimum of friendship ties in the simulation occurs just before opinions.

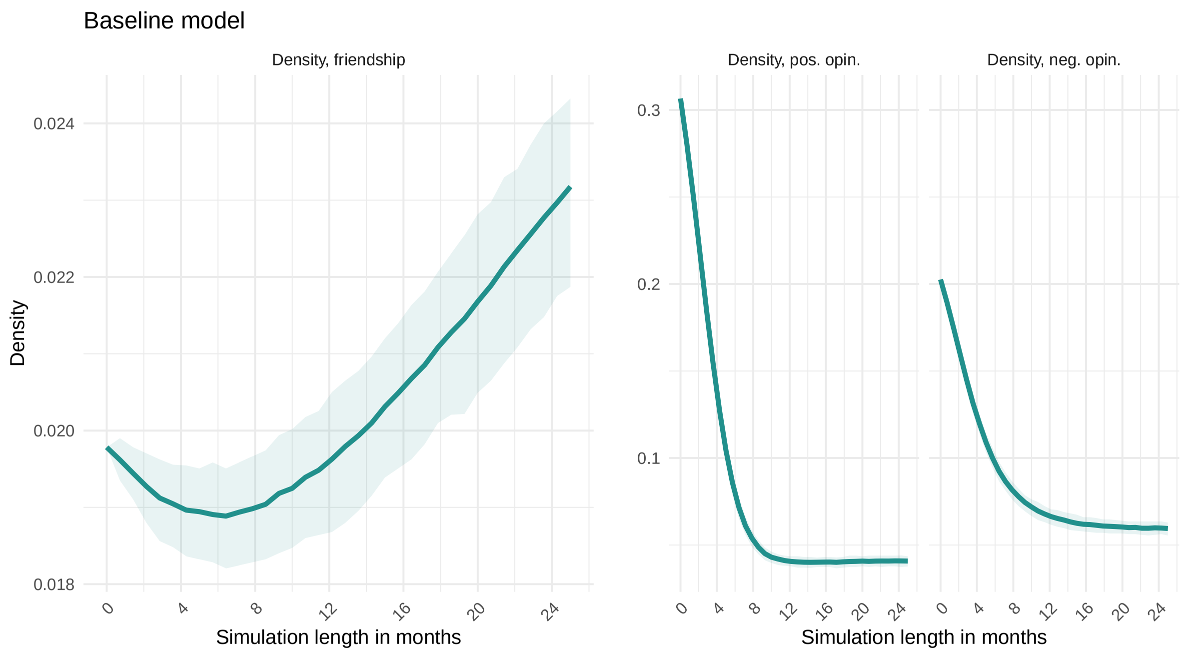


Figure 3.2: Densities of friendship, positive attitudes and negative attitudes over time in the baseline model alone. Ribbons indicate lower 25% to upper 75% quantiles.

In what follows, we show more complex plots, each one showing a different set of effects adjusted in the model. To allow easier comparison between manipulation types within each set of adjusted effects, we present manipulations of each set of parameters with the same y-axis scaling initially, followed by rescaled plots where appropriate. In each plot, multipliers are represented by colours, and the model estimation type is represented by different line types.

Moving to the results of the manipulations, we first examine influence, i.e. the tendency to adopt alters' opinions, presented in Figure 3.3. It is apparent here that synthetically increasing influence to five times its estimated value without re-estimating the model in any way causes increased density in both the opinion and the friendship networks. Zooming in on the regions where the rest of the values lay in Figure 3.4, even at this high resolution, does not show particularly separated densities under different conditions. Counter-intuitively, the re-estimation of density under a five times multiplier ends up with the lowest densities in all networks presented, whereas a zero times multiplier results in the highest in the two opinion networks.

Examining (homophilous) selection, i.e. the tendency to choose to connect to alters with more similar opinions, very similar effects can be seen in the opinion networks, but quite different ones in the social network, as seen in Figure 3.5 and with a truncated y-axis in Figure 3.6. In the opinion networks, again, a drop is seen. All conditions are quite tightly grouped. Looking at the friendship network densities, different results are apparent. Here, as the homophilous forces increase, the relative number of ties decreases. Beyond this, re-estimating models results in similar trends regardless of the method of re-estimation, with a reduction in selection resulting in more social ties, and an increase resulting in fewer ties over time. On the other hand, manipulating parameters without re-estimation results in the opposite effect, presumably owing to the fact that an isolated increase in a parameter representing an increased tendency to connect to others the more opinions they share, means that the average tie probability is increased. While the densities of the friendship networks without re-estimation largely follow the baseline and are ordered by their magnitude, a five times selection multiplier increases much more rapidly than others, then remaining roughly stable for the duration of the simulation.

Finally, the latent forces show the biggest changes in densities. With five times increased latent forces and without re-estimation, all networks rapidly increase in density, as shown in Figures 3.7 and 3.9. This particularly affects the opinion networks, which become completely saturated. Similar effects occur for the two times multiplier on latent convergence (Figure 3.7).

Where convergent latent factors are strengthened and (parts of) the model re-estimated, a rapid increase in frequencies of opinions are noticeable in the positive and negative networks. Interestingly, a larger multiplier typically leads to a relative decrease in final density in the opinion and friendship networks. A similar pattern is observed in the case of divergent latent factors; increased multipliers tend to mean lower overall densities at the end of simulation in the case of the opinion networks, though overall densities remain more similar across conditions. This inverse relationship between density and multiplier is not observed in friendship networks, and only under a five times multiplier does the trend diverge from the others in a faster increase.

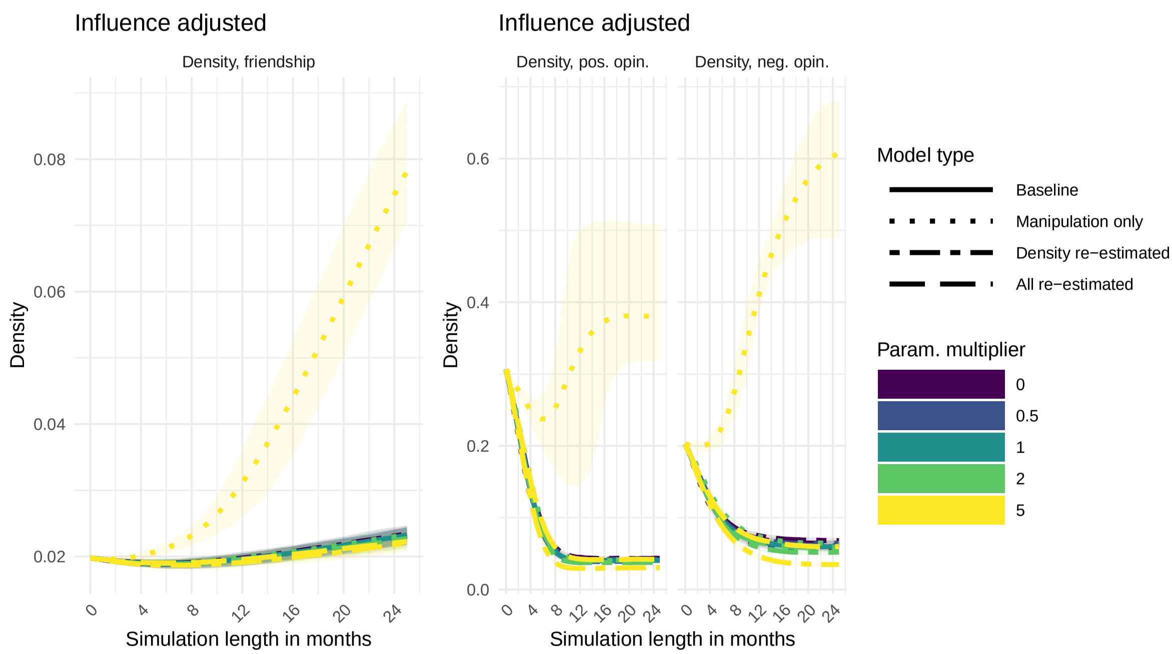


FIGURE 3.3: Densities of friendship, positive attitudes and negative attitudes over time when manipulating the influence parameters. Ribbons indicate lower 25% to upper 75% quantiles.

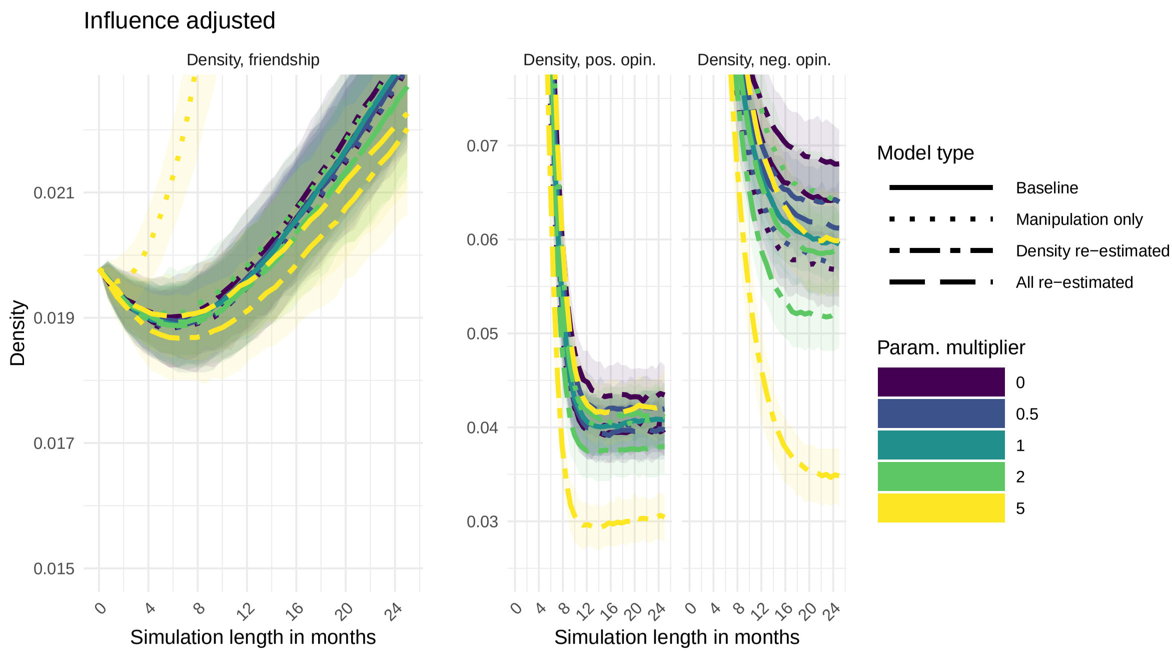


FIGURE 3.4: Truncated y-scale of densities of friendship, positive attitudes and negative attitudes over time when manipulating the influence parameters. Ribbons indicate lower 25% to upper 75% quantiles.

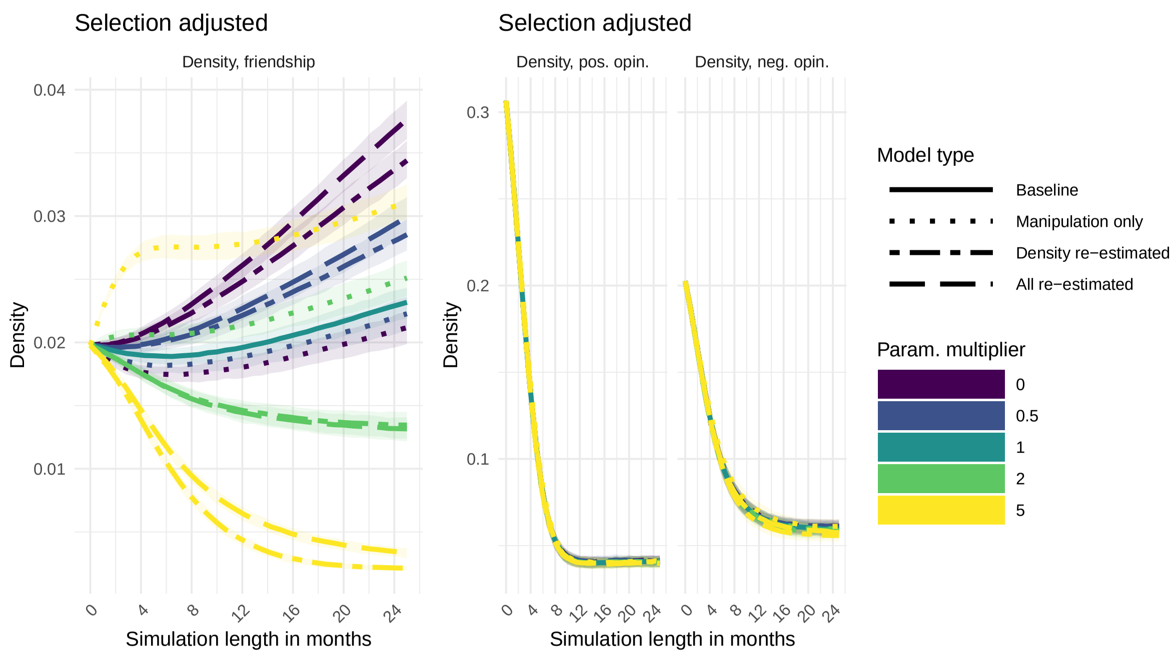


FIGURE 3.5: Densities of friendship, positive attitudes and negative attitudes over time when manipulating the selection parameters. Ribbons indicate lower 25% to upper 75% quantiles.

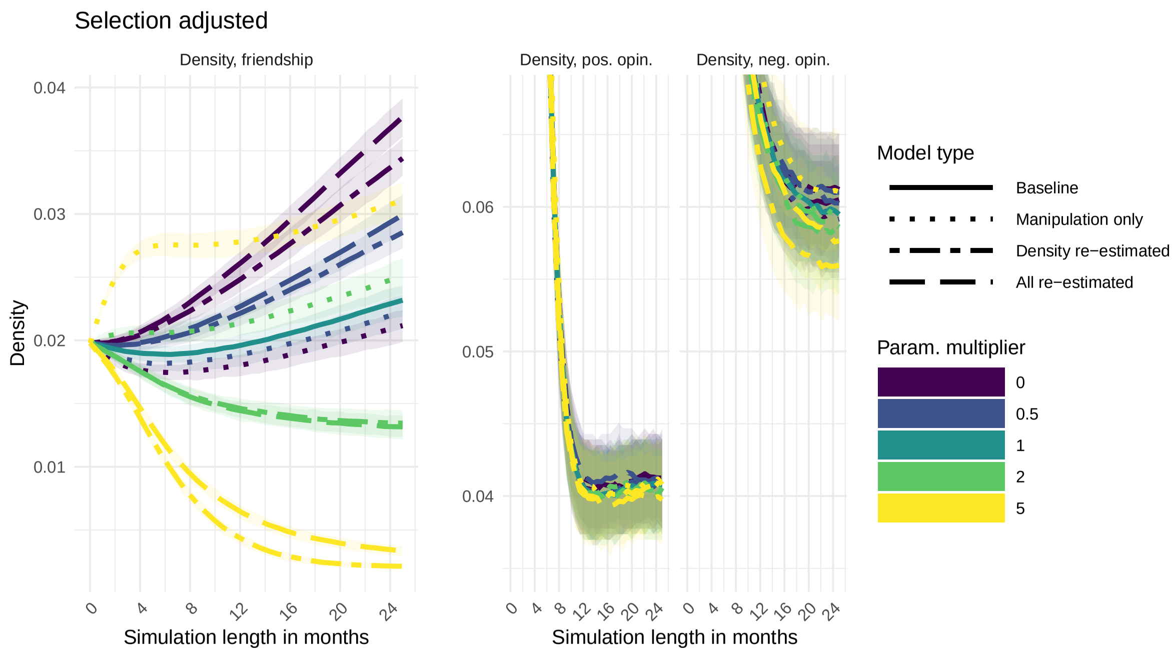


FIGURE 3.6: Truncated y-scale of densities of friendship, positive attitudes and negative attitudes over time when manipulating the selection parameters. Ribbons indicate lower 25% to upper 75% quantiles.

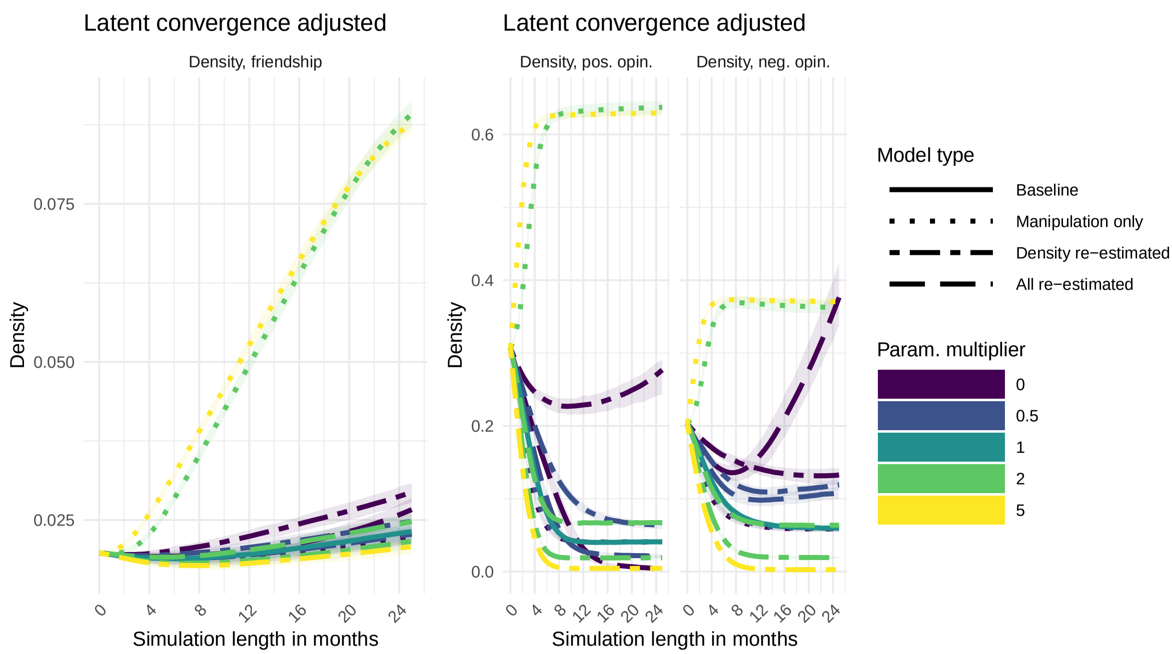


FIGURE 3.7: Densities of friendship, positive attitudes and negative attitudes over time when manipulating the latent convergence parameters. Ribbons indicate lower 25% to upper 75% quantiles.

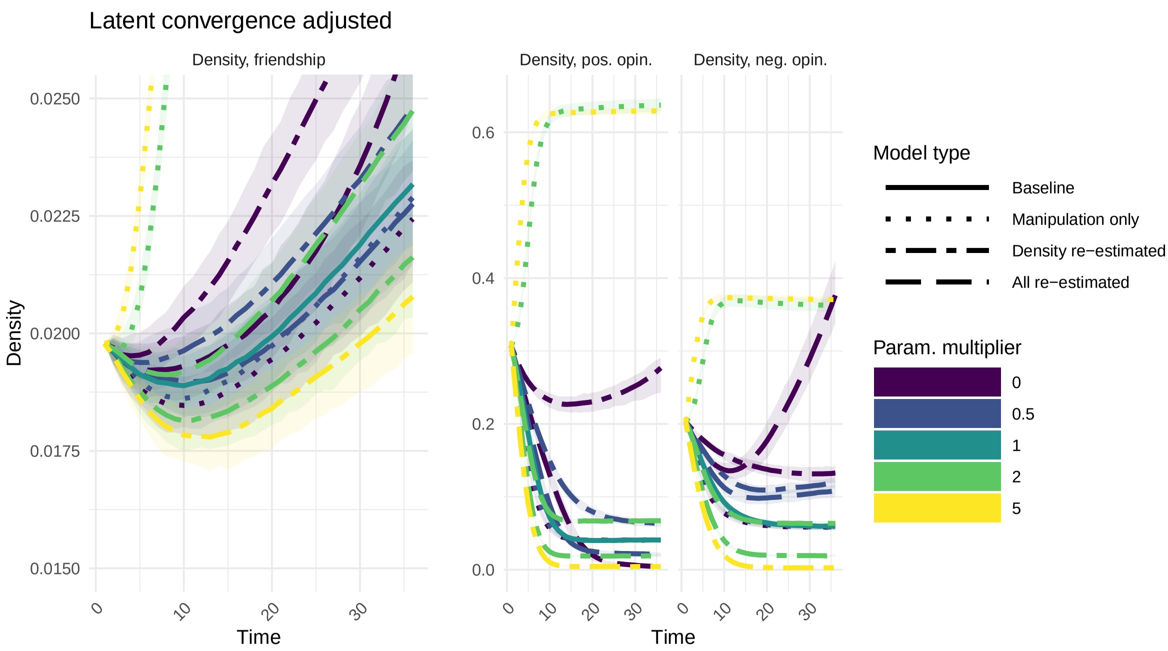


Figure 3.8: Truncated y-scale of densities of friendship, positive attitudes and negative attitudes over time when manipulating the selection parameters. Ribbons indicate lower 25% to upper 75% quantiles.

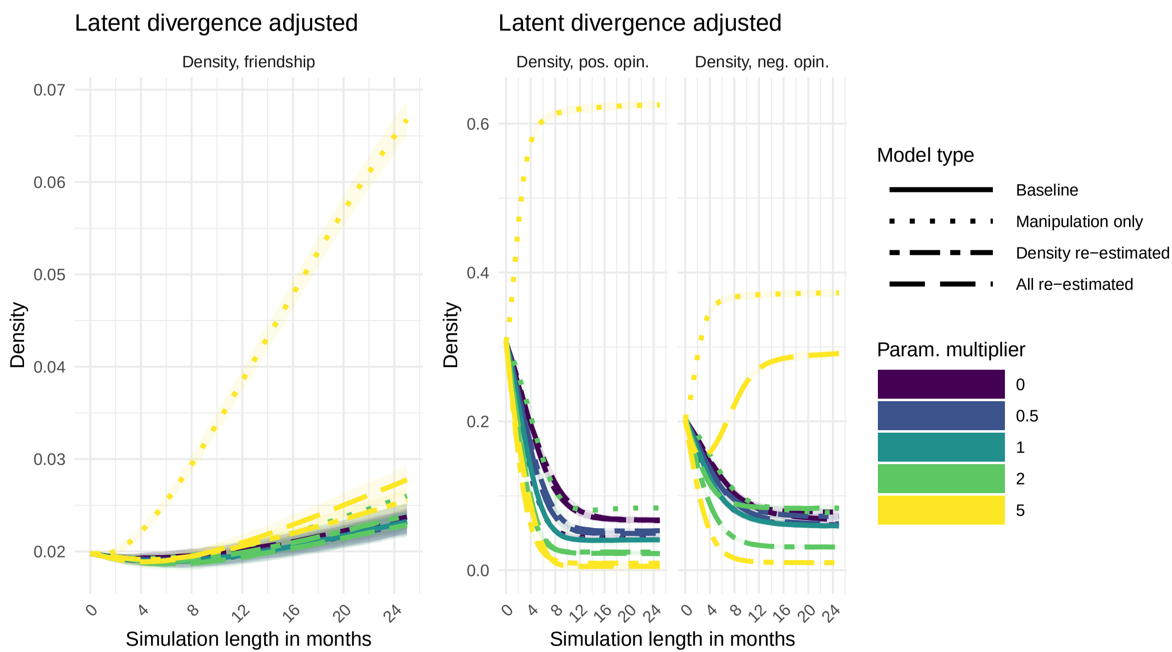


FIGURE 3.9: Densities of friendship, positive attitudes and negative attitudes over time when manipulating the latent divergence parameters. Ribbons indicate lower 25% to upper 75% quantiles.

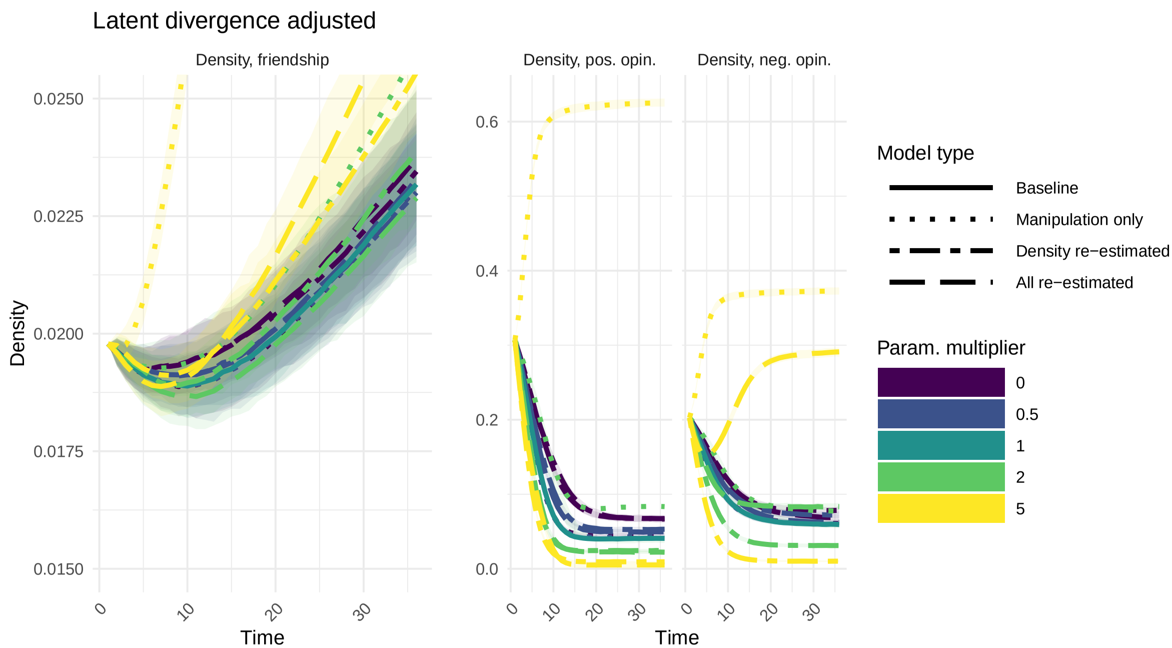


FIGURE 3.10: Truncated y-scale of densities of friendship, positive attitudes and negative attitudes over time when manipulating the selection parameters. Ribbons indicate lower 25% to upper 75% quantiles.

Summary of density results

Summarizing the effects of manipulations on the opinion networks, it is firstly apparent that the densities of the political attitude networks observed under the two kinds of model re-estimation (i.e. re-estimating density or re-estimating all non-focal parameters) are reasonably similar, more so than simulation without re-estimation. Secondly, the densities of the political attitude networks largely follow similar trajectories under manipulation of either selection or influence, with the tightest grouping under manipulations of selection – an effect which does not directly change the political attitude network. Thirdly, detailing more on the previous point, the density of the political attitude networks typically drops over the course of time from the observed data. This stabilizes between 0 and 0.1 when manipulating the social parameters (selection and influence), and typically between 0.0 and 0.2 when manipulating latent convergence and divergence. Several notable exceptions remain. When increasing the parameters of latent forces, the combined density of the two networks goes to one when the rest of the model is not re-estimated. When only the density parameter is re-estimated, the observed density drops near to zero. For the influence effects, the saturation of the political opinion networks occurs only when the parameter is multiplied by five, and is not reached by the end of the simulation. Interestingly, when re-estimating the model with the latent force parameters set to zero, the density increases rapidly too, but primarily for the positive opinion network. This is likely due to indegree effects on specific opinions, combined with a larger number of positive opinions than negative opinions at the start of observation.

Turning to the friendship network, it is first observed that manipulating selection leads to large differences in trajectories of densities, as might be expected: selection directly affects the social ties individuals hold. More notably, manipulating the latent forces has similar overall effects. Second, a similar pattern can be seen in friendship density as compared to political opinion density: under increased latent forces without re-estimation, density continually increases (but does not stabilize by the end of the simulation period). Under the manipulation of the influence parameters to five times their original size, without re-estimation of the model, friendship density also trends steeply upwards, although all other conditions have tightly grouped trajectories. Interestingly, under both re-estimation conditions when manipulating selection parameters, the smaller the parameter, the higher the final network density. The re-estimation method seems to matter less than the multiplier, as the conditions are tightly grouped. On the other hand, manipulation without re-estimation is grouped quite neatly around the baseline simulation, with the exception of the five times multiplier which is reasonably stable but higher than these other conditions.

Polarization

First, we again examine plots of the baseline model on both relational and ideological polarization in Figure 3.12 and Figure 3.11. These metrics range from zero to one[8].

In both cases, a drop in the overall level of polarization in both ideological and relational senses can be observed. This tendency appears to decelerate over time – indeed appearing somewhat stable by the end of the simulation for relational polarization, but less clearly for ideological polarization. Notably, relational polarization on average remains above the 0.5 expectation under a random distribution of attitudes.

Continuing to the computational experiment conditions, results of the simulations on the focal metrics, relational and ideological polarization, are presented in Figures 3.13, 3.15, 3.16 and 3.18.

As with the baseline model, the level of ideological polarization decreases over time under all manipulations of influence and selection. One exception to the general decrease in ideological polarization here is that when the rest of the model is not re-estimated and influence is multiplied by five, ideological polarization on average increases rapidly after an initial slight drop, eventually ending above the 0.5 midpoint of the measure. In the remaining two models with a five times multiplier, there is a slowing but not a reversal of the decrease

8 Note that ideological polarization is biased downwards by a shared component of the denominator in both the repulsion and attraction dimensions: namely, the asymmetric four-cycles where individuals agree on one topic but not another. Relational polarization is expected to be .5 under complete absence of sorting of attitudes amongst friends.

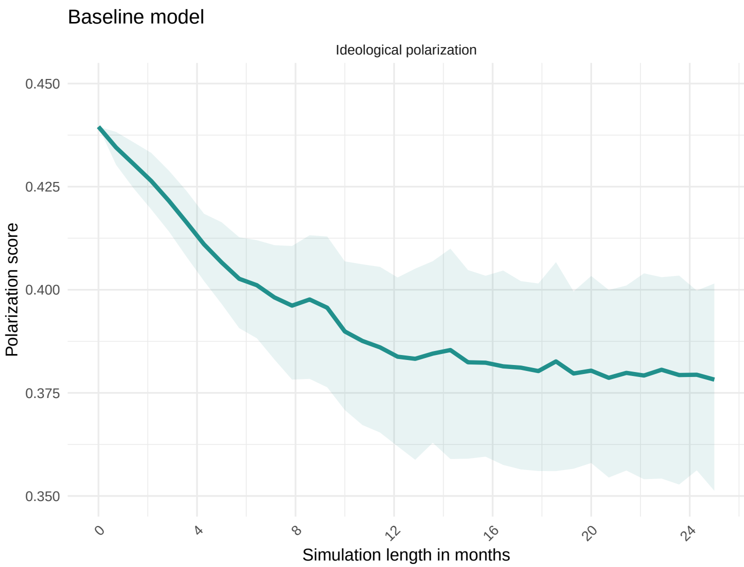


Figure 3.11: Change in ideological polarization under the baseline model. Ribbons indicate lower 25% to upper 75% quantiles.

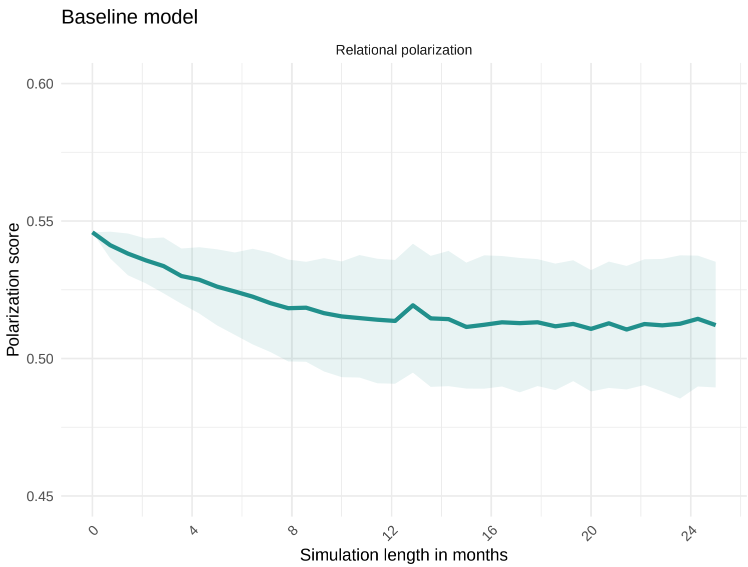


FIGURE 3.12: Change in relational polarization under the baseline model. Ribbons indicate lower 25% to upper 75% quantiles.

in ideological polarization under both re-estimated conditions. Both trendlines remain above all others during the initial drop, although this effect seems to wear off under the condition of re-estimated density towards the end of the simulation period, and for both of these models, the variability of the outcomes is large. Selection's effects on ideological polarization do not appear so impactful in forward simulation, with no clear differentiation on the timespan of the simulation.

Turning to the impact of latent convergence on relational polarization, it can again be observed that an increase in latent forces at the five times multiplier leads to starker differences relative to the social parameters of selection and influence. As mentioned before, only one of the two re-estimation methods resulted in a converging model. With density re-estimated, this multiplier leads to an initial increase followed by a brief drop. A similar pattern is shown also with a two times multiplier, and for both of these density-reestimated models, the level of relational polarization becomes quite unstable relative to the other conditions and baseline. Indeed, after month nine, most simulations for density-reestimated models with a five times multiplier return to the midpoint for relational polarization leading to the interval becoming very narrow[9]. This is contextualised by the results regarding network density: as described above, the number of political attitude ties becomes particularly low when increasing the effects of latent convergence and re-estimating model density, so few changes in attitudes can lead to larger changes in the level of polarization. Manipulating the parameters without re-estimation surprisingly leads to more stable changes, with the level of relational polarization showing an overall drop relative to the start of the simulation grouped tightly together, but nonetheless largely ordered according to the parameter multipliers. This suggests a near-universal trend towards decreasing of relational polarization to near-random levels, with slight differences in final values determined by the strength of the latent forces.

9 Further examination of the data revealed that these values tend to be bunched either in the middle or at the two outer edges of the distribution, hence the disappearing interval.

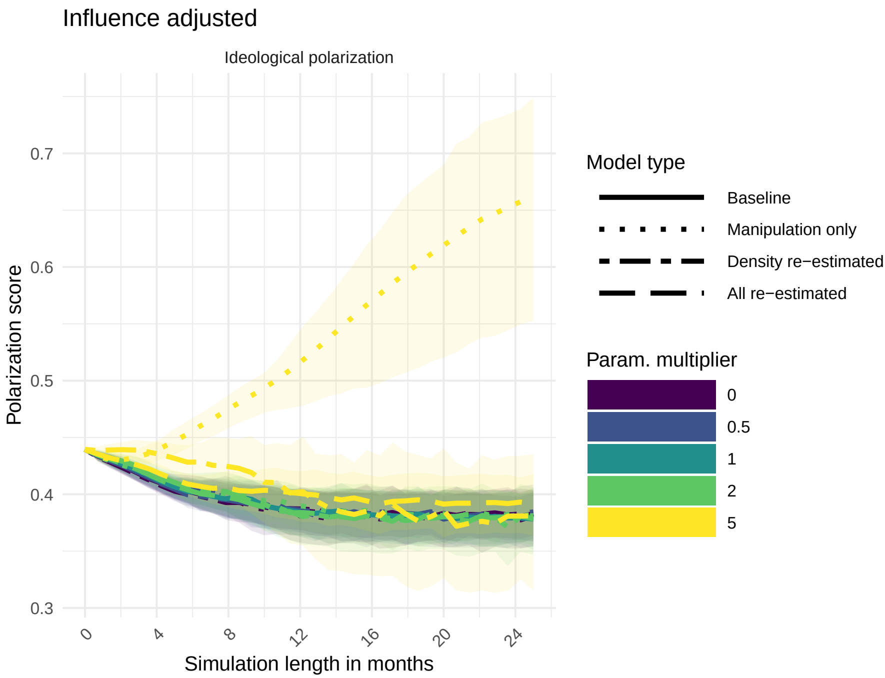


FIGURE 3.13: Change in ideological polarization when manipulating the influence parameters. Ribbons indicate lower 25% to upper 75% quantiles.

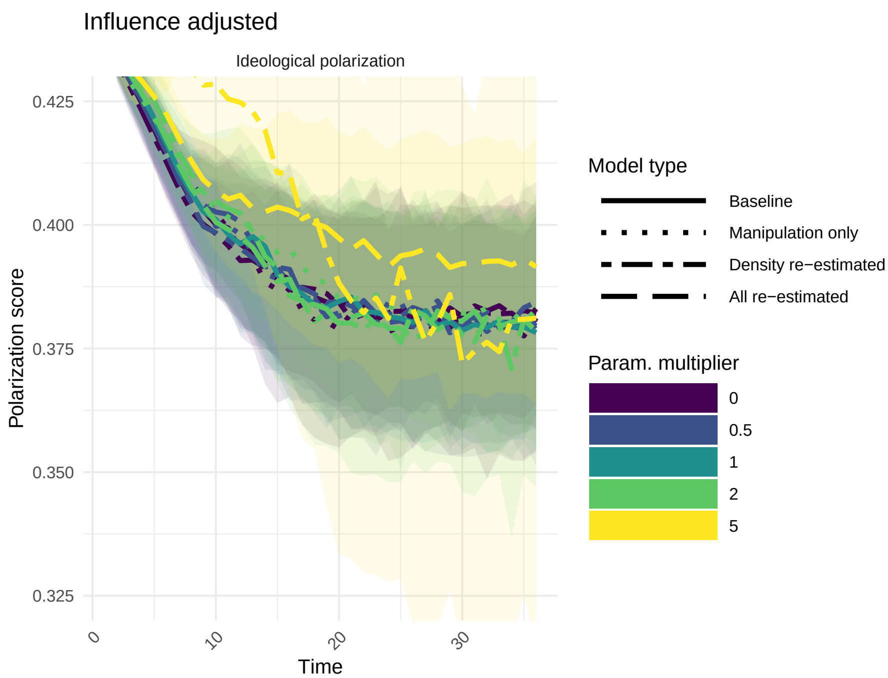


FIGURE 3.14: Truncated y-scale of change in ideological polarization when manipulating the influence parameters. Ribbons indicate lower 25% to upper 75% quantiles.

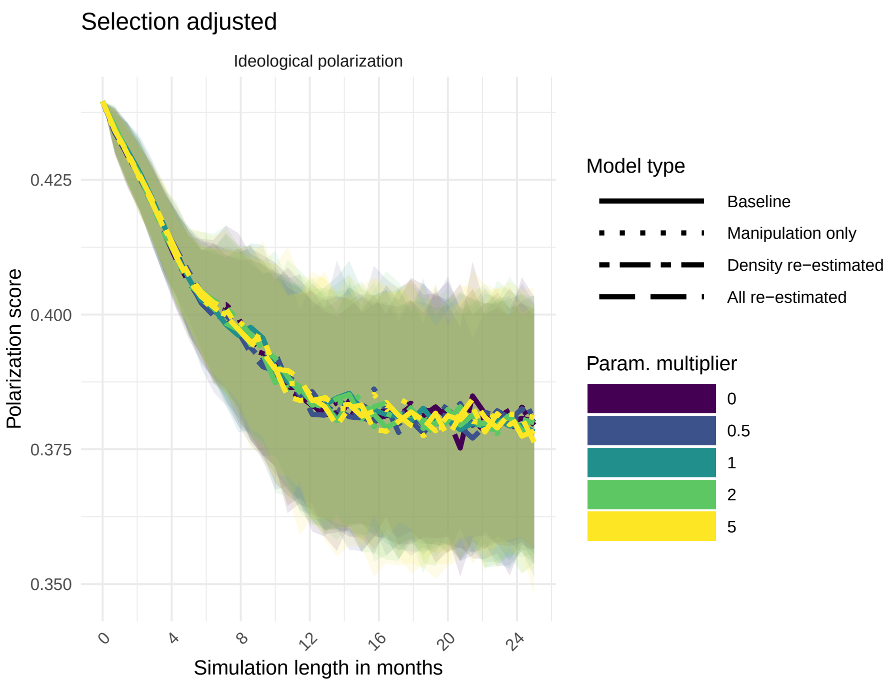


FIGURE 3.15: Change in ideological polarization when manipulating the selection parameters. Ribbons indicate lower 25% to upper 75% quantiles.

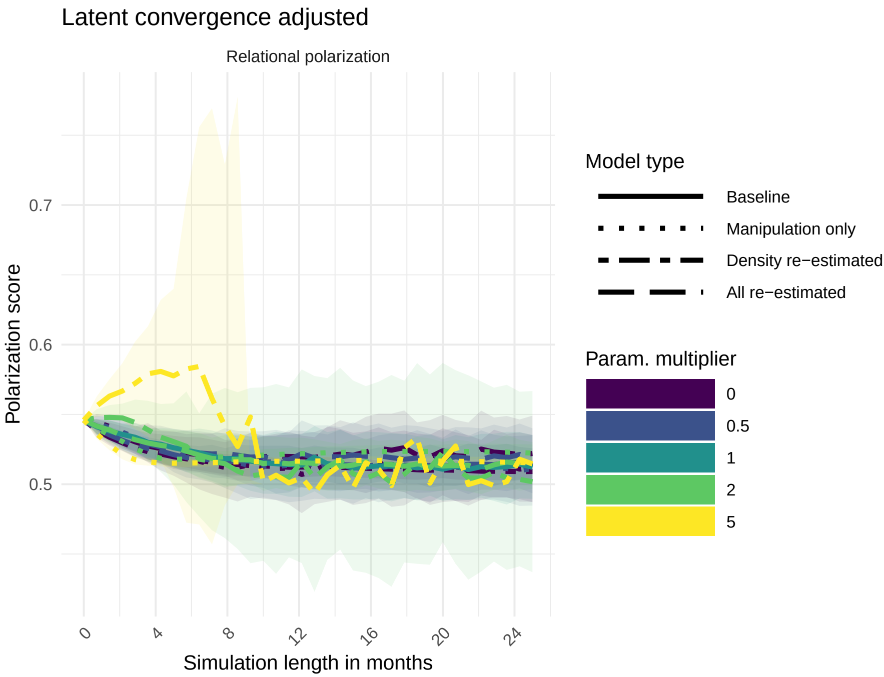


FIGURE 3.16: Change in relational polarization when manipulating latent forces. Ribbons indicate lower 25% to upper 75% quantiles. Due to near-zero densities of the opinion networks, relational polarization varies very little after the 9th month.

FIGURE 3.17: Truncated y-scale of change in relational polarization when manipulating latent convergence. Ribbons indicate lower 25% to upper 75% quantiles. Due to near-zero densities of the opinion networks, relational polarization varies very little after the 9th month.

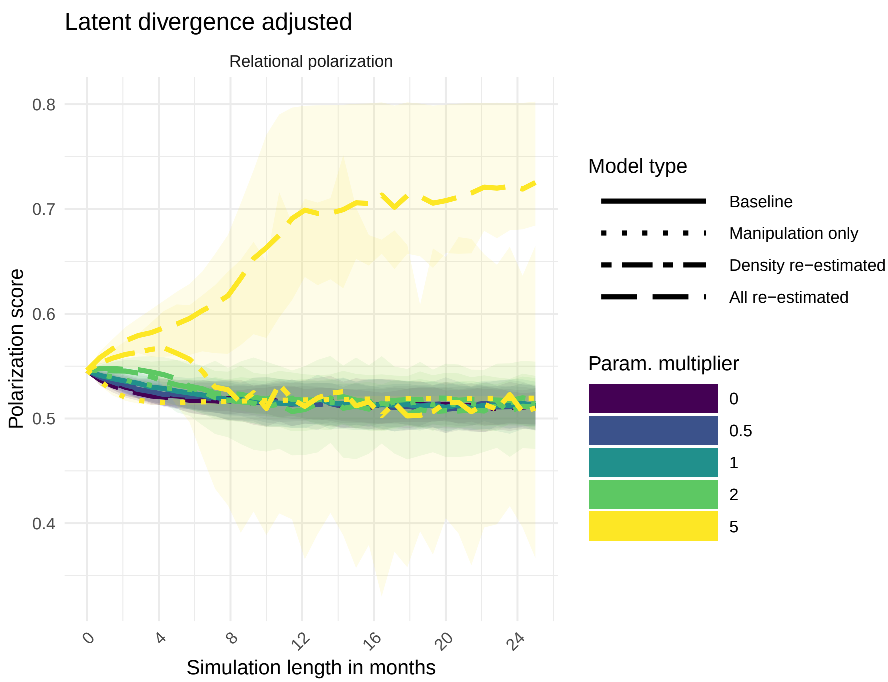


Figure 3.18: Change in relational polarization when manipulating latent divergence. Ribbons indicate lower 25% to upper 75% quantiles.

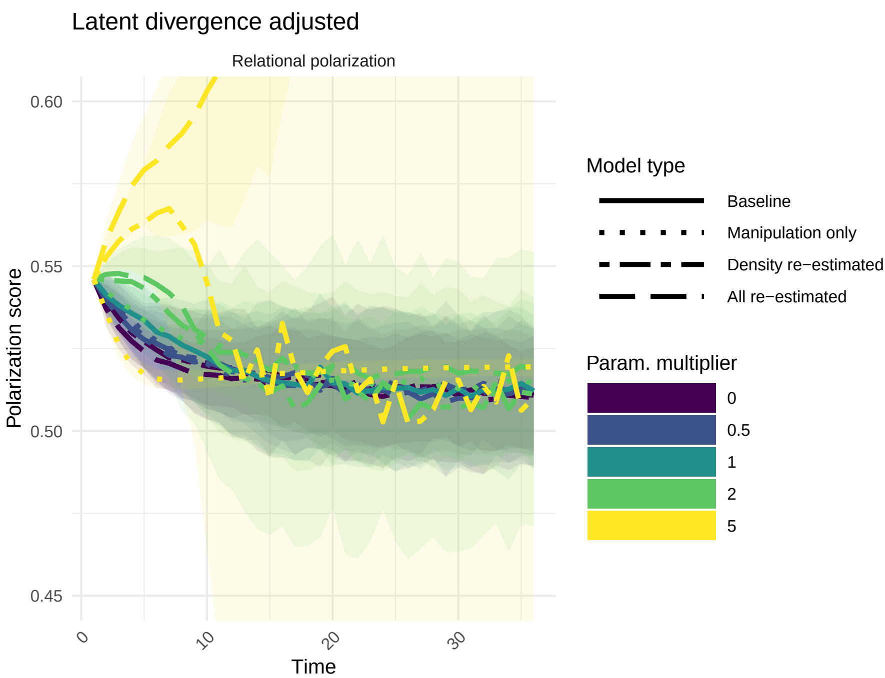


FIGURE 3.19: Truncated y-scale of change in relational polarization when manipulating latent divergence. Ribbons indicate lower 25% to upper 75% quantiles.

Summary of polarization results

Overall, polarization tends to go down under all manipulations and both in relational and ideological terms. This tendency is circumscribed by a drop in density of political attitudes, and an increase in friendships. Most differentiation between models is seen when increasing parameters substantially, whereas decrease and removal does not cause a substantial change over the course of the simulation.

DISCUSSION

In this study, we examined the consequences of empirically observed processes of political selection and influence, and latent attitudinal convergence and divergence in a small, real-world friendship network of students. Drawing on calls for a stronger specification of the reality underpinning agent-based models, we applied a stochastic actor-oriented model as a simulation engine and demonstrated three methods for its use. This offered us several advantages over standard ways of specifying ABMs. These include the testing of proposed mechanisms against real data, allowing an empirical estimate of the strength of these tendencies, a stricter specification of social ties, which are now also permitted to follow an endogenous and dynamic model, and finally, the possibility to specify a time frame based on the estimated rate of change in empirical data, allowing us to specify a time frame of two years. In what follows, we consider the results in light of the empirical case, followed by a discussion of the modelling methods applied.

Observed model outcomes

Using the SAOM, we assumed sequential updating with a probabilistic update function to examine the long-term consequences of the observed processes and the sensitivity of the system to change. Applying previously-defined statistics capturing both relational and ideological facets of polarization, we found that the group of students studied tended to decrease in polarization on both facets, preferring to maintain more friendships and fewer political opinions. While an increase in social influence would have slowed this drop in ideological polarization, the end result appeared quite similar with the exception of one re-estimated model and a model without re-estimation. The former of these suggested slightly more ideological polarization, and the latter dramatically more polarization by the end of the simulation. A change in homophilous selection had very little impact on polarization. Changes in latent forces had relatively large effects on polarization, with a clear association between the magnitude of parameters and the tendency towards polarization. Nonetheless, caution is warranted due to rapidly changing estimates brought about by a tendency towards extremely low densities under a density-reestimated specification. Overall, these results suggest two conclusions about this community: First, even with stronger tendencies towards forces often considered polarizing, the tendency would be to become less fractured by political attitudes. Simpler ABMs will tend to find e.g. consensus or polarization as a guaranteed outcome under some specifications (Flache et al., 2017), where we highlight an important point that in a reasonably specified timeframe, stable points of our metrics of polarization do not necessarily fall to the extremes. Second, relational polarization is affected much more by unobserved processes than ideological polarization is affected by observed ones, as indicated by the stronger sensitivity of polarization outcomes to changes in the four-cycle parameters indicating tendencies towards latent (dis)agreement maximization. This latter finding points towards exogenous factors (such as media environments) affecting how individuals are grouped with like-minded others.

Comparing modelling methods

Turning to variations in the model reestimation procedure, a key observation is that both density and full-model re-estimation methods tend to group together under equal parameter multipliers. This implies that these are producing similar networks for the metrics of interest, and that these alternative plausible explanations of the original data do not differ dramatically in consequence.

On the other hand, models produced via manipulation without re-estimation caused one case of extreme polarization, and resulted in fast increases in the density of the friendship network and saturation of the political attitude network. Heuristically, such extreme and rapid results suggest that these are the least plausible models of polarization presented in this research, while the less extreme results are more plausible. At least, this is the case if we assume that it is not a common occurrence that dramatic, multidimensional political fault lines emerge in a relatively short period amongst cohorts of students.

Although more plausible than a manipulation-only condition in terms of outcome, the reestimation methods applied deserve some attention: Questions remain about the validity of re-estimating a model with parameters partially fixed, and may depend on the desired purpose for using these methods. In the current paper, we have examined the effects of causes, not the causes of effects. That is to say that while we might validate our model against real data (i.e. 'effects') and seek to improve it via the addition of further structural terms (i.e. 'causes') to the model, we are doing something fundamentally different than examining what would happen to levels of polarization in the counterfactual case that moderating effects of context change the forces which adjust opinions and social ties. We do not yet truly understand the processes by which social and opinion networks form – instead taking a structural-statistical approach, and atheoretically re-estimating relevant parameters. Under what conditions would these changes to the strength of processes be most plausible? Do they continue to represent the network's development well over time? These questions are important but difficult to answer.

With this in mind, by interpreting models such as the ones presented in this work, we may fix parameters to understand consequences of a process. This does not carry with it the assumption that these manipulated parameters are the true effects, but shows how sensitive the outcome, polarization, is to changes in the theorized causal parameters. This does not guarantee that the observed data are explained by a model with these manipulated parameter values.

If the model does explain the data (i.e. if fit is adequate under re-estimation), however, it shows us that these potential parameter sizes could already be consistent with reality, and may shed some light on the real trajectory of the network in question.

Model interpretability and limitations

Some further points remain as to the interpretability of this collection of models, with the novel empirical calibration applied. Firstly, the polarization metrics tend to vary quite substantially under some conditions. In some cases, this indicates only a few opinions or a few individuals holding an opinion. This reduced-issue focus initially seems undesirable, but it may actually be more interpretable than at first glance. Particularly in the former case, it is entirely plausible that a single attitudinal object is the focus of some fracturing in a social group. Indeed, single-issue differences of opinion have been considered evidence of polarization, both in ABM research (DeGroot, 1974; French, 1956; Hegselmann, Krause, et al., 2002) and empirical research (Bramson et al., 2016; DiMaggio, Evans, & Bryson, 1996; Fiorina & Abrams, 2008). On the other hand, rapid switching of the level of polarization due to (few) polarized issues does suggest deviation from realism in either the model or the conceptual measure of polarization.

Secondly, the model we use explains the structure of the network in aggregated time well, but when examining wave-by-wave the fit is closer to a significant deviation from observation. This hampers the ability to draw a predictive interpretation in particular from forward simulation. However, we do believe it shines some light on what the estimated processes' consequences are in general in terms of their impact on the network's development.

Thirdly, the two-year period applied is somewhat arbitrary, determined on the basis of practical constraints such as computational manageability but also on what seems 'reasonable'. Not all metrics are stabilized under all models with this simulation period, unlike many other ABM research projects which simulate until stable. Nonetheless, keeping a reasonable simulation window when a time window *is* specifiable is a justified choice – we would like to know about tendencies for a (somewhat) foreseeable future. One could question whether we should indeed think of two years as a reasonable simulation period; the student community did not exist past its third year (and we simulate from the end of the first 9-month academic period), implying that the network boundaries would disintegrate then. Indeed, we know that the boundaries substantially changed in the period following that in which we collected the data used to estimate the model. This would have the potential to change network dynamics, particularly if the selection is non-random. For instance, one could ask whether students with unpopular opinions are less likely to integrate socially, and therefore drop out. In sum, these considerations suggest we should not look at this model as a direct predictive tool, but as a model giving some suggestion of ordering and magnitude of different effects' consequences in the near future.

If we nonetheless *were* to examine the models' consequences, we propose that we should consider the following: Firstly, we should carefully consider the duration of the simulation – if we estimate a model on data collected three months apart, we should not simulate 100 years ahead. Our cohort would have been almost entirely replaced, both in terms of their existence in the long term, but also in their short-term inclusion in the boundaries of the network due to the undergraduate setting. Indeed, data from later in the students' studies exist, but due to the substantial change in the composition of the network, and increasing non-response, it would be hard to justify applying the same model estimates to this data given the scaling issues in generative network models (Duxbury, 2021; Vörös et al., 2021). Secondly, we should also assume that these processes cannot continue indefinitely. At some point, any given issue may become nearly completely irrelevant and therefore less subject to change or important in selecting social ties, or social ties may be so strongly developed that we would no longer expect them to disappear. For at least these two reasons, shorter time frames make for more plausible extrapolation.

CONCLUSION

As argued in Chapter 2 of this dissertation, we believe that the data set used is one of the most appropriate in current existence to the type of network model estimated here. Given this, the observation that socially-driven effects occurring within the network are of relatively low impact on polarization is an important one. Aside from assumptions on time, tie types, and dynamics covered directly, this observation draws into further question the explanatory capabilities of many agent-based models of opinions in social networks: under which conditions should the commonly applied influence processes be expected to reflect reality, and at what timescale? Would such a timescale ever be realistic with regard to boundary changes in the network? Future research should aim to answer these questions more precisely. Extending to data sets similar to that applied in this paper would aid in understanding the replicability and generalizability of these results. More broadly, the use of the SAOM as an ABM tool is a promising step to link ABMs of actors in networks on varied topics to reality, and may benefit users of both approaches.

4

CHAPTER 4 – DOES POLITICAL DISCUSSION CHANNEL SHORT-TERM VOTE INFLUENCE?

This paper examines effects of political discussants and friends on individuals' choice to vote and choice of vote in the unique context of Switzerland's direct democracy. With a pre-post test design spanning multiple referenda, I test for assimilative influence on choosing to vote and how one votes. I furthermore examine whether undecided individuals are more likely to be influenced, and whether this effect is strongest amongst the politically interested. I additionally consider the effects of political knowledge as a cause of reduced influence from discussants. Testing theories of disagreement's effects on turnout, I also consider effects of being isolated in one's opinion. Overall, results of a permutation-based analysis of three cohorts of university students provide moderate support for any of the hypotheses: although correlations are found between individuals' and their discussants' vote choices, these disappear once controlling for pre-referendum intentions, making a suggestion that there is primarily homophilous selection. Amongst friends, controlling for prior voting intentions nonetheless suggests influence on choice of and choice to vote. Opinion isolation is too rare to be tested as a factor in voting behaviour in both networks, and no differences are observed between the undecided, the more interested, or knowledgeable in their susceptibility to influence.

This article is currently in preparation as:

Mepham, K. (TBD). *Do political discussion networks channel vote influence?*

INTRODUCTION

Political discussion is a fundamental part of a functioning democracy. Exchange amongst individuals may cause one to gather new information, understand other unconsidered sides of an issue, or apply heuristics ultimately to choose the 'correct' side of an issue or party to stand with; i.e. aligned with one's interests and values (Downs, 1957; Huckfeldt & Sprague, 1995; Richey, 2008). Beyond such benefits, political discussion may inform one of the social norms surrounding participation such as turnout, or foster a sense of civic engagement (Klofstad, 2015; McClurg, 2003, 2006), enhancing or depressing the societal consequences of one's position. In this paper, I examine whether there are potential normative and/or informational effects of one's political discussants on individuals' turnout (i.e., the choice to vote) and on the choice one makes once one has decided to vote (i.e., the choice of vote).

Three core contributions in this work are the following: firstly the unique whole-network longitudinal network data which overcomes the predominance of self-reports used elsewhere. Secondly, I use pre- and post-vote measures of voting intentions, which allows for stronger causal inference on social influence than cross-sectional designs (for instance, studies using the international Comparative National Election Project surveys, or the U.S. General Social Survey, but see Baker, Ames, & Renno, 2006; Huckfeldt & Sprague, 1995). Third, I contribute in a new voting context by using specific policy votes as the subject of influence: I study direct-democratic referenda in Switzerland.

In addition, I consider differential effects on the potentially impactful undecided voters. Previous investigations have often focused on the effects of media campaigns and individual attributes such as implicit attitudes in determining the undecideds' ultimate choices in voting, to mixed results (Friese et al., 2012; Hopmann et al., 2010; Raccuia, 2016, but see also Ohme, De Vreese & Albaek, 2018), limited to an atomistic analysis by omitting the potentially impactful functions of discussion networks. Finally, I explore potential moderating roles of political interest and knowledge. These tests are made in the case of the direct-democratic Swiss system, with the aim of understanding the effects of alters' opinions on voting behaviour, testing for social influence in terms of adoption of others' behaviour turnout. In a relatively short time window preceding and following referenda on the popular ballot, I examine whether there is a correlation between individuals' votes and that of their political discussants within the cohort of students in which they are embedded. I examine their choice to vote, and, conditional on their voting, their choice of vote.

The Swiss direct democracy offers an interesting case to study, and social network effects have received very little attention in this context. At four scheduled opportunities per year, there are popular votes on changes to the constitution, changes to laws approved by the federal assembly and against which at least 50,000 people register their opposition, and popular initiatives; changes to the constitution proposed by citizens to which at least 100,000 register their support. Turnout in the votes studied here ranged from 34.5% to 54.8% at the population level (Federal Statistical Office, 2017, 2018), though turnout amongst the respondents in our sample is much higher (see Table 4.1)[1]. Multiple issues can appear on the ballot at the same time, and turnout for these simultaneously-treated issues typically differs by less than half of a percentage point. Overall, this makes the question of turnout particularly interesting, as multiple topically independent issues have an impact on the representativeness of the voting outcome. The effects of political discussants are understudied in this context focused on such single issues. While electoral votes have been subject to frequent examination (Rolfe & Chan, 2017; Santoro & Beck, 2016) these are arguably *a priori* much more polarized. Where individuals' opinions and partisanship are stronger and deeper seated, they may apply motivated reasoning; simply picking and choosing the information they attend to, and interpreting opposed information in ways that can justify or otherwise reinforce their existing opinion (Lord, Ross, & Lepper, 1979; Taber & Lodge, 2006). Opinions surrounding referenda are less intense and more subject to change from outside influence than opinions on political parties, though some topics such as COVID measures attracted higher public interest as reflected in turnout (Federal Statistical Office, 2021).

1 This aligns with statistics on education and turnout in referendum votes in Switzerland, see VOTO (2022).

Table 4.1: Referendum topic, turnout, date of votes and date of surveys.

Referendum	Sample turnout	Vote date	Date pre survey	Date post survey
Energy law[a]	73.3%	21 May, 2017	13 Mar., 2017	29 May, 2017
Pension reform[a]	79.6%	24 Sep., 2017	29 Aug., 2017	11 Dec. 2017
Tax incr. for pension[a]	69.3%			
TV license fee	71.1%	4 Mar., 2018	11 Dec., 2017	12 Mar., 2018
Full money initiative	67.9%	10 Jun. 2018	28 May, 2018	28 Aug., 2018
Gambling act	69.1%			
Fair food initiative	67.9%	23 Sep., 2018	28 Aug., 2018	17 Dec., 2018
Food sovereignty	69.8%			
Bike initiative	69.2%			

Note.[a]Only recorded in Cohort 1

Whole networks and political outcomes

I use data from the Swiss StudentLife Study, a longitudinal, whole network panel study of three cohorts of undergraduate students at a Swiss technical university (Vörös et al., 2021). Data include measures of political discussion networks, voting behaviour, and other potentially relevant attributes of participants. Employing panel data, in particular around the period of the relevant vote in a way that allows for a baseline test of opinion and thereby stronger inference (e.g. Bello & Rolfe, 2014). The Swiss referendum context – where vote decisions may be made relatively late and are based on more volatile opinions (LeDuc, 2007) are a potential ground where discussant influence may be heightened, and in this university context where voting behaviour may be more susceptible to change (Bhatti & Hansen, 2012) this effect may be further enhanced.

Whole network data allow for better inference of influence, when considering that individuals tend to show a bias towards perceiving agreement in their online friendship network even given political discussion (i.e., agreement is perceived as even higher than the already increased baseline agreement between friends compared to random pairs of individuals; (Goel, Mason, & Watts, 2010; Laumann, 1969). This may be driven by preferential exposure of one's opinions to those expected to share that opinion, hiding one's opinion when it is expected to be unpopular, or indeed stereotyping (Cowan & Baldassarri, 2018; Goel, Mason, & Watts, 2010; Noelle-Neumann, 1974). Whole network designs such as the one advanced in the current study offer a similar advantage to the limited previous work applying snowball samples to electoral votes; providing self-reports of alters' opinions and choices (Baker, Ames, & Renno, 2006; Huckfeldt & Sprague, 1995).

Whole network studies in other university settings have found mixed effects of social ties on political attitudes and behaviours (Campos, Hargreaves Heap, & Leite Lopez de Leon, 2017; Lazer et al., 2010; Levitan & Visser, 2009) though typically these do not include explicit political discussion ties. In other work on this sample, evidence of friendship homophily when aggregated across multiple political issues over a year has been demonstrated (see Chapter 1 of this dissertation; Mepham, Vörös, & Stadtfeld, 2022). In general, however, evidence is mixed. Lazer et al. (2010) found no evidence of political partisanship-friendship homophily, although evidence of influence was found in their sample of public policy university students. Campos, Hargreaves Heap, and Leite Lopez de Leon (2017), in contrast, find no effect of peers on partisanship. They find that peer political engagement, with peers operationalized as those sharing a classroom, reduces partisan extremity and weak evidence that electoral turnout may be reduced. The context studied in Lazer and colleagues' as well as Campos and colleagues' studies related directly to partisanship, in opposition to the current study. Levitan and Visser (2009) demonstrated in a quasi-experimental design that network opinion heterogeneity may reduce attitudinal strength, potentially explaining results on opinion diversity and voting (Bello, 2012; Nir, 2011). Overall, evidence on relatively frequent, low-stakes single-issue votes in referenda, especially in combination with discussion networks, is absent.

Using the current design, I estimate the effect of one's local network on nine referendum issues. Single-shot, post-vote design estimates of alter (i.e., discussant) and ego (i.e. focal individual) vote choice influence effects are likely confounded by effects operating on longer time frames, and potentially homophily effects (Shalizi & Thomas, 2011). On the other hand, short-term effects as might be expected in the context of frequent referenda should be less affected by such confounding. In this way, the current study has several advantages over much previous work: capturing pre- and post- vote plans, using ego and alter self-reports of votes and opinions, and examining the effects of social networks on voting in the unique context of Swiss direct-democratic referenda. In what follows, I focus first on turnout, or the choice to vote, then turn to the choice of vote.

Turnout and the choice to vote

One factor explaining turnout may be the endogeneity of voting operating through the discussion network. Individuals' voting behaviours may be shaped by those of their political discussants. Some studies suggest a net benefit to turnout from political discussion in general (regardless of the intended behaviour of alters; Klofstad, 2007, 2015; McClurg, 2003). Another consideration here is related to the norms of individuals involved in the conversation[2]. If those who discuss politics are interested in politics, they are also more likely to vote. This may result in normative pressure encouraging egos to adopt the behaviour of their alters; as the proportion of voters amongst those with whom one discusses politics increases, so might one's own propensity to vote. The behaviour of one's (online) social contacts has previously provided evidence supportive of such effects (e.g. Bond et al., 2012; Jones et al., 2017), and experimental work has suggested that turnout is affected by one's household co-residents (Nickerson, 2008). A study of first-time voters in Denmark has suggested that changes in social networks associated with leaving the parents' home could also cause changes in turnout (Bhatti & Hansen, 2012). This hypothesis is of particular relevance in the current data, which concerns recently started undergraduate students. Assuming that a normative pressure may indeed affect voting behaviour, Hypothesis 1 is thus:

H1: As the proportion of voting alters increases, so does ego's propensity to vote[3].

2 While one function of discussant alters is providing information on how to vote (Klofstad, 2015; McClurg, 2003), postal voting is employed in Switzerland. Individuals receive the necessary documents to vote by mail, and may send them by mail too. This limits the potential of alters to provide information about the voting procedure itself.

3 Related to this issue, within the political discussion networks literature a debate has continued on the effects of disagreement. While a recent meta-analysis finds no evidence of an effect of cross-cutting exposure (Matthes et al., 2019), some suggest that disagreement reduces political participation including turnout (Klofstad, Sokhey, & McClurg, 2013; Mutz, 2002, 2006), others suggest an increase (Scheufele et al., 2004), and some find small-to-nil effects (Huckfeldt, Mendez, & Osborn, 2004; Nieuwbeerta & Flap, 2000). One potential explanation for these mixed results not considered in the aforementioned meta-analysis is given both by Jason Bello (2012) and Lilach Nir (2011): opinion isolation (also termed 'oppositional networks'), i.e. being the sole holder of a specific opinion in one's egocentric discussion network, reduces one's propensity towards political activity such as voting.Nir also notes weaker evidence that consensus results in such a reduction in the propensity to vote. Such results are potentially explained in two ways; in light of new information an actor might be less certain of their opinion and therefore more hesitant to vote, or in the case of publicly observable actions, might be concerned about the consequences of violating social norms (Hopmann, 2012). On the other hand, Nir's results surrounding consensus suggest that when an issue is already settled, individuals

Choice of vote

Beyond affecting turnout, political discussions in our social networks have long been considered crucial in the formation of mass opinions. Discussants may convey new information in the form of arguments, heuristic information, or social norms, all of which may affect individuals' opinions, and potentially, behaviour (Downs, 1957; Hopmann, Matthes, & Nir, 2015; Lazarsfeld, Berelson, & Gaudet, 1944). Early empirical examinations of social influence on voting in sociology include works suggesting that individuals look to their alters for political information. Seminal work by Katz and Lazarsfeld (1955), attempting to establish the effects of media campaigns, lead to a theory of a two-step flow of communication: Opinion leaders who are particularly engaged in a given domain mediate media campaigns' effects on many people. These individuals discuss with others, potentially influencing their opinions and decisions on the basis of their acquired information. In Switzerland, many people follow governmental recommendations in their choice of vote. However, not everyone attends directly to the information which the government offers (Trechsel & Sciarini, 1998), leaving open the possibility of alters to influence via sharing this, and other information.

One central debate surrounding a finding of cross-sectional homophily (i.e. greater than chance probability of individuals sharing an opinion; McPherson, Smith-Lovin, & Cook, 2001) in political networks questioned the direction of causality; are individuals influencing one another to become more similar, or are they selecting each other on the basis of political opinion[4]? While one

do not vote as they expect it to make no difference. Pattie and Johnston (2009) suggest that private forms of political participation (such as voting) on issues for which one is not a strong partisan may be suppressed when exposed to disagreement. In the context of the referendum voting, too, we may then expect an effect of isolation. Thus, a hypothesis was initially formulated that individuals isolated in their views are less likely to vote. However, this hypothesis proved untestable, as only one participant was found to be isolated in their views.

4 This stands aside from the issue of context; individuals are often nested in social contexts with reasonably high levels of homogeneity from the outset, meaning that random partner selection in one's local context may lead to the appearance of many homophilous ties if examined from a higher level; see e.g.Huckfeldt and Sprague (1995).

cannot say that this debate is fully settled, the evidence thus far has leaned in the direction of influence being a more reliably detectable force than selection, with some suggestion that political discussions are not particularly homophilous nor intentional, instead being mostly incidental to other relationships (Bello, 2012; Levitan & Visser, 2009; Minozzi et al., 2020; Sokhey & Djupe, 2011). Given the relative weakness of referendum issue opinions under study in this work, it is implausible that these kinds of political discussions are selected on prior shared beliefs on the subject matter. Interestingly, while seminal work on the matter of political discussant effects included opinions of the alters as reported by the alters themselves (Huckfeldt & Sprague, 1995), most research on this topic has operationalized the network-opinion context via egocentric reports. These have focused on the ego's perceived level of disagreement in the network, or on their beliefs about their alters' opinions (e.g., Bello & Rolfe, 2014; Huckfeldt, Mendez, & Osborn, 2004; Nieuwbeerta & Flap, 2000). This leaves a discrepancy between political influence research as practiced and two key findings in the literature. Firstly, research shows that individuals may be inclined to hide their true opinions when these are expected to be opposed to their alters', particularly if they expect to be in a minority (Cowan & Baldassarri, 2018; Noelle-Neumann, 1974). Furthermore, discrepancies have been found between perceived and actual opinions of one's social ties (Goel, Mason & Watts 2010; Huckfeldt & Sprague, 1995, pp.131; Laumann, 1969; Levitan & Visser, 2009), with a tendency to over-perceive alters' agreement with an ego. Understanding whether there is assimilative social influence, then, requires both the ego's and their alters' self-reports of their opinions and choices. Previous work has found evidence of an assimilative influence on electoral voting in countries such as Brazil, Germany, Japan, the Netherlands (Nieuwbeerta & Flap, 2000) US, and the UK (see Santoro & Beck, 2016, for a review). Similarly, I hypothesize that in this single-issue context that

H2: Ego's choice of vote will tend to change towards that corresponding to (the average of) their alters' support or opposition of the relevant issue.

A critical issue is how and whether people take sides when they are undecided. In the current context of often-divided votes (such as the Brexit referendum or the election of President Trump) these are of particular interest, as small shifts in the ultimate choices of the undecided could affect the outcome. Previous work has primarily focused on micro-level attributes such as implicit attitudes or macro-level campaign effects (Friese et al., 2012; Hopmann et al., 2010; Raccuia, 2016), with some distinction made between ambivalent (i.e. interested, but uncertain) and disinterested undecided voters (Ryan, 2017). Those who are undecided but aim to vote are obviously motivated to vote, but presumably insufficiently informed to make a final decision, holding potentially ambivalent attitudes or insufficient exposure to information to induce strong attitudes (Ryan, 2017). Given this combination of motivation and undecidedness, information should have a greater impact on their final choice than on those who are already decided. Political discussion alters may then also have a greater influence on them; the undecided's ultimate choice could be decided by alters' information about how others are voting; allowing for heuristic choices, about arguments relating to the vote, or how they should weigh the various relevant considerations. One might expect that the undecided intended voters are more likely subject to environmental influences such as that of their alters: to fulfill of casting a meaningful, 'correct' vote, they must gather information to make their decision.

H3: If ego is undecided, the assimilative effect of alters is stronger than for decided voters who had already decided at pre-test.

In early communications research, sender and receiver characteristics were suggested to be of importance in considering whether an influence attempt will be successful (Hovland, Janis, & Kelley, 1953; Hovland & Weiss, 1951). Later research has specified that those with stronger initial opinions, but also more informed individuals are less likely to change their mind in the face of new information (Kazee, 1981; Zhang, 2019). Reasons for this include a potential ability to apply stronger motivated reasoning, i.e. reasoning in favour of existing beliefs, and the potential that their opinions are already based on more information in general, so new arguments are less likely to outweigh the old.

H4a: Ego's pre-existing knowledge of an issue will make them more resistant to alters' influence.

Considering these individual attributes further, it stands to reason that the undecided and interested would be more likely to be more influenced by their environment: they would be more motivated to inform themselves when the opportunity arises. This bears some similarity to the division between the undecided but ambivalent and undecided but apathetic division (Ryan, 2017). Those most interested should not be apathetic and therefore should be more likely to seek out more information.

H4b: The effect on the undecided is strongest amongst those who are most politically interested.

While these hypotheses are made specifically on political discussants in line with prior literature, I also examine identical main models replacing political discussion variables with friendship-based equivalents. This is due to problems in the definition of political discussion ties, particularly for the respondents. For instance, one might ask what a discussion is, compared to one-sided divulgence of (political) information; or have difficulty determining what is 'political' (Hopmann, Matthes, & Nir, 2015). In addition, memory of interactions, in general, may be poor (Bernard & Killworth, 1977; Kashy & Kenny, 1990; Killworth & Bernard, 1976), and political discussions are often embedded in, and incidental to relations such as friendship (Bello & Rolfe, 2014; Levitan & Visser, 2009; Marsden, 1987; Minozzi et al., 2020; Sokhey & Djupe, 2011). Potentially the friendship measure captures political conversations where they go unmeasured, so it is therefore treated as a probably over-inclusive proxy for the focal independent variable.

METHODS

Data

[5]Data are drawn from the Swiss StudentLife Study, a study that monitored three cohorts of Swiss STEM undergraduate students at an elite technical university ($N_1 = 253$, $N_2 = 261$, $N_3 = 660$; Elmer et al., 2022; Vörös et al., 2021). This study monitored numerous aspects of the development of the students' social networks and social integration, including political opinions and behaviour. The data I use regard all three cohorts at six time points, taken from surveys spaced approximately three months apart (see Table 4.1). For Cohort 1, five planned voting opportunities are covered for nine referenda[6], while cohorts 2 and 3 cover six referenda at three planned voting opportunities. All referenda covered for cohorts 2 and 3 are a subset of those covered for Cohort 1.

5 This project was preregistered at https://osf.io/6acju. I deviate from this preregistration in several substantive ways. Firstly, an initial plan was made to apply ERGMs to understand the network's development. Given that the network is relatively sparse and highly embedded in friendships, consistent with other literature (Marsden, 1987; Minozzi et al., 2020), I do not model the network structure. With lower sparsity this might nonetheless be feasible; subsets of friendships may have endogenous tendencies towards political discussion occurring on top of the cross-network effects presented. Secondly, I do not use auto-logistic actor-oriented models nor QAP to analyze the network data. This is simply because the current method works for controlling for network effects (Krackhardt, 1988) while essentially being equivalent to a y-permuted quadratic assignment procedure regression and is much easier to estimate. Third, I opt for alters' support of a referendum over their choice to vote, as this allows for better inference where voting data are absent (e.g. if ineligible) but nonetheless gets at preferences in one's local network. Fourth, I incorporated individual characteristics (knowledge and interest) as potential moderators out of exploratory interest. Fifth, I do not exclude the question on TV licensing, as here is where the largest subset of observations upon which the analyses are based. Finally, I included the friendship analysis having seen the relative sparsity of the political discussion network, as explained at the end of the introduction.

6 These referenda concern topics of an energy law, change to taxes in support of pensions and a change in state pension conditions, the TV license fee, the gambling act, the full money initiative, the bike initiative, the fair food initiative and the food sovereignty initiative. A brief description of each topic is given in Appendix B.1, which shows an introductory text also shown to participants.

Responses

Due to the design of the data collection, the invited cohort decreases over time[7]. For the cross-sectional portion of the analysis, the number of unique individuals who are eligible to vote is 492, and they provide information on 1690 outcomes. Considering individuals observed at pre- and post-referendum time points, the corresponding number of individuals is 397, and 1362 outcomes. For Cohort 1, this stems from a 70% participation rate at the first used observation, down to 41% at the final used observation[8]. For Cohort 2, this ranges from 66% to 47%, and for Cohort 3, from 56% to 36%. Of the unique individuals represented in the dataset, respectively 125 (25.5%) of those observed post-referendum and 107 (27.0%) of those observed both pre- and post-referendum were women, whereas the total across the cohorts contained 22.4% women.

Variables and variable construction

Vote choice

From variables on nine Swiss referenda, I collected pre (i.e. intended) and post (i.e. retrospective report of) behavior[9]. A paragraph-length prompt informed participants of the proposed change to the law or constitution, and key arguments for or against the referendum. Participants were asked "Did [do] you [intend to] vote in the referendum on <topic>. Multiple choice answer categories included "Yes, I will vote against the referendum", "Yes, I will vote for the referendum", "No, I will not vote", and "No, I am not eligible to vote". For the final five of the referenda (at the last two time points, see Table 2), an additional pre-referendum answer category stated, "Yes, but I do not know whether I will vote for or against". Individuals who reported themselves as ineligible to vote were excluded from the main analysis, as their (intention to)

7 Students who drop out of their degree programme are ineligible for the study.
8 The participation rate is calculated here from the number of individuals who are registered as members of the cohort at the time of the survey.
9 First, a question surrounding perceived knowledge was asked, followed by the introductory paragraph, then support for the issue at hand, and finally the intended vote.

vote was not available [10]. From these responses I constructed the main variables:

Choice to vote was constructed as a binary variable, with "Yes, I will vote against the referendum", "Yes, I will vote against the referendum", and "Yes, but I do not know whether I will vote for or against", treated as one, and "No, I will not vote", as zero. In Models 1a and b, ego's post-referendum choice to vote is treated as the dependent variable, while in model 1b, *ego's choice to vote* at pre-referendum observation is used as a control.

As several referenda occur simultaneously, and the choice to vote is highly correlated at the same time points, each referendum opportunity was treated as the same choice. Out of 471 cases of individuals who reported on their voting behaviour at a time points in which multiple votes occurred simultaneously, only 24 (5%) reported inconsistencies in their choice to vote on separate issues.

Choice of vote was constructed as a binary variable, with all votes in favour treated as ones, and all votes in opposition to the referendum treated as zeroes. All other cases were excluded. In Model 2, post-referendum vote choice is used as the dependent variable.

Network ties

A question was asked about individuals' *political discussants*. The question translates from the original German to "With whom of your fellow students do you discuss political issues?"[11]. An additional question was asked about individuals' *friendships*. The question translates from the original German to "whom amongst your fellow students would you consider a friend?"

Using a name generator procedure with autocomplete and dropdown options, participants could name up to 20 cohort-mates as political discussants and/or friends. I used the ties at post-referendum, as these should correspond best to the political discussants in the period between pre- and post-

10 In exploratory analyses, vote-ineligible participants' overall support or opposition to a proposed measure is included.

11 In additional analyses, I made use of informal social groups of students to whom respondents reported belonging, in particular those reported to have political discussions as a focal activity. Results are comparable to those of the political discussion ties. See Appendix B.6.

referendum surveys. Since the intended tie corresponds to an event, I symmetrized an adjacency matrix such that in any pair of individuals, if one reported a tie, it was treated as a tie for both parties[12].

Referendum support

For each referendum, a question asked "Considering the aforementioned information [in the question introduction, reported in Appendix B.1], to what extent do you support the [referendum]?" Participants could respond on a 7-point Likert scale, with options 1=Strongly opposed, 2= Opposed, 3=somewhat opposed, 4=undecided, 5=somewhat in favour, 6=in favour, 7=strongly in favour. *Ego's support* of a given referendum at pre-referendum observation is used as a control variable in models 2b,d,f and h.

From these variables, composite variables were constructed testing the main hypotheses:

Proportion of voting alters was constructed by taking the choice to vote dummy variable, and taking its mean amongst alters at the previous time points. This is used in model series 1 to test Hypothesis 1.

Mean alter issue support was constructed as the mean average of support for a referendum amongst the alters of an individual at pre-referendum vote. This variable is used in model series 2 to test Hypothesis 2.

[13].

12 In the case of individuals who skipped the political discussion question of the survey, these ties were treated as missing information and excluded from the dependent side of the analysis (idem for analyses using friendship instead of political discussion ties). However, if they reported voting or issue support and others reported ties to them, their vote choice information was retained for calculating the alter's political discussion-voting environment.

13 Opinion isolation was determined starting from the number of alters amongst which support for the given referendum was on the opposite side of the neutral midpoint from the ego in question. Firstly, for each referendum all alters of each ego were checked. Egos who had only disagreeing alters amongst those who reported their support of an issue were valued at one, all others zero. For each referendum time point, the mean value of this binary variable was taken. I additionally calculated this as the same set of support-reporting alters and their disagreement, minus the number who held no opinion in either direction. This produced a slightly more inclusive measure of isolation, which nonetheless resulted in only one isolate being available for the analysis.

Control variables

[14]*Number of political discussion partners* was calculated as the total number of ties per individual. This variable relates to social-political engagement, a factor associated with increased political participation amongst US college students (Klofstad, 2015), and previously theorized to signal the quality of political discussions (Bello, 2012). This variable is included in model series 1 as a control variable, as an increase in voting intentions amongst all individuals might then correlate with political discussions and therefore create a spurious correlation between ego and alter voting.

Political interest has been previously shown to be a factor positively related to turnout (Bello, 2012; Hopmann, 2012), and may confound turnout if individuals additionally are homophilous on political interest. A four-point measure taken from the European Social Survey was used, with the following prompt: "How interested would you say you are in politics? Are you. . . " and answer categories 1=Very interested, 2=Quite interested, 3=Hardly interested, 4=Not at all interested[15]. This variable is included as a control in model series 1, and as a subcomponent of interaction effects in model series 2.

Ego issue knowledge at pre-test is of potential impact, as previous work has suggested that knowledge impacts one's propensity to update opinions in the face of new information (Zhang, 2019). At the same time, it is more proximal to the issues being voted on than a measure of general political interest and may therefore also capture more localized interest in a set of issues. This was assessed by an item asking "To what extent do you agree with the following statement: I know a lot about referendum <X>?" Participants could answer on a 7-point Likert scale: 1=Strongly disagree, 2=Disagree, 3=Somewhat disagree, 4=Neither agree nor disagree, 5=Somewhat agree, 6=Agree, 7=Strongly agree. This question was asked prior to the text introducing key arguments for and

---

14 In analyses presented in Appendix B.5, I additionally include demographic controls of gender, financial status as a proxy of SES, and the main language in the region of origin within Switzerland, which do not change the substantive results. Women, however, are found to be more likely to vote than men in both models, and people who do not report a canton of origin are less likely to vote than those from the German-speaking reference category.

15 This was imputed with the last observed wave for cohort 1 in the final referendum time point, as the question was not included in the questionnaire at this point.

against the referendum issue. This issue knowledge, prior to the vote, is included in model series 1 as a control for any more localized political interest effects, and in model series 2 as a subcomponent of interaction effects.

I include fixed effects for *cohort*, and in the models of choice of vote, each *referendum* has another fixed effect.

In both the models of the choice to vote and choice of vote, I first test the model without including a control for individuals' pre-referendum vote intentions in model series 1, and pre-referendum issue support in model series 2. This is due to the potential that influence may be occurring prior to the window of the pre-test, thus controlling for pre-test opinion of ego may be controlling out the effects for which I am testing. The former model may capture slower-operating indirect effects of discussants, such as an assimilative change in overall political preferences that may have downstream consequences for voting. The latter provides a more direct test of the hypotheses presented, considering the theoretically specified shorter time frame in which the referendum issues at hand may be discussed (LeDuc, 2007).

Missing data

Complete case analysis is applied. Due to dependencies in network data, acceptable imputation techiques would increase the computational complexity of already-intensive permutations by orders of magnitude (Krause et al., 2020).

Analysis method

Due to the dependence of tie-based variables inherent in social network analysis, I first conduct a preliminary check of the association between political discussants' choice to vote and choice of vote using a QAP. Hypothesis testing analyses are conducted with *y*-permuted logit tests. This is performed in three steps. First, I compute the sample logit estimates on the relevant dependent variable (i.e. choice to vote, or choice of vote), pooling all available data to do so after listwise deletion of missing entries on the dependent variable and deletion of individuals with no ties. Second, I permute the dependent variable within each cohort by referendum category, generating 5000 permutations of the data. Third, I calculate the permutation logit estimates for each permutation, against which I can compare the estimate from the observed data. I calculate the absolute values of centered estimates, and use the percentile at which the observed estimate appears amongst these permutations as a non-parametric, two-sided p-value. This procedure is repeated for each model. This procedure preserves the features of the network, while controlling error rates for the non-independence of network data[16].

RESULTS

Descriptive analysis

Discussion network

Table 4.2 shows descriptive statistics of political discussion networks at time points relevant to this analysis, with exemplary network plots showing the networks and vote choices in Figures 4.1 and 4.2. The political discussion network is quite sparse, with the average respondent indicating 1.24 ties in the lowest case (Cohort 1, final wave) and 1.51 (Cohort 2, second post-referendum wave). Interestingly, the level of reciprocity is exceedingly low, at .41 at most, .24 at least, and .32 on average. These are lower than seen in friendship ties in the same dataset, a commonly used social network. Comparatively, the political discussion network is sparse, less stable and less clustered than friendship, of which the descriptives can be seen in Appendix B.2. Political discussion tie nominations are largely embedded in friendship nominations, ranging from 76% to 91%, with the mean proportion across all cohorts and included time-points of 84%.

It is clear from Figure 4.1 that those who have more political discussion ties are more likely to be voters; this association may well be a confounded

16 Tests in this style are somewhat conservative (Dekker, Krackhardt, & Snijders, 2007). Non-independence of the repeated measurements of the same actors across time is not controlled in this framework.

one driven by political interest, or may be caused by discussions resulting in greater political participation as suggested by Klofstad (2015). The Jaccard index, measuring tie stability as the proportion of ties present in two subsequent waves out of all ties appearing in both, averages .46, suggesting that ties are somewhat stable over time.

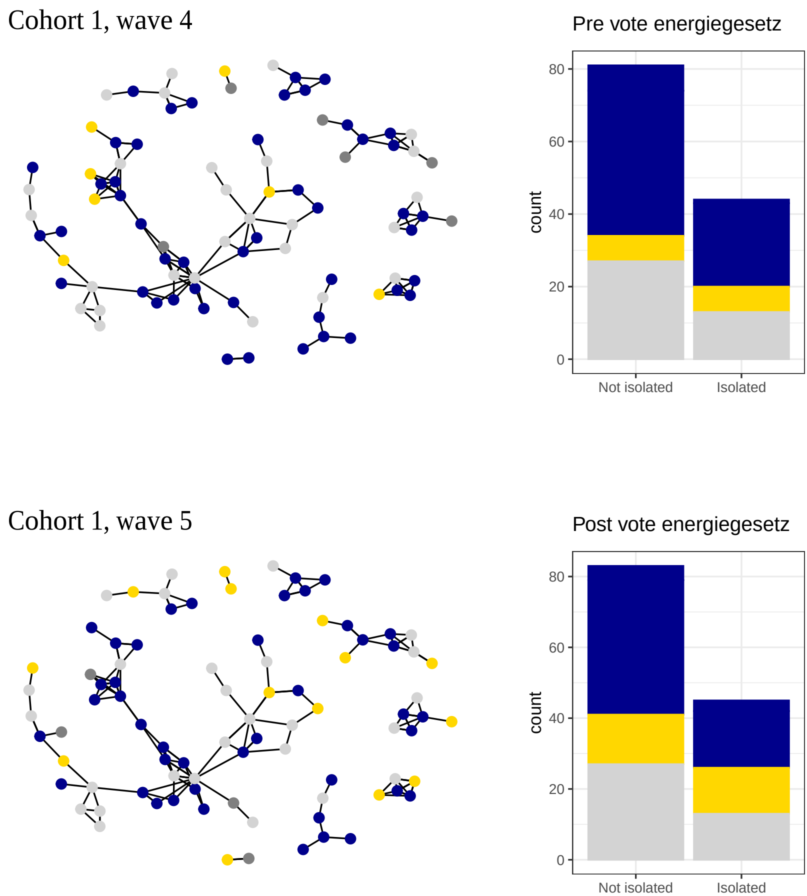


FIGURE 4.1: Exemplary political discussion network of cohort 1 with the choice to vote at two sequential time points, where ties are political discussions and nodes are people. Yellow indicates non-voting, blue voting, light grey ineligibility, and dark grey indicates nonresponse. Only non-isolates are plotted. The bar charts to the right indicate that individuals who choose to vote (blue nodes) are more common amongst those who have political discussions (left bar) compared to those who do not (right bar).

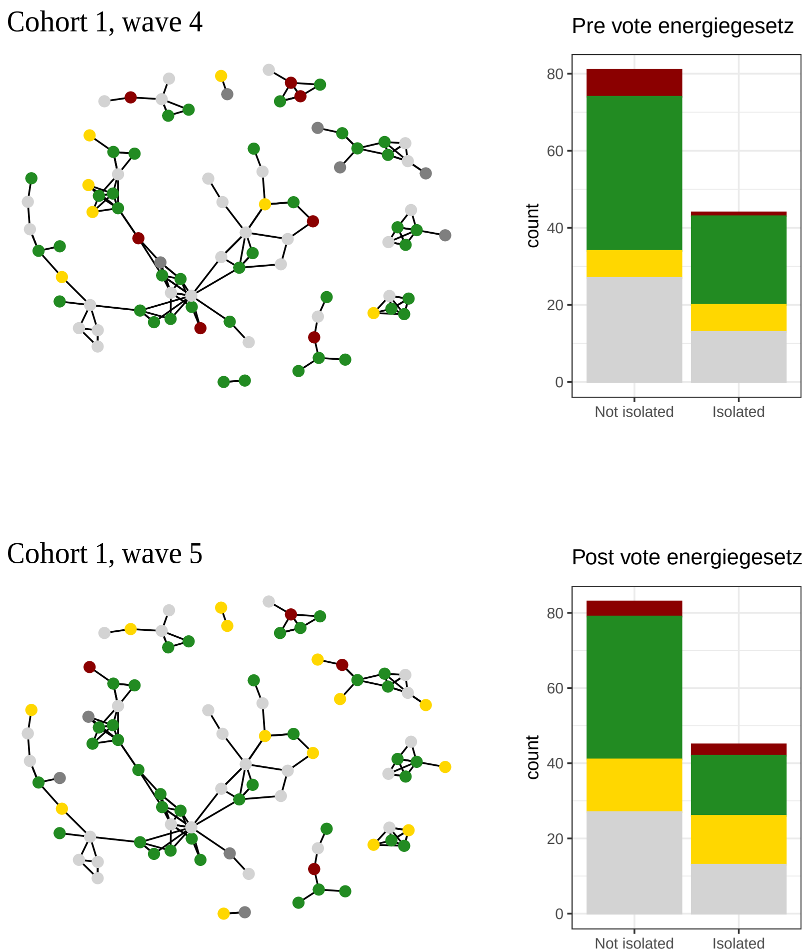


FIGURE 4.2: Political discussions as ties and individuals with their choice of vote as coloured nodes. Only non-isolates are plotted. It appears that some individuals changed their minds by time point two, and the bar chart on the right shows that the proportion of contra-voters decreases amongst discussants (left bar), but increases amongst isolates (right bar).

TABLE 4.2: Raw political discussion network descriptives by (post-referendum) wave and cohort

Cohort	Obs. month	N^a	N ties	N isolated	Mean outdegree	Max. outdegree	Reciprocity	Cluster coeff.	Jaccard t to $t-1$	N miss.
C1	May '17	194	185	88	1.43	6	0.31	0.45	0.45	65
	Dec '17	135	99	71	1.39	5	0.31	0.53	0.39	64
	March '18	133	98	74	1.36	5	0.33	0.48	0.53	61
	Aug '18	132	87	72	1.43	5	0.39	0.37	0.56	71
	Dec '18	142	63	93	1.24	6	0.41	0.43	0.49	91
C2	March '18	215	208	88	1.48	7	0.24	0.35	0.32	74
	Aug '18	210	151	96	1.51	6	0.25	0.28	0.51	110
	Dec '18	137	95	73	1.36	6	0.38	0.46	0.42	67
C3	March '18	620	444	277	1.25	8	0.29	0.39	0.41	265
	Aug '18	603	271	357	1.30	7	0.30	0.28	0.50	394
	Dec '18	339	198	163	1.31	9	0.36	0.28	0.48	188

[a]Estimated number of valid cohort members based on university records.

Table 4.4: Choice to vote at pre and post referendum survey aggregated across all referendum time points and cohorts.

Post Pre	Vote	No vote	Ineligible	Missing
Vote	574	129	6	291
No vote	38	68	4	50
Ineligible	4	4	267	128
Missing	119	62	69	

Choice to vote and choice of vote

Intended behaviour at pre-test, in particular intending to vote, is typically realized at post-test. The largest change is opting not to vote. Out of 1362 observations of pre- and post-test voting, 251 cases of an intention to vote prior to the referendum instead opt not to vote. Similarly, only relatively few instances occur of a change in the planned vote to the opposite position (a total of 79). Of 434 planned votes without a decided position, 174 were realized as votes against, and 106 as votes for the given issues, with the remaining 154 opting not to vote. Of 150 instances of an intention not to vote, 47 cases were observed where individuals nonetheless voted. A full table of the transitions including missing responses is given in Table 4.4 for choice to vote and Table 4.5 for choice of vote.

Examining the frequency table presented in Table 4.6, it is evident that testing an opinion isolation hypothesis is not possible with the current data; there is only one case of an individual being an isolate in their opinion; notably, there are also none in the friendship operationalization of the data (not shown here). As seen in Table 4.7, there is nonetheless never a consensus in the cohorts on the way to vote, and in 27.6% of observed cases, people are undecided at the pre-referendum survey. Given the above information, I continue with inferential analysis.

Examining the constructed variables testing Hypothesis 1, a bias towards zero and one proportions of alters voting is apparent (see Figure 4.3. Exploratory checks not reported here suggest that this feature of the variable

TABLE 4.5: Vote behaviour at pre and post referendum survey aggregated across all referenda and cohorts.

Post Pre	Vote against	Vote for	No vote	Ineligible	Missing
Vote against	373	22	42	2	162
Vote for	57	229	55	4	149
Vote undec.	174	106	154	3	328
No vote	35	12	103	6	104
Ineligible	1	3	7	426	273
Missing	118	80	119	135	

TABLE 4.6: Categorical variable frequencies for models of choice to vote

Variable	Category	Frequency
Cohort	1	185
	2	156
	3	396
Opinion isolate	0	625
	1	1
Choice to vote	0	153
	1	584
Issue knowledge (t-1)	1	110
	2	134
	3	118
	4	69
	5	129
	6	49
	7	20
Choice to vote (t-1)	0	69
	1	539

Table 4.7: Categorical variable frequencies for models of choice of vote, and vote choices frequencies per referendum

Variable	Category	Frequency
Referendum	Tax increase for pension reform	28
	Pension reform	27
	TV license fee	253
	Energy law	45
	Food sovereignty	116
	Fair food initiative	116
	Gambling act	142
	Bike initiative	116
	Full money initiative	143
Cohort	1	247
	2	232
	3	507
Undecided dummy	0	602
	1	229
Choice of vote	0	615
	1	371
Issue knowledge (t-1)	1	119
	2	184
	3	158
	4	94
	5	192
	6	70
	7	30
Choice of vote (t-1)	0	329
	1	238

becomes less apparent when considering only individuals with >2 alters, suggesting that this is not entirely due to a bias of individuals with one voting alter. A similar feature of the distribution is found both when using friendship and political discussion networks, though less severe when using the denser friendship networks.

Turning to the constructed variable of mean alter issue support, distributions are more symmetric and particularly in the case of friendship, more normal (see Figure 4.4). Nonetheless, there is a spike in a mean discussant issue support of zero.

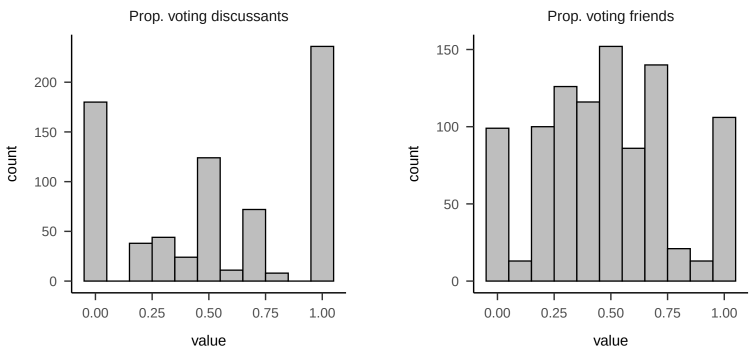


FIGURE 4.3: Distributions of the proportion of voting alters

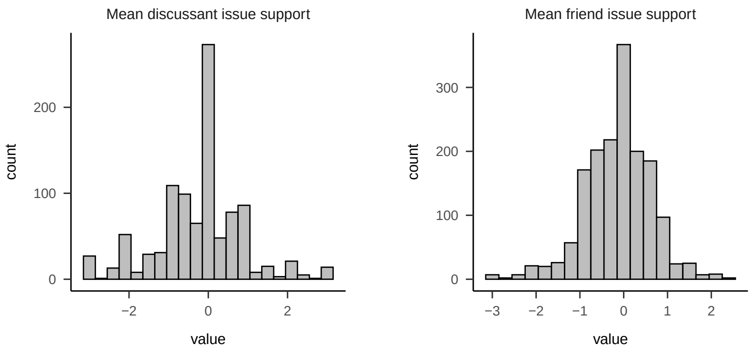


FIGURE 4.4: Distributions of mean alter issue support

Associations between the main and control variables are presented for choice to vote in Table 4.8 for political discussions, and Table 4.10 for friendship. I apply Spearman's r, since these variables are (partially) on an ordinal or binary scale. I do not account for the networked aspects of the data, so significance should be interpreted with caution.

The choice to vote is positively related to general political interest, issue knowledge, and the number of discussants but does not appear related to the proportion of voting discussants in this simple correlation; using friendship a small positive association is found between friends' choice to vote and the dependent variable.

Moving to the choice of vote, Table 4.12 shows Spearman's r associations between main and control variables for political discussants, and Table 4.14 for friendship. Here, we find that choice of vote is positively related to mean issue support of both discussants and friends. Being undecided is negatively associated with general political interest and issue knowledge. I note that further correlations are hard to interpret here, as there is no specific directional expectation for the choice of vote and, for instance, issue knowledge; these factors will vary from issue to issue, and only the interaction terms presented in the inferential models will have intuitive explanations.

As a check of the possibility of influence accounting for the networked nature of the data, I first test for the significance of the expected relation be-

TABLE 4.8: Pairwise Spearman's correlations of main choice to vote variables with discussants

	1.	2.	3.	4.	5.
1. Choice to vote (DV)	–				
2. Prop. voting discussants	0.05	–			
3. Choice to vote at $t-1$	0.38***	−0.01	–		
4. Political interest	0.31***	−0.02	0.3***	–	
5. Issue knowledge at $t-1$	0.31***	0.02	0.27***	0.47***	–
6. *N* discussants	0.18***	0	0.18***	0.24***	0.14***

Note. Pairwise cases selected, after selecting only cases with valid values for the dependent variable and prop. voting alters

TABLE 4.10: Pairwise Spearman's correlations of main choice to vote variables with friendship

	1.	2.	3.	4.	5.
1. Choice to vote (DV)	–				
2. Prop. voting friends	0.07*	–			
3. Choice to vote at $t-1$	0.35***	0.04	–		
4. Political interest	0.36***	0	0.33***	–	
5. Issue knowledge at $t-1$	0.35***	0.02	0.29***	0.46***	–
6. *N* friends	0.06	−0.03	0.01	0	−0.01

Note. Pairwise cases selected, after selecting only cases with valid values for the dependent variable and prop. voting alters

TABLE 4.12: Pairwise Spearman's correlations of main choice of vote variables with discussants

	1.	2.	3.	4.	5.
1. Choice of vote (DV)	–				
2. Mean discussant issue support	0.29***	–			
3. Issue support at $t-1$	0.66***	0.28***	–		
4. Undecided	0.01	0.13***	0.14***	–	
5. Political interest	0.02	−0.02	−0.11***	−0.17***	–
6. Issue knowledge at $t-1$	−0.08*	−0.2***	−0.2***	−0.41***	0.46***

Note. Pairwise cases selected, after selecting only cases with valid values for the dependent variable and prop. alters voting in favour

TABLE 4.14: Pairwise Spearman's correlations of main choice of vote variables with friendship

	1.	2.	3.	4.	5.
1. Choice of vote (DV)	–				
2. Mean friend issue support	0.33***	–			
3. Issue support at $t-1$	0.63***	0.32***	–		
4. Undecided	0.01	0.18***	0.14***	–	
5. Political interest	0	−0.02	−0.12***	−0.14***	–
6. Issue knowledge at $t-1$	−0.08*	−0.18***	−0.22***	−0.41***	0.45***

Note. Pairwise cases selected, after selecting only cases with valid values for the dependent variable and prop. alters voting in favour tween the existence of political discussion ties, and individuals displaying the same voting behaviour. Due to the non-independence of observations, I apply a multiple group quadratic assignment procedure regression (QAP) with 1000 permutations to generate a non-parametric p-value of the logit regression coefficient (Elmer & Stadtfeld, 2020; Krackhardt, 1988). Under this procedure, I test whether at the dyad level (i.e. between each pair of individuals) the same behaviour is more likely if two individuals are discussion partners. I include fixed effects for cohorts and referenda. Firstly, I check whether all categories except ineligible (a category that does not represent a choice) are more likely to occur between discussion partners. Indeed, this is the case ($B = .58, p < .001$). However, this result is somewhat hard to interpret as it includes individuals co-voting for, co-voting against, and not-voting. I then separate the results into voting versus not voting, and voting for versus against. Repeating the analysis for these two groupings, I find that the association is of similar magnitude for the choice to vote ($B = .47, p = .014$), but weaker and nonsignificant for the choice of vote ($B = .30, p = .082$). When controlling for pre-existing similarity, the results weaken substantially ($B_{all\, votes} = 0.36$, $p = .120$; $B_{choice\, to\, vote} = .47$, $p = .041$; $B_{choice\, of\, vote} = .20$, $p = .374$).

Considering friendship as the network channeling influence, similar results are found. Firstly, across all categories of vote choice, excluding ineligibility, a positive and significant association is found ($B = .59$, $p < .001$), a significant and positive association is found between friends and their choice to vote ($B = .47$, $p = .006$) a nonsignificant positive association is found between friendship and similarity of vote choice ($B = .21$, $p = .074$). Similarly, when including prior similarity these effects are typically nonsignificant ($B_{all\, votes} = 0.18$, $p = .005$; $B_{choice\, to\, vote} = .13$, $p = .25$; $B_{choice\, of\, vote} = .15$, $p = .319$).

y-permuted logit regression models

Choice to vote. In Table 4.16, I present the results of the two models of choice to vote, alongside the non-parametric p-values. McFadden's R^2, a pseudo measure of explained variance, AIC, a measure of the parsimony-explanatory power, and accuracy, capturing the proportion of correct classifications of voting or not, all indicate better model fit when including planned vote behaviour as a predictor (Model 1b). With regards to H1, that the proportion of alters voting affect's one's choice to vote, I find no support in either model as captured by the parameter "Proportion of voting alters"[17]. In the equivalent friendship-based models in Table 4.17, even after the inclusion of pre-test vote choice, it appears the proportion of voting friends is positively associated with one's own chances of voting, in support of H1. General political interest is a significant and positive predictor of voting, and there are significant positive associations between voting and issue knowledge. These results hold for both the models using the discussion networks and friendship networks. I find a positive and significant association between voting and the number of discussion alters an individual has only in Model 1a when using political discussion and in both 1a and b for friendship.

Choice of vote. In Table 4.18, I present the results of the models for choice of vote and their corresponding non-parametric p-values. Overall, there is a positive and significant preference to vote in the same direction as one's alters in two models for political discussion (Models 2a and e), which disappears when including ego's support of an issue at pre-survey (Models 2b, d, f, h), and when including ego knowledge of the issue and its interaction with alter support (Model 2c), partially supporting H2. On the other hand, the analysis of the friendship network supports H3 consistently with a positive coefficient even when controlling for ego's prior issue support, except for Model 2g which includes all interaction terms and their sub-components, and which has an inferior AIC to the equivalent model with only the interaction term for support and knowledge (Model 2c), furthermore suggesting it is not a parsimonious model.

Regarding H3, that the undecided are more susceptible to alter's influence, no support is found using the interaction term of alter's average support and being undecided in any model. H4a is also not supported, with a nonsignificant interaction between alters' mean support and being more knowledgeable, where a negative one was expected, and several cases of a positive and signif-

17 Tests using the number, rather than the proportion of alters voting, I similarly find no effect.

TABLE 4.16: Logit regression of discussant choice to vote on ego choice to vote

	Model 1a	Model 1b
Intercept	-0.7***	-1.36***
Prop. voting alters[a]	0.29	0.2
Number of alters	0.15*	0.05
Ego choice to vote[a]	–	1.67***
Ego knowledge[a]	0.39***	0.32***
Political interest	0.72***	0.64***
Accuracy	0.82	0.84
AIC	533.21	473.93
McFadden's R^2	0.17	0.22
Valid cases	629	608

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. Fixed effects for cohort excluded. For these effects, see Appendix B.4. [a]Variable constructed from individual data at t-1 [b]Here, 96.5% of individuals are predicted to vote by the model, while 80.8% did so.

TABLE 4.17: Logit regression of friend alter choice to vote on ego choice to vote

	Model 1a	Model 1b
Intercept	-1.18***	-1.76***
Prop. voting alters[a]	0.82**	0.71*
Number of alters	0.08**	0.07**
Ego choice to vote[a]	–	1.24***
Ego knowledge[a]	0.41***	0.35***
Political interest	0.84***	0.7***
Accuracy	0.78	0.8
AIC	763.72	703.16
McFadden's R^2	0.19	0.21
Valid cases	825	792

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. Fixed effects for cohort excluded. [a]Variable constructed from individual data at t-1.

icant coefficient under the discussant version of the model (Model 2d, f, h) [18]. Political interest does not appear to moderate the effect of being undecided, either, in contrast to hypothesis H4b: the three-way interaction of political interest, mean alter support/mean alter vote, and being undecided, is never significant (Model 2g and h).

18 I additionally rerun the analyses using groups of individuals reported as a social group which discusses politics within the cohort, treating co-group membership as a tie. Results with regards to the hypotheses are identical. Finally, separating results by time point or referendum yields too few observations to be meaningfully interpreted (Harrell, 2001, pp. 72-73).

TABLE 4.18: Logit regression of discussant alter support on ego choice of vote

	Model estimate							
	2a	2b	2c	2d	2e	2f	2g	2h
Intercept	0.85**	0.72**	0.73	0.59	0.79**	0.64	0.94	0.36
Mean alter issue support[a]	0.25**	0.16	0.05	-0.09	0.32**	-0.17	0.08	-0.19
Ego support[a]	–	1.35***	–	1.36***	–	1.36***	–	1.38***
Ego knowledge[a]	–	–	0.01	0.06	–	0.09	0	0.01
Ego undecided[a]	–	–	–	–	0.01	0.24	-0.82	-0.1
Ego undec.*mean alter support	–	–	–	–	-0.14	0.14	0.39	0.48
Mean alter support*ego knowl.	–	–	0.06	0.09*	–	0.1*	0.08	0.13*
Political interest	–	–	–	–	–	–	-0.06	0.33
Ego undec.*political interest	–	–	–	–	–	–	0.51	0.17
Mean alter support*political interest	–	–	–	–	–	–	-0.04	-0.07
Mean alter support*ego undec.*political interest	–	–	–	–	–	–	-0.26	-0.2
Accuracy	0.77	0.87	0.79	0.87	0.78	0.86	0.79	0.86
AIC	1010.26	529.53	850.57	531.08	835.07	521.85	841.15	524.31
McFadden's R^2	0.24	0.55	0.27	0.55	0.26	0.55	0.27	0.56
Valid cases	986	846	847	846	831	830	831	830

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. Fixed effects for cohort excluded. For these effects, see Appendix B.4. [a]Variable constructed from individual data at t-1

TABLE 4.20: Logit regression of friend alter support on ego choice of vote

	Model estimate							
	2a	2b	2c	2d	2e	2f	2g	2h
Intercept	0.67***	0.43***	0.58	0.21**	0.59**	0.17**	0.89	0.29
Mean alter issue support[a]	0.4***	0.31**	0.62*	0.69**	0.36**	0.62*	0.7	0.8*
Ego support[a]	–	1.2***	–	1.2***	–	1.19***	–	1.19***
Ego knowledge[a]	–	–	-0.02	0.06	–	0.07	0	0.05
Ego undecided[a]	–	–	–	–	0.02	-0.01	-0.85	-0.62
Ego undec.*mean alter support	–	–	–	–	0.06	0.16	0.43	0.34
Mean alter support*ego knowl.	–	–	-0.07	-0.12	–	-0.11	-0.05	-0.06
Political interest	–	–	–	–	–	–	-0.18	-0.03
Ego undec.*political interest	–	–	–	–	–	–	0.51	0.36
Mean alter support*political interest	–	–	–	–	–	–	-0.09	-0.2
Mean alter support*ego undec.*political interest	–	–	–	–	–	–	-0.28	-0.1
Accuracy	0.76	0.85	0.78	0.86	0.78	0.85	0.78	0.86
AIC	1223.73	688.76	1025.76	689.89	1009.54	678.62	1015.24	683.73
McFadden's R^2	0.24	0.51	0.26	0.51	0.25	0.51	0.26	0.51
Valid cases	1192	1016	1021	1016	1000	996	1000	996

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. Fixed effects for cohort excluded. [a]Variable constructed from individual data at t-1

DISCUSSION

In this study I used new, purpose-collected whole-network longitudinal data to examine whether the choice to vote and the choice of vote are affected by one's political discussant and friend alters, in the unique context of the Swiss direct democracy. I additionally considered a possible differential role of undecided voters and moderating effects of individual attributes of political knowledge and interest.

Regarding choice to vote, I found evidence for an assimilative effect of alters' choice to vote in the short term only amongst friends. Very few individuals were isolated in their opinion at any of the referenda, meaning that one intended hypothesis regarding opinion isolation could not be tested.

Regarding choice of vote, I found evidence that people tend to vote in line with the opinions of their political discussants, but this association disappeared after controlling for pre-test vote intentions. For friends, This effect was consistently found. This is moderate evidence in favour of Hypothesis 2, as will be outlined below. Regarding Hypothesis 3, the critically important undecided were no more likely to be influenced by their alters than others, neither were the politically interested and undecided, failing to support Hypothesis 4a. Similarly, people with more political knowledge were no less likely to be influenced in their vote choice by their alters, in opposition to Hypothesis 4b – indeed, in one case being more likely to show signs of influence.

Several lines of questioning may help to understand the mixed results here, and should be examined in further research. Firstly, why do the results from the friendship network support H1 and H2, while H1 is unsupported and H2 is partially supported by the political discussion network? Potentially, an issue of statistical power is at play here – there are far fewer political discussion ties observed than friendship ties, and therefore fewer cases tested against. Additionally, potential biases in memory of interactions (Bernard & Killworth, 1977; Kashy & Kenny, 1990; Killworth & Bernard, 1976) might mean that the measurement of alters' preferences are more unreliable when using political discussion networks compared to friendship networks.

Second, this issue may also be compounded by political opinions or voting behaviour being driven by unobserved ties outside the network boundaries studied, such as family discussants, or friends external to the cohort. While collecting whole network data is inherently difficult (see e.g. Robins, 2015; Vörös et al., 2021), the inclusion of alters outside the boundary of the cohort studied would be useful to establish whether there are limited effects of discussion, or only limited effects of discussion inside the cohort. That said, given the nature of the intensive study program, the assumption that most of the students' discussion occurs within the confines of the university cohort seems reasonable, and is consistent with the evidence given in analyses using the friendship network.

Third, it is possible that the members of the cohort discuss politics, but not specifically referendum voting or issues. This could be due to a paradox: while referendum issues are subject to more volatile and less deep-seated opinions, possibly permitting stronger influence when discussion occurs, they may be subject to less discussion when people discuss politics for the same reason (Cowan & Baldassarri, 2018). This relates to a broader debate about what exactly we measure when we ask about political discussions (Hopmann, Matthes, & Nir, 2015); quality (e.g. content) and quantity (e.g. duration, frequency) are assumed. If, for instance, one cuts off disagreeing political conversations before they turn into discussions, it may be that one does not report all alters from whom one receives political information when mentioning discussants, but does report the broader category of friends.

Fourth, examing individual characteristics expected to predict differences in turnout and rigidity of beliefs further could be useful. Similarly to the arguments regarding political discussion and referenda, perhaps the nonspecificity of the political interest measure means that individuals are interested in politics in general, but not the issues at hand. This again relates to the nature of referendum votes compared to electoral votes. However, it is unlikely that an effect of knowledge would be found once a null effect of alters' opinions was noted. If people do not appear to be influenced, it is unsurprising that knowledge would not affect how they react to others' discussions. Again, a power issue could be at play, as the required number of cases for an interaction term is much higher than a standard effect.

Limitations

A key limitation of this research sample is that it is relatively unique. While it broadens our knowledge surrounding influence on voting habits, it has its own quirks. It offers some diversity away from ubiquitous US samples, but it remains a WEIRD (white, educated, industrialized, rich, and democratic) population. The latter of these is to some extent a necessity for a study of voting, but extrapolation from this highly educated sample is clearly not simple; for instance, turnout is particularly high amongst those responding to the survey. This aside, the direct democratic voting system in Switzerland is a rarity. However, examining how people vote, rather than simply their self-reported attitudes, helps us to understand both a more consequential form of behaviour which is both impactful on society, as well as a metric that reveals preferences where they may be discordant with beliefs.

A second limitation concerns the statistical analysis; while I can control for the effects of the dependence of alters' opinions in the network with the *y*-permutation test performed here, this test does not handle the dependence of individuals appearing multiple times and the dependence of ties within simultaneous referenda. On the other hand, *y*-permutation tests tend to be somewhat conservative and may lead to the underdetection of true effects (Dekker, Krackhardt, & Snijders, 2007).

Further research

One particularly interesting avenue for further research concerns the tests that could not be made in these data: so few individuals are isolated in their opinion. Bello (2012) and Nir (2011) both use a measure based on participants' perception of alter's preferences, with a limited number of political discussants. This may mean some misestimation of true isolation. Some individuals will not name all of their political discussants due to the limit of the five, four and three-person alter reports that were used. Furthermore, some individuals will falsely believe they are not isolated when they are due to biases in perception and disclosure (Cowan & Baldassarri, 2018; Goel, Mason, & Watts, 2010; Laumann, 1969; Levitan & Visser, 2009). While it is natural to think that it is mere perception of alters' beliefs by an individual that may drive influence, subtler cues that align with alters' true beliefs may too carry weight. To test for these two different flavours of the effects of opinion isolation, and to get a comprehensive view of the networks involved we need better data: comprehensive measures stemming from both ego and alter reports of their beliefs, perceptions of their alters' beliefs, and complete egocentric networks. Additionally, we could use this to test whether at a larger scale alters' beliefs or ego's perceptions of alters' beliefs are more influential.

Effects of a context such as media campaigns that may have differential times of onset but have a potentially powerful impact (Sciarini & Tresch, 2011). This may look like influence where it is actually a shared context effect. Nonetheless, such context effects would be hard to disentangle from social influence: contextual factors (in particular exposure to media campaigns) have some effect precisely via the ties under study. Indeed, this was a key finding in early work on the subject (Katz & Lazarsfeld, 1955). Increased research into the longitudinal pathways by which discussion may indirectly affect individuals' vote choices in these direct democratic referenda would contribute to our understanding of how democracy operates.

Conclusion

Overall, it seems that there is minimal reason to believe that the political discussants in this sample have strongly affected one another; we should not expect communities at this scale to have large impacts on direct democratic outcomes in the short term. Nonetheless, questions remain surrounding the effects of opinion isolation as a perceived or objective phenomenon, and long-term, potentially indirect network effects on voting behaviours.

5 SUMMARIES, CONTRIBUTIONS AND LIMITATIONS

This dissertation concerned political selection and influence in interpersonal networks, and the resulting outcomes. I argued that political opinions are intrinsically linked to our social relations. I used and extended methods for social network analysis, leveraging the strengths of offline social network data to understand selection and influence and their interaction with endogenous network processes. This provided some useful extensions of ways to understand implications for political opinions and voting behaviour from social network ties and vice versa.

Two focal aims were made in this dissertation. Firstly, building on micro-level studies of interpersonal influence and selection in political attitudes and behaviour, and macro-level studies of political polarization, I aimed to build closer bridges between these two types of research and study the evolution of opinions and social networks within a community. Secondly, I aimed to develop further understanding of these processes and polarized outcomes in ways suited to contexts unlikely to meet traditional bipolar understandings of polarization; i.e., in a multiparty political system.

In what follows, I first present a summary of the research chapters. Next, I will summarize the main contributions of this dissertation, followed by its limitations. I highlight avenues for future research that could build on this work, before providing concluding remarks.

5.1 RESEARCH CHAPTER OVERVIEW

5.1.1 Chapter 2

In Chapter 2, a model was presented capturing selection and influence on a variety of political issues embedded in individuals connected by friendship, alongside endogenous network effects. A mixed one-mode and two-mode network approach to describe individuals, their friendships, and their attitudes towards policies in the stochastic actor-oriented model framework was applied, new model effects were developed to represent micro-level processes of selection and influence, as well as latent factors producing homogenizing and polarizing attitudes.

Linking these processes to the macro-level outcome, a new metric of network polarization was introduced. This comprised structures related to these micro-level processes. This metric aimed to capture the level of polarization in the network in a relational sense (whether friends are more likely to be of the same opinions, or non-friends of opposed opinions) and ideological sense (whether agreeing or disagreeing policy attitudes to one policy reflect similarly configured attitudes to another) and a test was devised based on its deviation from a baseline distribution of opinions across policies and individuals.

These new structural model effects, metrics and tests demonstrate a method of micro-macro linkage of selection and influence in a single network model, which has the potential to be generalized across further attitudinal objects. Examining two cohorts of Swiss students, their friendships and their attitudes towards varied political topics over a one-year period, it was found that a modest but significant level of polarization was present in an ideological sense, but it did not clearly increase over time. Furthermore, evidence for selection was found, but only limited evidence for influence was only found, while latent factors consistently structured issue attitudes.

5.1.2 Chapter 3

In Chapter 3, the aims were to understand the link between selection and influence in the model from Chapter 2 and the resulting levels of polarization, while developing the use of the RSiena software as a tool to bind agent-based models to an empirical reality in a more direct way. To achieve these aims, the model from Chapter 2 was taken as a baseline for simulation and a 4x4x3 design was applied to a two-year forward simulation period.

The three variables adjusted in the model included the choice of four sets of effect parameters which were manipulated (selection, influence, or latent convergence and divergence), the sizes of these parameters (multiplied by zero, halved, doubled or quintupled) and variants on how the remainder of the model, including various endogenous effects of the friendship network, was estimated. For this latter variable, three different perspectives on the implications of the model were considered: Firstly, taking the model as approximately correct and simulating from it with only the variable manipulation. Secondly, re-estimating the density (i.e. intercept) of the model, taking a view of there being some set tendencies towards frequencies of friendships and opinions from which one might deviate due to the network context. Thirdly, with full model re-estimation, representing a view that network effects can only be understood in concert with one another.

Overall, many of these models had similar results – a tendency to reduce political opinions, possibly in favour of maintaining social ties in the face of conflicting opinions. Notable exceptions include firstly that, particularly under models with only a manipulation, latent convergent and divergent forces produced very high opinion densities. Most other models tended towards much less saturated sets of opinions, but generally tended towards denser friendship networks. These findings suggest that perhaps the manipulation-only models produced less realistic results than other model types. Secondly, slight differences in the shape and speed of arriving at a stable level of relational polarization appeared when manipulating latent (dis)agreement maximizing processes. Implausible changes also occurred when making them five times as strong as observed, when (partially) re-estimating the underlying model. These results show the predominant insensitivity of the community to processes commonly understood as polarizing.

5.1.3 Chapter 4

Finally, Chapter 4 examined further data on the pre- and post-referendum voting intentions, examining the roles of firstly political discussion ties, and secondarily friendship ties. The aim here was to zoom in on a dyadic behaviour specifically transferring political information, and examine the consequences on a short-term change on the political outcome of voting. Questions were asked of whether and how individuals were influenced in their behaviour in the short term, on issues of only brief temporal salience. More precisely, it was examined to what extent individuals' political discussion networks showed assimilative influence on their choice to vote, and their choice of vote on 9 referenda occurring in Switzerland. As an alternative measure, friendship was also examined due to limits to memory of discussion and events and questions surrounding the interpretation of the applied political discussion measure.

A correlation between alters' choice to vote with ego's, and of alters' attitudes towards an issue and ego's choice of vote were found in the political discussion network. However, these associations disappeared after controlling for their prior planned vote. This hints at selection processes being a more likely candidate for causing this similarity between individuals than influence – possibly due to the friendships in which political discussions were typically nested being selected for political similarities, as found in Chapter 2. Alternatively, influence may have occurred on a longer time frame, but this seems unlikely in the context of these issues which often attract relatively little public engagement (LeDuc, 2007). On the other hand, friendship alters' intention to vote and attitudes towards a political issue positively predicted ego's choice to and of vote – possibly this discrepancy is due to known memory effects (Bernard & Killworth, 1977; Killworth & Bernard, 1976), ambiguity in interpretations of the network measuring question (Hopmann, Matthes, & Nir, 2015), differences in confounding variables in the two networks (Shalizi & Thomas, 2011), or the relative sparsity and thereby power of the statistical analysis in the case of the discussion compared to the friendship network.

A hypothesis on the effects of being an attitude isolate (i.e. being alone in one's opinion in one's ego network) could not be tested, due to the infrequency with which individuals were indeed isolated in their opinions. Although not directly testable, this highlights a potential gap between people's perceptions and their true opinion environment when contrasted with previous studies (Bello, 2012; Nir, 2011). Partially this may occur due to a smaller sample size in the data applied in this dissertation, but also due to measurement – the previous studies asked about ego's perceptions of agreement or disagreement, rather than the state of agreement or disagreement constructed from self-reports of opinions. Comparing these two could yield further insights.

5.2 CONTRIBUTIONS

In this dissertation, I made scientific contributions on multiple fronts. In particular, I contributed to the development of micro-macro linking of networks and political opinions, considered additional multivariable approaches to the empirical study of political behaviour, and extended methods for developing realistic agent-based models. In terms of empirical information, I note that selection appears more consistently than influence, but there is some evidence for both processes – possibly affecting voting behaviour.

5.2.1 Micro-macro framework for networks and political opinions

I built on prior work describing macro-level, societal ideological and affective political polarization by examining polarization as a networked phenomenon built out of social micro processes. A key contribution here is in bridging these processes and outcomes in a single empirical framework. With my coauthors, we defined new measures of polarization by representing friendships between individuals and their attitudes towards political issues as a mixed one-mode and two-mode network. We modelled relevant processes in the stochastic actor-oriented model, and due to the longitudinal framework were able to disentangle selection and influence on political attitudes in two university communities. In doing so four new model terms were contributed to the RSiena software in which the stochastic actor-oriented model is implemented.

5.2.2 Multivariable approach

I examined a variety of issues; 22 topics and 9 referendum votes. Given the complexities of multiparty systems such as that occurring in Switzerland, accounting for multiple issues in examining selection and influence may help draw closer to reflected complexities of identities and interactions where a more bipolar perspective focused on two-party political systems falls short. In addition, these issues were selected on potential relevance rather than prior knowledge of more extreme and/or deep-seated preferences such as party voting, showing how attitudes about relatively less important issues can affect and be affected by our networks. In both empirical Chapters 2 and 4, I included both ego and alter reports of political attitudes, providing one of few empirical examples where disagreement can be observed absent biases of projection and selective disclosure occurring in interactions (Cowan & Baldassarri, 2018; Goel, Mason, & Watts, 2010; Huckfeldt & Sprague, 1995; Laumann, 1969).

5.2.3 Realism of agent-based models

I extended the strengths of agent-based modelling in understanding how micro-processes occurring between individuals result in societal outcomes, contributing to the literature by responding to calls for better empirical grounding (e.g. Baldassarri & Page, 2021; Boero & Squazzoni, 2005; Mäs, 2019; Sobkowicz, 2009; Stadtfeld, Takács, & Vörös, 2020). In particular, the use of an empirical data set in calibrating and validating an agent-based model allowed the demonstration of a method to tackle problems of time, theoretical tie specification, and accounting for dynamic and endogenous social networks, applying these to a model of opinion dynamics. Doing so helped to identify that the studied community would rarely polarize in a two-year time period, with only increases in latent tendencies towards (dis)agreement producing more polarization than other conditions, but still trending downwards.

5.2.4 Selection is more visible than influence

In both Chapters 2 and 4, primarily evidence of selection was found, whereas evidence for influence was much more limited – indeed, not all that much change was present in many attitudes and behaviours. This contrasts with e.g. Lazer et al. (2010), who had opposite results when examining a single item measure of ideology in a group of students particularly engaged in public policy, but is consistent with a similar cohort studied in a similar manner by Wang, Lizardo, and Hachen (2020). This is an interesting point for agent-based modellers who primarily assume influence, with selection being indirectly implemented via e.g. bounded confidence models which assume influence stops beyond a certain level of difference (Deffuant et al., 2000; Hegselmann, Krause, et al., 2002).

Notably, the fewer observations of issues in the two-mode network coupled with the relatively low amount of change amongst them may imply lower power to detect influence effects as compared to selection effects amongst the much higher number of individuals (Stadtfeld et al., 2020), though it could be argued that tie change is particularly high in the emerging networks studied in the data I applied in this dissertation.

The more reliable appearance of selection than influence remains intriguing, suggesting that further development of a dynamic network-based approach within the agent-based modelling context would be an important improvement.

5.3 LIMITATIONS

The chapters presented are also subject to a number of limitations; primarily these relate to the unobserved qualities of ties and information channelled through them, limited accounting for node attributes (i.e. individual and issue-specific variability of effects), and the case-study nature of the analyses presented.

5.3.1 Tie content and quality

The content of the interactions between individuals cannot be captured with the data used. Particularly in the case of political discussion ties, this presents difficulties in knowing the effects of discussions (but not of discussion partners). If one does not talk about the topic at hand, it is less likely that influence would occur (outside of heuristic markers of opinions that may reveal themselves in other day-to-day interactions, as implied by DellaPosta et al., 2015; Goldberg & Stein, 2018 and Sokhey & Djupe, 2011).

Similarly, qualities of the ties are not included, such as frequency or duration of interactions, or strength of the friendship ties, which may also provide variation in the amount of influence. The friendship ties considered in Chapters 2, 3, and 4 may channel influence, but there are likely a number of mediators (including political discussions as events), and moderators (such as the level of trust between friends, or dyadic perceptions of knowledge) in these relations which could be of importance. A deeper, multiplex perspective on social ties may be useful in exploring this issue of quality further (Vörös & Snijders, 2017).

5.3.2 Issues, individuals and variance of effects

The issues considered in Chapter 2, and votes in Chapter 4, obviously do not cover the span of political issues and votes that could have been selected. Amongst these, there is almost certainly some variance in the extent to which selection and influence occur. Some issues are likely to be more salient than others, and polarization, for instance, may occur at a greater rate around such issues (Baldassarri & Bearman, 2007). Similarly, between individuals these effects may vary too, with some individuals being more influential, some more likely to be influenced, and some selecting more strongly depending on their individual characteristics (Bello & Rolfe, 2014; Boutyline & Willer, 2017; Hovland & Weiss, 1951; Katz & Lazarsfeld, 1955). Incorporating explanations of this variance, especially via the inclusion of additional variables, would be a substantial improvement to the explanatory power of the models presented, but would likely require much more data to estimate

5.3.3 Case study

Finally, whole social network studies such as these are necessarily case studies, as they do not feature a random sampling of individuals but an entire bounded network. This is amplified by the fact that the two empirical studies in Chapters 2 and 4 overlap in two of the three cohorts examined. Therefore, generalizability is questionable if variables important for the communities studied differ from other populations.

5.4 FURTHER RESEARCH

Interesting avenues for further research include the further inclusion of identity and variables associated with ideology, incorporation of more nuanced forms of influence, and effects of different contexts.

5.4.1 Political identities in multiparty systems

A key contribution that I highlighted related to the use of multivariable attitudes in considering polarization. While ideology is a core part of polarization (Rogowski & Sutherland, 2016; Webster & Abramowitz, 2017), identity-based approaches are extremely relevant, particularly when it comes to affective conceptualizations of polarization. Differences in evaluations of ingroups and outgroups used in the definition of affective polarization are rooted in the identity-based perspectives from social psychology (see e.g. Brewer, 2017; Druckman et al., 2020; Tajfel, 1981). This, of course, is more complicated in measurement and conceptualization outside of the two-party system of the United States, or the near two-party system of the United Kingdom (Reiljan, 2020; Wagner, 2021). Understanding how the sets of political attitudes relate to political identities, and how these identities are constructed and relate to one another would be a difficult but potentially valuable task in understanding the levels and causes of polarization; in particular where these identities may mediate between ideology and effects on and of social ties, as is implied by work on policy versus identity and affective polarization (Dias & Lelkes, 2022). I expect that the structure of these identities may be quite complicated: as noted by Zuckerman and Kroh (2006) individuals may have affiliations to specific segments of the political party system, within which they choose. Are political identities in these complicated systems similarly hierarchically nested? How much do they overlap? These and many other identity-related questions should be examined in future research on socio-political dynamics (Huddy, 2001).

5.4.2 Extension to other cultural preferences

The emergence of cultural clusters is a common theme in studies of political divides both in empirical and theoretical settings (DellaPosta, Shi, & Macy, 2015; Flache et al., 2017; Goldberg & Stein, 2018; Mäs et al., 2013). Extension of the variables used in the models presented in chapters 2 and 3 of this dissertation to include non-political cultural markers, that may provide heuristic cues about one's ideology, may help to separate out effects of explicit discussion and indirect observations. These could help to answer questions such as whether and which sets of similar features may indicate political points of view, like 'latte drinking' and 'bird hunting' are said to in the United States (DellaPosta, Shi, & Macy, 2015), and whether these have consequences for polarization in a relational or ideological sense. Interviews digging into people's perceptions around these kinds of variables may be useful as a precursor to such a study, although this cannot inform us about variables that will later come to be affiliated with political ideology.

5.4.3 Complexities of influence in observational contexts

In the current projects, I focused on simple conceptualizations of assimilative, dyadic social influence, based on simple di- or trichotomized attitudes or voting. More nuanced forms could be incorporated both in terms of using less coarse data treatment, and in terms of distinguishing different functional forms of influence – how is an individual affected by another individual who has a very strong positive attitude, as compared to various individuals with milder attitudes of a uniform direction? How does this affect the influence of these individuals on the overall network? Are some individuals becoming more extreme in some attitudes in response to others, rather than assimilating to their possibly moderate point of view? Under the current framework, these questions are difficult to answer. Though social psychology has already extensively examined conditions for influence, it has focused on experimental work which does not necessarily provide ecological validity (Turner, 1991).

Beyond assimilative and/or extremetizing forces, repulsive forces are necessary for polarization in simulated opinion models (Flache et al., 2017). These have been observed in a large-scale online context (Bail et al., 2018), but are not as easily observed in offline interpersonal tests (Takács, Flache, & Mäs, 2016) in line with theories of minimal but consistent change in some opinion domains (Levendusky, 2009). To what extent do these possible effects exist, and when do they generalize to observational contexts where selection is possible and bounded similarity effects may take hold (Deffuant et al., 2000; Hegselmann, Krause, et al., 2002)?

With increasing data and model capabilities these features may be explored further with potential consequences for the micro and macro links presented. Combining the sense of real time given in the approach presented in Chapters 2 and 3 with varied functional forms of influence we could also, for instance, account for the speed of changes in macro-level outcomes. Accounting for the effects of biased opinion sharing while acknowledging that selective disclosure of an unpopular attitude may have different effects due to heuristic cues would greatly improve the realism of the models presented.

5.4.4 Relative consequences online and offline

A comparison of online and offline contexts could be useful, particularly in understanding where interventions may be targeted to reduce polarization where it is problematic. As noted in the introduction, offline relations are stronger than purely online ones (Antheunis, Valkenburg, & Peter, 2012; Mesch & Talmud, 2007), and stronger influence on political opinions could occur via known than unknown individuals (Weeks & Gil de Zúñiga, 2021), but the scale of the internet may mean that its ultimate effects are larger. How much are people channeling information and opinions discovered from online media via their in-person relationships? How does this compare to impersonal mass communications over similar media, and how does the context of observation (i.e. on the platform versus in-person discussion) affect influence? These questions linking the perspective presented in this dissertation and the common online focus on political polarization's causes could be fruitful avenues for further research.

Extension of the network models presented in this dissertation to online contexts, applying digital, relational event data such as interaction events between individuals combined with survey-based observation to understand relational ties and relational events (Butts, 2008; Stadtfeld & Block, 2017) may be fruitful. This would enable us to understand the consequences of relations that vary in the extent to which they are offline and online in interplay with political opinions. The structural feature-based approach considered in Chapters 2 and 3 – focusing on triads of individuals and issues, connected by social ties and attitudes respectively, or tetrads of two individuals and two issues connected by attitudes – extend with relative ease to online communication data. Combining digital communication events with survey opinions, or where reliable, attitudes constructed from observations on the digital communication platforms, may allow for comparison of online, offline, and mixed effects using appropriate frameworks analogous to the SAOM framework used in this dissertation, such as dynamic network actor models (Stadtfeld & Block, 2017). In such a framework, communication event frequency and relationships such as friendship could be used to extend to weighted networks, and interaction terms to account for online and offline ties would help compare different connections' effects. More efficient testing methods for the proposed polarization index should then be developed, as permutation-based testing requires substantial computational power which increases non-linearly with network size.

5.5 IN CLOSING

Political attitudes and social networks are intertwined. This dissertation acknowledged this, and examined the reciprocal effects between the two, developing new methods and concepts to build our understanding of societally relevant outcomes such as voting and polarization. Overall, findings presented in this dissertation do not suggest dramatic effects of selection and influence on polarization in the short term – for which we can be thankful, and perhaps not too surprised: one does not regularly observe intense ideological conflicts within single cohorts, though protests involving violence have been known to occur. Neither do the current results suggest strong effects of discussants and friends on vote choices. Whether this is good or bad is of course entirely dependent on the qualities of information transmitted by one's alters. Do they attempt to persuade in their own favour, or inform in the name of allowing their contacts to make a choice reflecting their own values? In the former case, it is desirable that they are unsuccessful.

How consequential processes are that may contribute to polarization, to what extent they are present, and which methods should be applied to understand them, are all questions tackled in this dissertation. More can and should be done to understand how and when socio-political fractures may occur and how they impact us.

A

APPENDIX FOR CHAPTER 2

A.1 STRUCTURES USED IN THE POLARIZATION DEFINITION AND MODEL

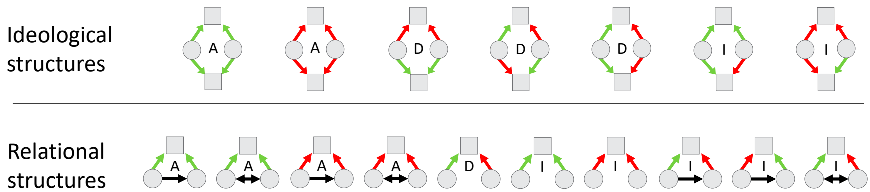


FIGURE A.1: Structures used in defining polarization. Four-cycles (upper half) are used in defining ideological, triads in the lower half are used in defining relational polarization. "A" and "D" indicate agreeing and disagreeing structures, consistent with polarization. "I" indicates structures which are inconsistent with our concept of polarization.

A.2 NETWORK POLARIZATION CALCULATION

Using the networks of bipartite attitudes and people, and people amongst one another, we define the metrics capturing network polarization. Figure 2.3 shows stylized forms of both a non-polarized network, and a polarized network under our metrics. These metrics are constructed from specific network structures - the census of mixed triads and the census of mixed, complete four-cycles of people, friendships, political statements and attitudes. Censuses of network structures have a relatively long history in the social networks literature both in descriptive and in inferential analyses (for seminal publications, see J. A. Davis, 1970; Holland & Leinhardt, 1970). We thus take a census of the possible configurations of these complete four-cycles and triads, and examine whether their prevalence relative to one another would suggest polarization in our data.

To do this, we also define an expectation. In the ideological aspect of polarization, we define this expectation from a sample of synthetic opinion networks with the same degree distribution (i.e. all nodes have the same degree) as the observation. To find these networks, we perform a rewiring procedure. Step by step, this occurs as follows:

1. Select a random individual *a1* (weighted by number of ties)
2. Select a random attitudinal object *b1* it is connected to
3. Check the valence of the tie (positive or negative)
4. Select a random connected individual-attitude pair *a2b2* of which neither is connected to the individual-attitude pair from steps 1 and 2, with the same valence as tie *a1b1*
5. Swap the ties so that one goes from a1 to b2, and the other goes from a2 to b1
6. Repeat for 3 times the number of ties observed in both the positive and negative network

For the relational aspect of polarization we take a simpler approach, and simultaneously permute the rows and columns of the friendship adjacency matrix, while holding the opinion network constant. This implies an identical structure of the friendship network, but a changed relationship between the opinion and friendship network.

We generate 1000 synthetic opinion networks, and 1000 synthetic friendship networks using this procedure. All code to generate these is available on the github repository: [omitted for review, available on request] Functions 'rewire3' and 'permuteNet' generate the networks following the procedures above.

To capture the ideological facet of polarization, defined in Formula A.1 we count four-cycles of opinions, specifically we take a census of complete tetrads containing two individuals and two topics. Only opinion ties are considered, while friendship ties are ignored in these configurations. For both the observed and synthetic opinion networks, we calculate statistics given below.

i and j are individuals where $i \neq j$, k and m where $k \neq m$ are political issues, and w_+ and w_- indicate positive or negative opinion ties between individuals and political issues.

$$s_{\text{agree}} = \sum w_{ik_+} w_{jk_+} w_{im_+} w_{jm_+} + \sum w_{ik_-} w_{jk_-} w_{im_-} w_{jm_-} + \sum w_{ik_+} w_{jk_+} w_{im_-} w_{jm_-} \tag{A.1a}$$

defines the count of tetrads where individuals completely agree, while

$$s_{\text{disagree}} = \sum w_{ik_+} w_{jk_-} w_{im_+} w_{jm_-} + \sum w_{ik_+} w_{jk_-} w_{im_-} w_{jm_+} \tag{A.1b}$$

defines the count of tetrads where individuals completely disagree. Finally,

$$s_{\text{inconsistent}} = \sum w_{ik_+} w_{jk_+} w_{im_+} w_{jm_-} + \sum w_{ik_-} w_{jk_-} w_{im_-} w_{jm_+} \tag{A.1c}$$

defines the count of tetrads where individuals agree on one issue but disagree on another.

Each term is then normalized, by dividing by the expectation. This is the mean of each statistic across the 1000 synthetic networks described above. For example, s_{agree} is divided by $E(s_{agree})$.

After normalization, we define each metric as follows:

$$s_{\text{ideological attraction}} = \frac{s_{\text{agree}}}{s_{\text{agree}} + s_{\text{inconsistent}}} \tag{A.1d}$$

$$s_{\text{ideological repulsion}} = \frac{s_{\text{disagree}}}{s_{\text{disagree}} + s_{\text{inconsistent}}} \tag{A.1e}$$

$$s_{\text{ideological polarization}} = \frac{s_{\text{ideological attraction}} + s_{\text{ideological repulsion}}}{2} \tag{A.1f}$$

To capture the social facet of polarization, represented in Formula A.2, we start with the census of triads containing two people, one attitudinal object, and two opinion ties. We then calculate the following statistics for both the observed network of friendships and opinions, and the combined observed opinions with 1000 permuted friendship networks.

x_{ij} represents the binary (0 for no tie, 1 for a tie) value of directed friendships between individuals ij, where $i \neq j$. w_{k_+} represents positive attitudinal ties to issue k, and w_{k_-} represents negative attitudinal ties to issue k.

$$s_{\text{friends agree}} = \sum x_{ij} w_{ik_+} w_{jk_+} + \sum x_{ij} w_{ik_-} w_{jk_-} \tag{A.2a}$$

defines the count of triads of friends with valenced agreement on a topic,

$$s_{\text{friends disagree}} = \sum x_{ij} w_{ik_-} w_{jk_+} + \sum x_{ij} w_{ik_+} w_{jk_-} \tag{A.2b}$$

defines the count of triads of friends with valenced disagreement on a topic,

$$s_{\text{nonfriends disagree}} = \sum -1 * (x_{ij} - 1) w_{ik_-} w_{jk_+} + \sum -1 * (x_{ij} - 1) w_{ik_+} w_{jk_-} \tag{A.2c}$$

defines the count of triads of unconnected individuals with valenced disagreement on a topic

$$s_{\text{nonfriends agree}} = \sum -1 * (x_{ij} - 1) w_{ik_+} w_{jk_+} + \sum -1 * (x_{ij} - 1) w_{ik_-} w_{jk_-} \tag{A.2d}$$

defines the count of triads of unconnected individuals with valenced agreement on a topic.

We normalize all statistics by the expectation, as above. Then, we calculate the relational polarization indices as follows:

$$s_{\text{social attraction}} = \frac{s_{\text{friends agree}}}{s_{\text{friends agree}} + s_{\text{nonfriends agree}}} \tag{A.2e}$$

$$s_{\text{social repulsion}} = \frac{s_{\text{nonfriends disagree}}}{s_{\text{nonfriends disagree}} + s_{\text{friends disagree}}} \tag{A.2f}$$

$$s_{\text{relational polarization}} = \frac{s_{\text{social attraction}} + s_{\text{social repulsion}}}{2} \tag{A.2g}$$

Both metrics theoretically range from zero to one. It has not been explored yet to what extent the structure of the observed networks may affect the empirically possible range of the metrics.

A.3 POLITICAL STATEMENT BATTERY

1. There should be an annual upper limit to the uptake of new asylum seekers.
2. Switzerland should be able to set quotas for the immigration of foreign workers.
3. Operators of internet sites should be legally mandated to remove Fake News that they are made aware of.
4. Children must be vaccinated against contagious illnesses.
5. All banks should be nationalized.
6. There should be a quota for the number of women on supervisory boards of listed companies.
7. High wealth should be taxed.
8. Arms exports from Switzerland should be prohibited, without exception.
9. The regulated sale of cannabis should be generally permitted.
10. Bank customer secrecy privileges in Switzerland should be loosened to combat tax evasion.
11. The state should be allowed to invade the privacy of citizens more forcefully to protect against terrorism.
12. Video surveillance in public spaces should be expanded.
13. Radio and TV license fees should be abolished and replaced by commerical financing.
14. The expansion of renewable energy sources should receive permanent financial support from the state.
15. Electromobility should be promoted more than conventional mobility technologies.

16. More bike paths should be built on public roads in Switzerland.
17. Organic agriculture should be promoted more than conventional agriculture.
18. Food should only be allowed to be imported if it meets Swiss standards of production regarding sustainability, animal rights, and workers' rights.
19. There should be a universal basic income in Switzerland.
20. The state should offer more resources for the construction of social housing.
21. There should be an upper limit to foreign students at Swiss universities.
22. University studies should be strictly free of charge.

A.4 POLITICAL STATEMENT DESCRIPTIVES

TABLE A.1: Cohort 1, means, standard deviations, and valid responses to policy items

	Wave 1			Wave 5		
Item nr.	N	M	SD	N	M	SD
1	154	3.35	1.95	118	3.64	1.85
2	154	3.62	1.76	118	3.67	1.79
3	154	5.60	1.51	118	5.42	1.57
4	154	5.41	1.67	118	5.51	1.63
5	154	3.29	1.52	118	3.19	1.49
6	154	2.91	1.62	118	2.94	1.66
7	153	5.44	1.43	118	5.45	1.36
8	154	4.32	1.76	118	4.34	1.73
9	153	4.98	1.82	118	4.74	1.81
10	154	4.23	1.57	118	4.11	1.55
11	154	3.13	1.53	118	3.27	1.53
12	153	3.44	1.63	118	3.41	1.64
13	154	3.68	1.68	118	2.69	1.58
14	154	5.86	1.27	118	5.76	1.30
15	154	5.51	1.30	118	5.40	1.33
16	153	5.39	1.38	118	5.28	1.42
17	154	5.36	1.26	118	5.32	1.31
18	154	5.34	1.23	118	5.45	1.17
19	154	3.88	1.89	118	3.81	1.82
20	154	4.31	1.42	118	4.62	1.47
21	154	2.44	1.46	118	2.58	1.71
22	154	4.69	1.75	118	4.76	1.77

Note. Ranges from 1 (strongly disagree) to 7 (strongly agree)

TABLE A.2: Cohort 2, means, standard deviations, and valid responses to policy items

	Wave 1			Wave 5		
Item nr.	N	M	SD	N	M	SD
1	393	3.77	1.83	272	4.00	1.73
2	391	3.83	1.59	272	3.93	1.61
3	393	5.50	1.63	272	5.18	1.75
4	392	5.56	1.56	272	5.38	1.61
5	392	2.84	1.53	271	2.79	1.43
6	393	2.89	1.58	270	2.77	1.49
7	393	4.82	1.64	272	4.89	1.50
8	390	3.79	1.84	272	3.80	1.69
9	393	4.77	1.86	272	4.77	1.76
10	393	3.94	1.56	272	3.77	1.48
11	391	3.25	1.67	272	3.38	1.66
12	392	3.65	1.71	271	3.54	1.70
13	392	3.62	1.68	272	3.24	1.69
14	391	5.84	1.36	272	5.62	1.41
15	393	5.47	1.52	272	5.30	1.48
16	393	5.21	1.54	272	5.18	1.51
17	392	5.20	1.32	272	5.12	1.35
18	393	5.11	1.45	272	5.16	1.43
19	393	3.40	1.86	273	3.32	1.74
20	391	4.19	1.41	273	4.26	1.38
21	392	2.59	1.61	273	3.01	1.59
22	393	4.72	1.74	273	4.56	1.80

Note. Ranges from 1 (strongly disagree) to 7 (strongly agree)

A.5 ROBUSTNESS CHECKS

A.5.1 Model results

While we believe the best model and data specification is the main model presented, we conducted five robustness checks. Results from all of these models can be found in tables below, followed by summaries.

In the text that follows, we refer to these models by codes R1-R5. Four of the alternative specifications were based on data modifications:

First, since our data extended beyond the one-year period used in the main analysis, we additionally tested until the end of the students' second year of their undergraduate studies (R1; i.e. nine waves of data). This was not included in the main analysis because of substantial change in the size and particularly participation of the cohorts after the end of the first year. Participation dropped from 60% at Wave 5 to 48% at Wave 6 for Cohort 1, and from 48% to 36% in the same period for Cohort 2 (Vörös et al., 2021).

In a second test, we estimated the model for the first to second observation (R2). The first wave of data had been initially excluded due to difficulties it caused in estimating the model, likely due to the high instability of the friendship network from Wave 1 to 2 in both cohorts.

Third, we estimated the model with a response to an attitude item of "no opinion" as missing, rather than a neutral attitude (R3). Notably, this model misfit significantly on the four-cycles examining the ideological part of polarization (p = .037) for cohort two (tending to generate more four-cycles than observed).

Fourth, we estimated models with different thresholds for coding ties in the attitude network (R4). We considered the alternatives one point away from the neutral option on the scale (i.e. a less extreme threshold) and those at the outer ends of the scales (i.e. a more extreme threshold). The former of these two models did not converge, likely due to minimal tie changes – most people who held a valenced attitude on a topic either maintained an attitude of that valence, becoming less extreme or becoming neutral, rather than switching to the other side.

Lastly, focusing on alternative model specification, and in line with previous literature (Lazer et al., 2010; Wang, Lizardo, & Hachen, 2020), we considered the effect of political orientation as a univariate dimension predicting social ties which may correlate with other political attitudes (R5). We thus included a homophily term for a simple left-right political orientation scale.

Full Stochastic Actor Oriented Model parameter estimates.

	Main, C1	**Main, C2**	**Waves 2-9, C1**	**Waves 2-9, C2**	**Waves 1-2, C1**	**Waves 1-2, C2**	**No opinion NA, C1**	**No opinion NA, C2**	**High thresh., C1**	**High thresh., C2**	**Political. or., C1**	**Political. or., C2**
Rate friend W1-W2					11.12***	8.86***						
					(1.51)	(0.83)						
Rate friend W2-W3	6.87***	5.98***	6.84***	6.01***			6.93***	5.96***	6.89***	5.95***	6.85***	5.96***
	(0.55)	(0.36)	(0.69)	(0.46)			(0.55)	(0.41)	(0.55)	(0.35)	(0.60)	(0.42)
Rate friend W3-W4	6.97***	6.48***	7.02***	6.65***			6.94***	6.40***	7.01***	6.54***	6.97***	6.47***
	(0.58)	(0.36)	(0.53)	(0.32)			(0.57)	(0.35)	(0.65)	(0.38)	(0.54)	(0.38)
Rate friend W4-W5	4.30***	4.94***	4.33***	5.10***			4.25***	4.89***	4.26***	5.02***	4.29***	4.95***
	(0.37)	(0.28)	(0.36)	(0.33)			(0.37)	(0.33)	(0.40)	(0.27)	(0.42)	(0.32)
Rate friend W5-W6			3.71***	4.68***								
			(0.41)	(0.38)								
Rate friend W6-W7			8.46***	7.24***								
			(2.33)	(0.90)								
Rate friend W7-W8			3.00***	3.53***								
			(0.40)	(0.33)								
Rate friend W8-W9			1.95***	2.50***								
			(0.27)	(0.23)								
Density friend (intercept)	-3.10***	-3.57***	-3.08***	-3.39***	-3.23***	-2.60***	-3.15***	-3.61***	-3.08***	-3.50***	-3.12***	-3.57***
	(0.15)	(0.10)	(0.13)	(0.09)	(0.42)	(0.28)	(0.15)	(0.11)	(0.14)	(0.10)	(0.14)	(0.11)
friend: reciprocity	3.49***	3.85***	3.42***	3.94***	3.92***	4.49***	3.48***	3.85***	3.54***	3.87***	3.48***	3.85***
	(0.18)	(0.14)	(0.16)	(0.13)	(0.40)	(0.32)	(0.19)	(0.14)	(0.18)	(0.14)	(0.18)	(0.15)
friend: transitive triplets	0.81***	1.22***	0.85***	1.20***	1.27***	1.63***	0.81***	1.23***	0.82***	1.22***	0.81***	1.22***
	(0.05)	(0.04)	(0.05)	(0.04)	(0.19)	(0.12)	(0.06)	(0.04)	(0.05)	(0.04)	(0.06)	(0.04)
friend: transitive recipr. triplets	-0.62***	-1.12***	-0.64***	-1.11***	-0.77$^{\cdot}$	-1.39***	-0.62***	-1.12***	-0.64***	-1.14***	-0.62***	-1.12***
	(0.09)	(0.07)	(0.09)	(0.06)	(0.46)	(0.15)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)
friend: indegree - popularity	0.07**	0.06***	0.07***	0.06***	0.07	0.07	0.07*	0.06***	0.07*	0.06***	0.07**	0.06***
	(0.03)	(0.01)	(0.02)	(0.01)	(0.04)	(0.05)	(0.03)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)
friend: outdegree - popularity	-0.18***	-0.20***	-0.19***	-0.20***	-0.25***	-0.23*	-0.18***	-0.20***	-0.17***	-0.18***	-0.17***	-0.20***
	(0.03)	(0.02)	(0.03)	(0.02)	(0.06)	(0.10)	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)	(0.02)
friend: outdegree - activity	-0.02*	-0.01*	-0.02**	-0.02***	-0.01	-0.08*	-0.02*	-0.01*	-0.02*	-0.01*	-0.02**	-0.01*

	(0.01)	(0.01)	(0.01)	(0.00)	(0.03)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
friend: anti-isolate	-1.70***	-1.41***	-1.27***	-1.21**	-1.96*	-2.52**	-1.69***	-1.41***	-1.73***	-1.45***	-1.68***	-1.42***
	(0.39)	(0.35)	(0.35)	(0.37)	(0.77)	(0.97)	(0.44)	(0.42)	(0.41)	(0.31)	(0.34)	(0.32)
friend: language	0.23**	0.33***	0.27***	0.33***	0.35**	0.06	0.23**	0.33***	0.24**	0.33***	0.23**	0.32***
	(0.08)	(0.06)	(0.07)	(0.05)	(0.12)	(0.09)	(0.08)	(0.06)	(0.08)	(0.06)	(0.08)	(0.06)
friend: same major	0.50***		0.43***		0.81***		0.50***		0.50***		0.50***	
	(0.06)		(0.05)		(0.09)		(0.06)		(0.07)		(0.07)	
friend: gender alter	0.05	0.27***	0.10·	0.28***	0.06	0.15	0.05	0.28***	0.08	0.26***	0.05	0.27***
	(0.07)	(0.07)	(0.06)	(0.06)	(0.11)	(0.14)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)
friend: gender ego	-0.16·	0.49***	-0.19**	0.36***	-0.39*	0.03	-0.16·	0.50***	-0.16·	0.48***	-0.16·	0.49***
	(0.08)	(0.08)	(0.07)	(0.07)	(0.16)	(0.24)	(0.09)	(0.08)	(0.08)	(0.07)	(0.09)	(0.07)
friend: same gender	0.15*	0.42***	0.19***	0.29***	0.21*	-0.10	0.16*	0.42***	0.17*	0.42***	0.15*	0.42***
	(0.07)	(0.07)	(0.06)	(0.06)	(0.10)	(0.14)	(0.07)	(0.07)	(0.07)	(0.07)	(0.06)	(0.07)
friend: similar pol. orientation											0.20	-0.05
											(0.24)	(0.19)
friend: birthyear similarity	0.44	0.45*	0.55*	0.24	0.68	0.23	0.48	0.45·	0.53·	0.40·	0.45	0.45·
	(0.29)	(0.22)	(0.25)	(0.18)	(0.43)	(0.32)	(0.30)	(0.24)	(0.28)	(0.24)	(0.29)	(0.24)
friend: neg.att-pos.att two-paths[a]	-0.05	0.05	-0.03	-0.00	-0.09	-0.06	-0.03	0.03	-0.14	0.13	-0.05	0.05
	(0.05)	(0.04)	(0.05)	(0.04)	(0.08)	(0.07)	(0.05)	(0.04)	(0.21)	(0.09)	(0.06)	(0.04)
friend: pos.att-neg.att two-paths[a]	-0.04	0.03	-0.01	0.04	0.07	0.07	-0.01	0.03	0.03	0.02	-0.03	0.03
	(0.05)	(0.03)	(0.04)	(0.03)	(0.08)	(0.05)	(0.05)	(0.04)	(0.15)	(0.12)	(0.06)	(0.03)
friend: pos.att agreement	0.05*	0.02	0.04*	0.03	0.08*	0.02	0.06**	0.03	0.14**	-0.06	0.06*	0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.05)	(0.02)	(0.03)	(0.05)	(0.06)	(0.02)	(0.02)
friend: neg.att agreement	0.07*	0.06*	0.05·	0.06*	0.04	0.06	0.05	0.09**	0.03	0.09	0.07·	0.06*
	(0.03)	(0.03)	(0.03)	(0.03)	(0.05)	(0.04)	(0.03)	(0.03)	(0.07)	(0.06)	(0.04)	(0.03)
Rate pos W1-W2					6.60***	7.42***						
					(0.39)	(0.50)						
Rate neg W1-W2					4.71***	4.97***						
					(0.38)	(0.30)						
Rate pos W2-W3	6.30***	7.32***	6.36***	7.36***			5.20***	6.46***	4.88***	5.68***	6.29***	7.32***
	(0.38)	(0.33)	(0.41)	(0.33)			(0.38)	(0.36)	(0.45)	(0.38)	(0.40)	(0.32)
Rate neg W2-W3	4.79***	5.43***	4.82***	5.42***			4.67***	4.90***	3.30***	4.10***	4.78***	5.42***
	(0.35)	(0.27)	(0.33)	(0.35)			(0.32)	(0.32)	(0.35)	(0.32)	(0.32)	(0.28)

Rate pos W3-W4	7.37***	7.12***	7.39***	7.12***			6.00***	6.08***	5.97***	5.75***	7.35***	7.13***
	(0.46)	(0.30)	(0.49)	(0.35)			(0.46)	(0.29)	(0.47)	(0.39)	(0.49)	(0.28)
Rate neg W3-W4	4.97***	5.46***	5.02***	5.47***			4.75***	5.49***	2.95***	4.68***	4.97***	5.48***
	(0.39)	(0.27)	(0.44)	(0.24)			(0.38)	(0.31)	(0.34)	(0.31)	(0.41)	(0.27)
Rate pos W4-W5	6.52***	6.73***	6.56***	6.76***			5.62***	6.19***	5.07***	4.67***	6.53***	6.72***
	(0.43)	(0.29)	(0.44)	(0.42)			(0.41)	(0.30)	(0.41)	(0.35)	(0.49)	(0.29)
Rate neg W4-W5	3.55***	4.31***	3.59***	4.37***			3.49***	4.49***	2.78***	3.38***	3.55***	4.31***
	(0.28)	(0.22)	(0.29)	(0.21)			(0.30)	(0.30)	(0.24)	(0.26)	(0.27)	(0.19)
Rate pos W5-W6			5.50***	5.45***								
			(0.50)	(0.29)								
Rate neg W5-W6			3.98***	4.14***								
			(0.34)	(0.28)								
Rate pos W6-W7			4.81***	5.18***								
			(0.63)	(0.38)								
Rate neg W6-W7			3.51***	4.18***								
			(0.50)	(0.40)								
Rate pos W7-W8			5.50***	5.71***								
			(0.49)	(0.37)								
Rate neg W7-W8			3.83***	4.48***								
			(0.37)	(0.30)								
Rate pos W8-W9			4.62***	5.30***								
			(0.43)	(0.30)								
Rate neg W8-W9			3.79***	4.18***								
			(0.35)	(0.25)								
pos.att: outdegree (density)	-1.53***	-1.47***	-1.08***	-1.24***	-1.42***	-1.71***	-1.61***	-1.41***	-1.41***	-1.84***	-1.53***	-1.47***
	(0.12)	(0.07)	(0.10)	(0.06)	(0.25)	(0.16)	(0.13)	(0.08)	(0.13)	(0.08)	(0.12)	(0.08)
neg.att: outdegree (density)	-1.75***	-1.61***	-1.25***	-1.49***	-1.10***	-1.82***	-1.56***	-1.48***	-1.78***	-1.95***	-1.75***	-1.61***
	(0.15)	(0.08)	(0.10)	(0.08)	(0.25)	(0.17)	(0.18)	(0.10)	(0.17)	(0.11)	(0.14)	(0.09)
pos.att: friend to agreement	0.06*	-0.01	0.05*	0.05***	-0.08	0.05	0.02	-0.03	0.08·	0.08*	0.06*	-0.01
	(0.02)	(0.02)	(0.02)	(0.01)	(0.06)	(0.05)	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)	(0.02)
neg.att: friend to agreement	0.06·	-0.01	0.06*	0.03·	0.06	0.03	0.04	0.00	0.12·	0.05	0.07·	-0.01
	(0.03)	(0.03)	(0.03)	(0.02)	(0.09)	(0.07)	(0.04)	(0.03)	(0.07)	(0.05)	(0.04)	(0.03)
pos.att: friend and neg.att two-paths[a]	-0.11*	-0.02	-0.06	-0.03	0.09	-0.08	-0.06	0.01	-0.13	0.05	-0.10*	-0.02

	(0.05)	(0.04)	(0.04)	(0.02)	(0.13)	(0.11)	(0.06)	(0.04)	(0.12)	(0.09)	(0.05)	(0.04)
neg.att: friend and pos.att two-paths[a]	0.08$^{\cdot}$	-0.07$^{\cdot}$	0.05	-0.07*	-0.24	-0.14	0.12*	-0.08*	-0.20	-0.33*	0.08$^{\cdot}$	-0.07$^{\cdot}$
	(0.05)	(0.04)	(0.04)	(0.03)	(0.17)	(0.11)	(0.05)	(0.04)	(0.19)	(0.14)	(0.04)	(0.04)
pos.att: 4-cycles (1)[b]	0.31***	0.13***	0.32***	0.14***	0.35***	0.14***	0.28***	0.12***	0.01***	0.69***	0.31***	0.13***
	(0.03)	(0.01)	(0.02)	(0.01)	(0.05)	(0.02)	(0.03)	(0.01)	(0.00)	(0.06)	(0.03)	(0.01)
neg.att: 4-cycles (1)[b]	0.63***	0.26***	0.66***	0.26***	0.83***	0.34***	0.55***	0.21***	0.02***	0.01***	0.63***	0.25***
	(0.07)	(0.02)	(0.06)	(0.02)	(0.14)	(0.04)	(0.08)	(0.02)	(0.00)	(0.00)	(0.07)	(0.02)
pos.att: shared neg.att (1) to agreement[b]	0.07	0.01	0.07	0.01	0.24*	0.01	-0.01	-0.03$^{\cdot}$	0.32	0.06	0.07	0.01
	(0.05)	(0.02)	(0.04)	(0.01)	(0.10)	(0.03)	(0.06)	(0.02)	(0.20)	(0.07)	(0.05)	(0.02)
neg.att: shared pos.att (1) to agreement[b]	0.18***	0.07***	0.13**	0.04**	0.26**	0.05	0.15**	0.03	0.25	0.24*	0.18***	0.07***
	(0.05)	(0.02)	(0.04)	(0.02)	(0.09)	(0.03)	(0.06)	(0.02)	(0.19)	(0.10)	(0.05)	(0.02)
pos.att: indegree - popularity[b]	0.65**	0.19**	-0.16	0.01	0.10	0.23*	1.02***	0.21***	0.10	0.32**	0.64**	0.19**
	(0.20)	(0.06)	(0.18)	(0.05)	(0.47)	(0.10)	(0.23)	(0.06)	(0.43)	(0.10)	(0.22)	(0.07)
neg.att: indegree - popularity[b]	0.14	0.06	-0.62*	-0.01	-0.01$^{\cdot}$	0.07	0.06	0.07	0.39	0.23	0.16	0.06
	(0.37)	(0.09)	(0.25)	(0.08)	(0.01)	(0.14)	(0.46)	(0.11)	(0.57)	(0.15)	(0.36)	(0.10)
pos.att: indegree neg.att popularity[b]	-0.27	-0.57***	-0.02***	-0.69***	-0.24	-0.19	-0.12	-0.70***	-1.01	-0.50*	-0.28	-0.57***
	(0.31)	(0.09)	(0.00)	(0.07)	(0.59)	(0.17)	(0.33)	(0.10)	(0.66)	(0.22)	(0.34)	(0.10)
neg.att: indegree pos.att popularity[b]	-1.02**	-0.46***	-0.02***	-0.64***	-0.67	-0.23	-0.01*	-0.55***	-0.01	-0.27	-1.02**	-0.47***
	(0.40)	(0.09)	(0.00)	(0.08)	(0.66)	(0.16)	(0.00)	(0.10)	(0.01)	(0.20)	(0.38)	(0.10)
pos.att: different neg.att (1) to disagreement type 1[ab]	0.87***	0.38***	0.87***	0.34***	0.92***	0.38***	0.72***	0.31***	0.04***	0.01***	0.88***	0.38***
	(0.12)	(0.04)	(0.10)	(0.03)	(0.26)	(0.06)	(0.14)	(0.04)	(0.01)	(0.00)	(0.13)	(0.04)
neg.att: different pos.att (1) to disagreement type 1[ab]	0.66***	0.31***	0.69***	0.28***	0.56***	0.33***	0.51***	0.29***	0.03***	0.01***	0.66***	0.31***
	(0.09)	(0.03)	(0.07)	(0.02)	(0.17)	(0.05)	(0.10)	(0.03)	(0.01)	(0.00)	(0.09)	(0.03)
pos.att: different neg.att (1) to disagreement type 2[ab]	0.57***	0.37***	0.69***	0.37***	0.40$^{\cdot}$	0.36***	0.54***	0.40***	0.03***	0.02***	0.57***	0.37***
	(0.12)	(0.04)	(0.10)	(0.03)	(0.22)	(0.06)	(0.13)	(0.04)	(0.01)	(0.00)	(0.12)	(0.04)
neg.att: different pos.att (1) to disagreement type 2[ab]	0.47**	0.36***	0.37**	0.42***	0.45	0.29**	0.32$^{\cdot}$	0.35***	0.03*	0.02***	0.47**	0.36***
	(0.15)	(0.05)	(0.14)	(0.04)	(0.29)	(0.09)	(0.18)	(0.05)	(0.01)	(0.00)	(0.16)	(0.06)
Overall max. t-ratio	0.186	0.145	0.224	0.236	0.234	0.293	0.229	0.215	0.204	0.249	0.250	0.249

Param. max. t-ratio	0.040	0.036	0.068	0.068	0.083	0.079	0.055	0.083	0.060	0.091	0.082	0.060
Iterations	8147	8147	2296	3116	3267	3246	2910	3223	2910	2954	2953	2851

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\cdot}p < 0.1$. SEs in brackets. Bolded variable names correspond to hypothesis-testing parameters. [a]Newly programmed effect. [b]Estimate and SE multiplied by 100

As with the main results, we interpret both the model parameters themselves, as well as multivariate Wald tests presented in the table at the end of this section, which test the sum of all parameters representing a single hypothesis. This results table is followed by goodness-of-fit plots on the focal structures.

Regarding H1a, that people will tend to have friendships with others with whom they share political attitudes, results are largely consistent with the main model. One of the two parameters was consistently positive and significant in Cohort 1, and never significant in Cohort 2, as in the main model. The second of the two, which was positive and significant in the main model for both cohorts, was not significant in either cohort in R2 or R4 (i.e. in the first period or with a higher binarization threshold). In R1, R5 and R3, the parameter was reduced to non-significance for Cohort 1. For Cohort 2, the parameter remained positive and significant for R1, R3 and R5. Using Wald tests grouping the parameters, results are significant except for Cohort 2 R2 and R4, though the χ^2 statistic tends to decrease relative to the main model.

H1b, that people will tend to avoid friendships with people with opposed attitudes, was consistent with the main model; not supported with either parameter in any robustness check model. Similarly, Wald tests of the grouped parameters testing this hypothesis found no significant effect of selection. The majority of the models contained at least one opposite-expectation parameter sign.

H2a, that people will tend to hold attitudes that their friends hold, was largely consistent with the main model, with some exceptions. As in the main model, a positive and significant effect was found for influence on positive attitudes towards a topic in Cohort 1 in R1 and R5, but no significant effect in R2, R3, and R4. In contrast, for Cohort 2, the expected effect was positive and significant in R1 and R4, unlike the non-significant effect in the main model. For the second effect testing H2a on negative attitudes, the positive non-significant effect in the main model for Cohort 1 was significant in R1, but not in any other model. For Cohort 2, which did not have a significant parameter in the main model, it remained unsupported in other models. Wald tests supported the hypothesis in both cohorts for R1 (as opposed to the main model where the hypothesis was only supported in Cohort 1). In R4 and R5, the hypothesis was supported with a similar pattern to the main model, i.e. only in Cohort 1. In R2 and R3 the hypothesis was not supported by the Wald tests.

H2b, that people will tend to avoid holding attitudes that are opposed to their friends', was tested by two parameters. For the first of these, the robustness check models replicated the positive significant effect in Cohort 1 only in R5. Consistent with the findings for Cohort 2, no effect was found in any of the robustness check models. For the second parameter, the non-significant parameter found in Cohort 1 was increased to positive significance (counter to expectation) in R3, but was non-significant in R1, R2, R4, and R5. The negative parameter found in Cohort 2 became significant in R1, R3 and R4, in line with expectation, while staying non-significant in R5 and R2. Wald tests yield results similar to the main model: Only in the case of R1 is the test of the parameters significant, and only for Cohort 2.

Moving to H3a, that individuals sharing some pre-existing attitude are more likely to share more attitudes in the future, results are more straightforward. Of the first two of four parameters, both remained positive and significant in all robustness checks for both cohorts, as in the main model. The third parameter, which was not significant in either cohort in the main model, became positive significant in R2 for Cohort 1 and non-significant in R3 for Cohort 2. The fourth parameter, positive and significant for both cohorts in the main model, was supported fully in R1 and R5. In R2 and R3 it was only supported for Cohort 1, and In R4 for Cohort 2. Wald tests support the hypothesis in both cohorts for all robustness checks.

For H3b, that individuals with opposing attitudes will continue to have more opposing attitudes in the future, there are again four parameters, showing robust results. In both cohorts, for all four parameters, results are the same as in the main model – parameters are positive and significant with three exceptions. Both of the the parameters become non-significant in Cohort 1 in R2, while for the same cohort in R3 one parameter becomes non-significant. Wald tests support the hypothesis in all cases.

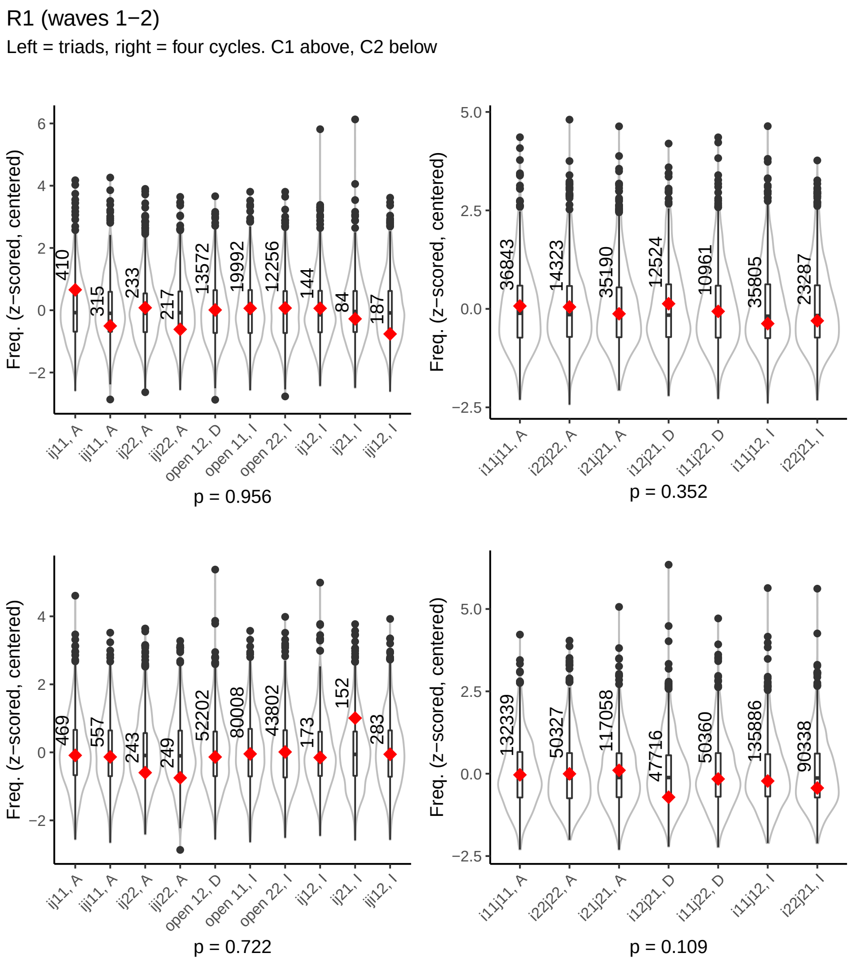
Overall, mixed results on the hypotheses largely retain the level of support across robustness checks, although patterns are not identical. In sum, it appears that positive selection (H1a) is robustly supported, negative selection (H1b) is unsupported, positive influence (H2a) is weakly supported, negative influence (H2b) is not supported, and latent-cause similarity (H3a) and dissimilarity (H3b) are robustly supported.

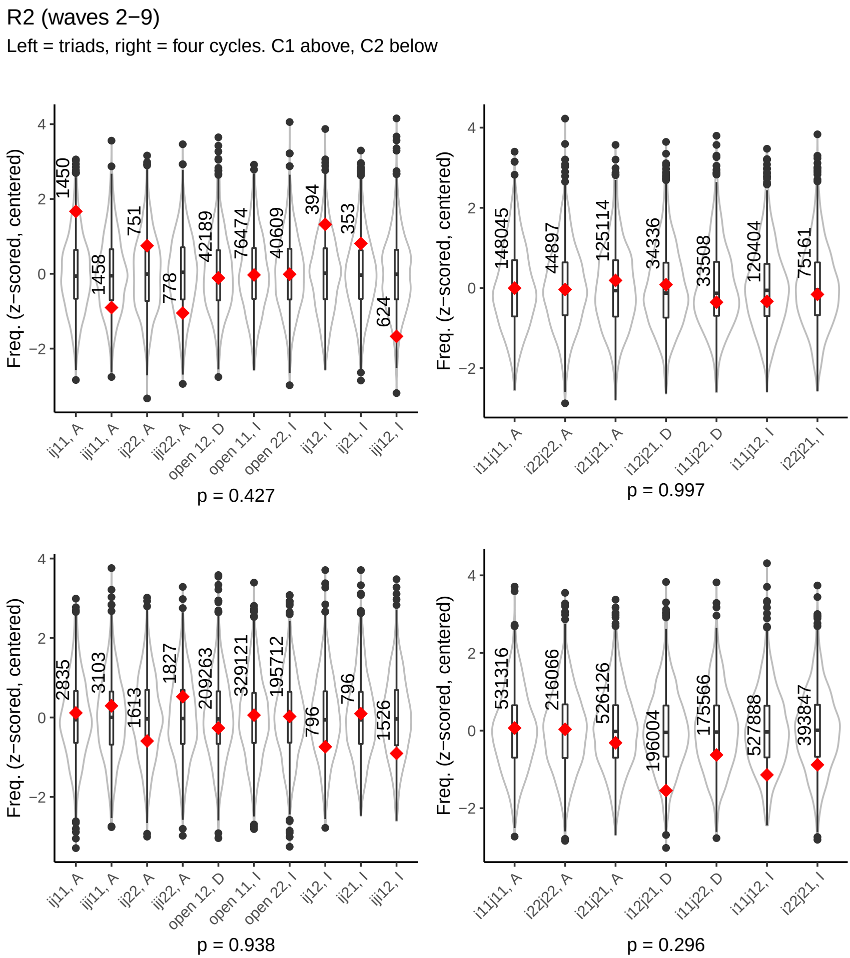
TABLE A.3: Wald tests of summed parameters

	H1a	H1b	H2a	H2b	H3a	H3b
C1, Main	13.67(1)***	1.68(1)	9.06(1)**	0.12(1)	110.65(1)***	135.28(1)***
C2, Main	5.65(1)*	*2.73(1)*	*0.39(1)*	3.29(1)	368.04(1)***	195.78(1)***
C1R1, W1-2	8.33(1)**	0.42(1)	10.43(1)**	0.02(1)	155.3(1)***	192.2(1)***
C2R1, W1-2	8.88(1)**	0.55(1)	13.16(1)***	6.63(1)*	472.93(1)***	243.48(1)***
C1R2, W2-9	4.47(1)*	0.02(1)	0.05(1)	0.37(1)	24.51(1)***	73.01(1)***
C2R2, W2-9	1.34(1)	0.03(1)	0.99(1)	2.1(1)	91.95(1)***	69.56(1)***
C1R3, no opin. as missing	10.29(1)**	0.38(1)	1.86(1)	0.68(1)	53.66(1)***	67.76(1)***
C2R3, no opin. as missing	10.63(1)**	*1.19(1)*	0.59(1)	1.65(1)	287.65(1)***	82.4(1)***
C1R4, high tie thresh.	5.47(1)*	0.26(1)	5.75(1)*	1.99(1)	64.16(1)***	4.75(1)*
C2R4, high tie thresh.	0.14(1)	*1.03(1)*	3.71(1)	2.77(1)	161.93(1)***	56.02(1)***
C1R5, pol. orientation	11.5(1)***	0.88(1)	8.6(1)**	0.1(1)	102.5(1)***	131.48(1)***
C2R5, pol. orientation	5.96(1)*	*2.11(1)*	*0.38(1)*	2.8(1)	299.88(1)***	173.52(1)***

Note. Values are χ^2 for sum of parameters with degrees of freedom in brackets. Italicized values indicate all parameters being in a direction counter to expectation. * : p<.05,** : p <.01,*** : p<.001.

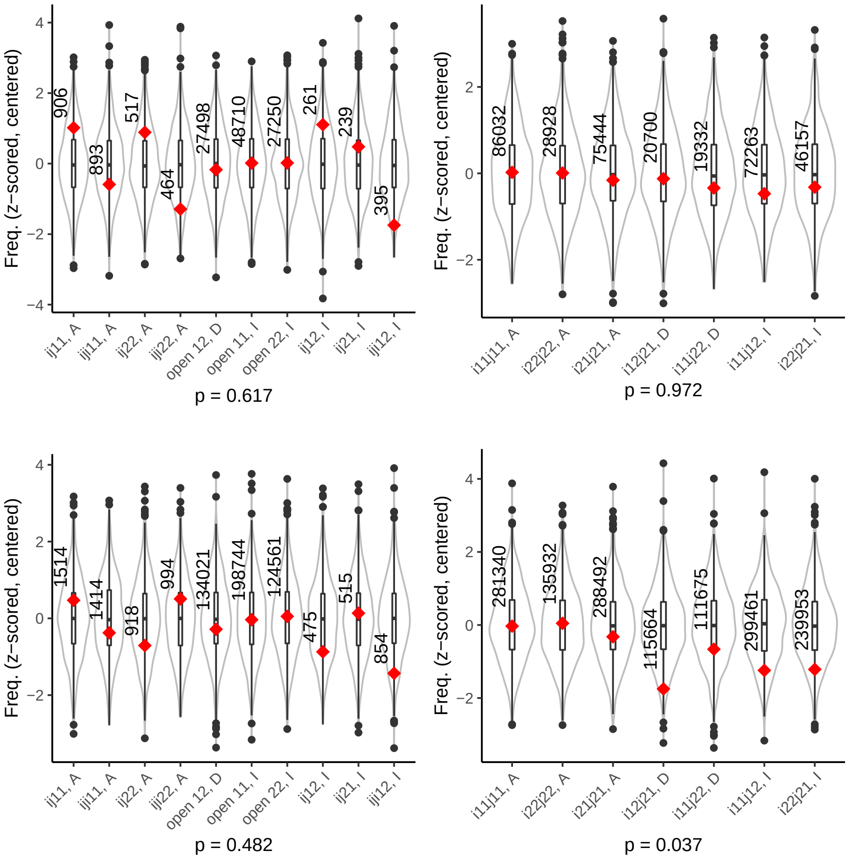
A.5.2 Main goodness of fit statistics for robustness checks




R2 (waves 2−9)
Left = triads, right = four cycles. C1 above, C2 below
Freq. (z−scored, centered)
ij11, A
iji11, A
ij22, A
iji22, A
open 12, D
open 11, I
open 22, I
ij12, I
ij21, I
iji12, I
1450
1458
751
778
42189
76474
40609
394
353
624
p = 0.427
i11j11, A
i22j22, A
i21j21, A
i12j21, D
i11j22, D
i11j12, I
i22j21, I
148045
44897
125114
34336
33508
120404
75161
p = 0.997
2835
3103
1613
1827
209263
329121
195712
796
796
1526
p = 0.938
531316
216066
526126
196004
175566
527888
393847
p = 0.296

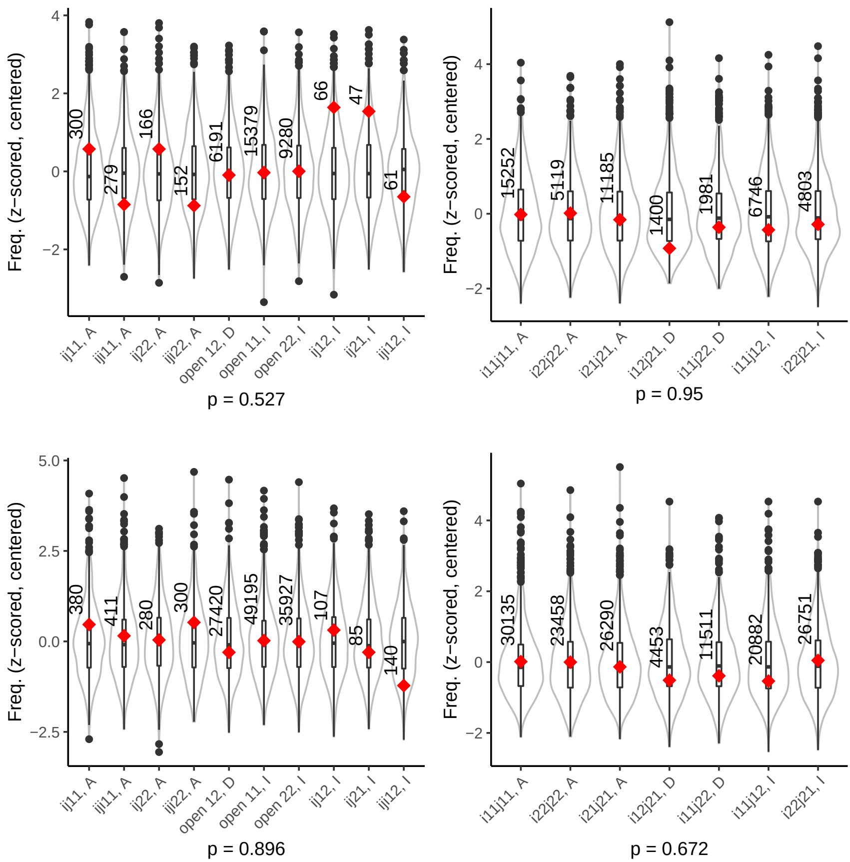
R3 (no opin as missing)

Left = triads, right = four cycles. C1 above, C2 below



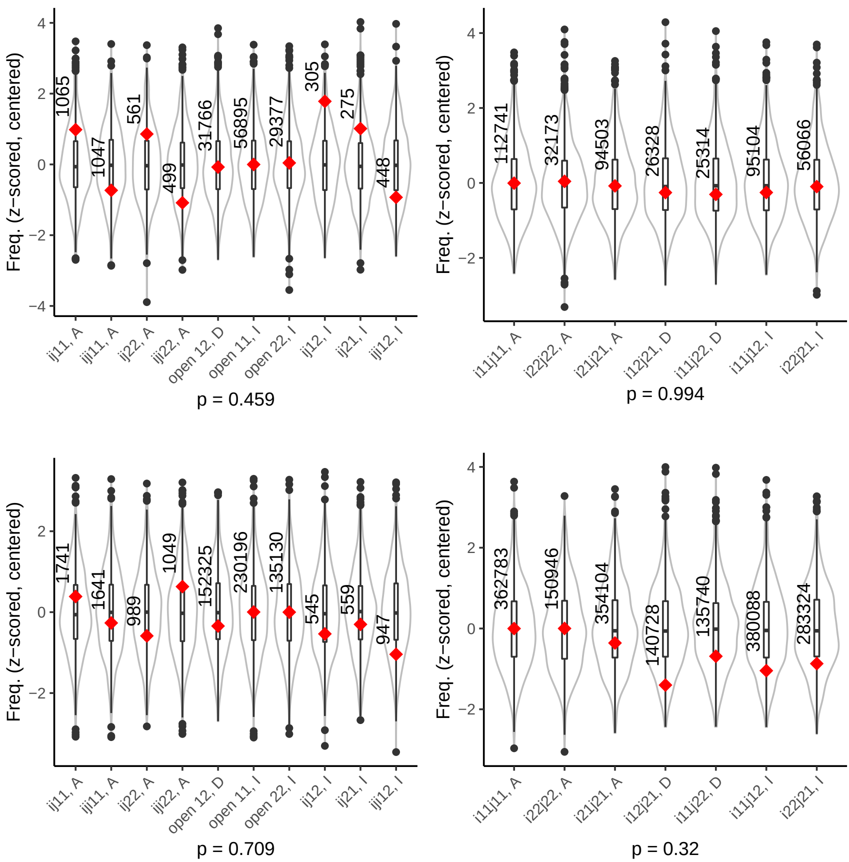
R4 (higher tie threshold)

Left = triads, right = four cycles. C1 above, C2 below



R5 (political orientation covar)

Left = triads, right = four cycles. C1 above, C2 below



A.6 SUMMARY OF EFFECTS AND GOODNESS OF FIT NEWLY ADDED TO RSIENA

A.6.1 Effects

SAOM effects for three networks with mixed two-mode and one-mode node-sets were not implemented in the RSiena software at time of writing. Thus, four new effects were added to RSiena for the purposes of our study. First of these were two, three-network, mixed triangle effects. These are (using the conventional RSiena shortnames) the "fromDiff" and "toDiff" effects, respectively, in analogy to "from" and "to" effects for two first mode nodes having a common bipartite node to which they connect in the same network type, already present in RSiena.

For the "fromDiff" effect, the dependent tie is the first-mode tie between two actors, dependent on their having a different connection type to a second-mode node they have in common. These correspond to the parameters reported in Table A.4 as "Heterophily type 1" and "Heterophily type 2".

For the "toDiff" effect, the dependent tie is a second-mode tie of a different type than the one held by another actor to which the first actor has an outgoing tie in the first mode. This effect was used in the estimation of parameters indicated by "Influence, pos. att and neg. att" and "Influence, neg. att and pos. att".

In language more familiar to Siena users: "fromDiff" is the effect of the mixed 2-instar to second-mode node m from i in network W and from j in network V, on the directed tie from i to j in network X.

The "toDiff" effect is the effect of the directed twopath from actor i to node m, via first-mode node j with ties in network X and W respectively, on actor i's connection to m in network V.

The second two new effects relate to the mixed four cycles in the bipartite networks - in particular the oppositional kinds. These are named "oppCycle4.1" and "oppCycle4.2" effects. The .1 and .2 indicate that these are two different forms of oppositional four-cycles, as for our purposes we saw no reason to distinguish between the two conceptually. For both of the statistics below, X represents the dependent network, W the independent.

oppCycle4.1 is the clockwise path of ties $XWWX$, starting at an actor sending tie in network X to the second mode node. This effect was used to estimate parameters reported in Table A.3 as "Latent divergence type 2, pos. att." and "Latent divergence type 2, neg. att.".

oppCycle4.2 is the clockwise path of ties $XWXW$, with the same starting position. This effect was used to estimate parameters reported in Table A.3 as "Latent divergence type 1, pos. att." and "Latent divergence type 1, neg. att.".

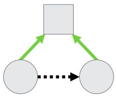
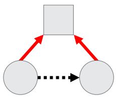
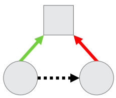
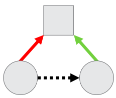
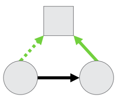
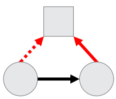
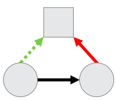
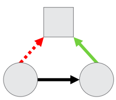
A.6.2 Auxiliary statistics (Goodness of fit)

We added two kinds of new auxiliary statistics, which we used for goodness-of-fit tests in SAOMs. First, we defined a census of mixed triads of three nodes, two first-mode and two second-mode, including up to three networks. Ties can only be reciprocal in the pair of first-mode nodes. Additionally, the bipartite networks are disjoint (i.e. there cannot be a tie present in both networks, for one dyad at the same time). This census is in the exclusive form, i.e. a closed triad is not additionally counted in any separate twopath configuration. Similarly, we count a reciprocal tie between two nodes only once.

Secondly, we created a new census of complete (i.e. where all pairs of nodes which has one in the second mode and one in the first mode set have a connection) four-cycles for configurations of two first-mode nodes, two second-mode nodes, and two network types connecting them, which are again disjoint.

A.7 MAIN SAOM RESULTS

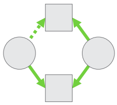
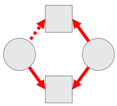
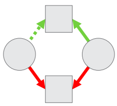
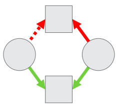
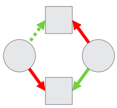
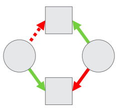
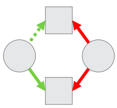
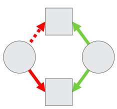
Table A.4: Main Stochastic Actor Oriented Model parameter estimates and standard errors: Selection and influence

	Effect	Cohort 1	Cohort 2
	Homophily from pos. att.	0.05* (0.02)	0.02 (0.02)
	Homophily from neg. att.	0.07* (0.03)	0.06* (0.03)
	Heterophily type 1[a]	−0.05 (0.05)	0.05 (0.04)
	Heterophily type 2[a]	−0.04 (0.05)	0.03 (0.03)
	Influence, pos. att.	0.06* (0.02)	−0.01 (0.02)
	Influence, neg. att.	0.06 (0.03)	−0.01 (0.03)
	Influence, pos. att. and neg. att.	−0.11* (0.05)	−0.02 (0.04)
	Influence, neg. att. and pos. att.	0.08 (0.05)	−0.07 (0.04)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Standard errors in parentheses. Dependent tie is dotted line.

[a]Newly programmed effect.

TABLE A.5: Main Stochastic Actor Oriented Model parameter estimates and standard errors: Convergence and divergence

	Effect	Cohort 1	Cohort 2
	Latent convergence type 1, pos. att.[b]	0.31*** (0.03)	0.13*** (0.01)
	Latent convergence type 1, neg. att.[b]	0.63*** (0.07)	0.26*** (0.02)
	Latent convergence type 2, pos. att.[b]	0.07 (0.05)	0.01 (0.02)
	Latent convergence type 2, neg. att.[b]	0.18*** (0.05)	0.07*** (0.02)
	Latent divergence type 1, pos. att.[ab]	0.87*** (0.12)	0.38*** (0.04)
	Latent divergence type 1, neg. att.[ab]	0.66*** (0.09)	0.31*** (0.03)
	Latent divergence type 2, pos. att.[ab]	0.57*** (0.12)	0.37*** (0.04)
	Latent divergence type 2, neg. att.[ab]	0.47** (0.15)	0.36*** (0.05)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Standard errors in parentheses. Dependent tie is dotted line.
[a]Newly programmed effect. [b]Estimate and SE multiplied by 100.

A.8 INTERPRETATION OF AUXILIARY RESULTS OF THE MAIN MODEL

For the results described below, see the tables at the end of this section. The rates in the friendship networks ranged from 4.30 to 6.97 for both cohorts. In the political attitude network, rates ranged from 3.55 to 7.37. This means that on average, individuals are estimated to have considered changing one of their outgoing friendship ties and attitude ties in each network around 4 to 7 times between subsequent observations.

Examining endogeneity in the friendship networks of both cohorts, we find positive and significant effects of reciprocity and transitivity. Furthermore, their interaction was negative and significant, as shown previously by Block (2015). Indegree popularity was positive and significant, suggesting individuals prefer to connect to alters who have more incoming friendship ties. The related effects of outdegree popularity and outdegree activity were both negative and significant, suggesting that individuals prefer to connect to alters who send fewer friendship nominations, and that they prefer to nominate fewer people as friends as well.

Turning to homophily, we find positive and significant effects of speaking the same language, being from the same study major (in Cohort 1), and gender. A positive, yet non-significant effect for age homophily was found in Cohort 2. In Cohort 2 there were additional positive significant effects of gender on friendship, with women more likely to nominate, and be nominated as friends than men. Indegree popularity for positive attitude about a political statement was positive significant in both samples, but not significant in either sample for negative attitude. The effect of many positive attitudes on others' negative attitude on a statement and vice versa was negative and significant, but only in Cohort 2.

TABLE A.6: Full Stochastic Actor Oriented Model parameter estimates.

	Cohort 1	Cohort 2
Rate friend W2-W3	6.87 (0.55)***	5.98 (0.36)***
Rate friend W3-W4	6.97 (0.58)***	6.48 (0.36)***
Rate friend W4-W5	4.30 (0.37)***	4.94 (0.28)***
Density friend (intercept)	−3.10 (0.15)***	−3.57 (0.10)***
friend: reciprocity	3.49 (0.18)***	3.85 (0.14)***
friend: transitive triplets	0.81 (0.05)***	1.22 (0.04)***
friend: transitive recipr. triplets	−0.62 (0.09)***	−1.12 (0.07)***
friend: indegree - popularity	0.07 (0.03)**	0.06 (0.01)***
friend: outdegree - popularity	−0.18 (0.03)***	−0.20 (0.02)***
friend: outdegree - activity	−0.02 (0.01)*	−0.01 (0.01)*
friend: anti-isolate	−1.70 (0.39)***	−1.41 (0.35)***
friend: ethnicity	0.23 (0.08)**	0.33 (0.06)***
friend: same major	0.50 (0.06)***	
friend: gender alter	0.05 (0.07)	0.27 (0.07)***
friend: gender ego	−0.16 (0.08)	0.49 (0.08)***
friend: same gender	0.15 (0.07)*	0.42 (0.07)***
friend: birthyear similarity	0.44 (0.29)	0.45 (0.22)*
friend: neg.att-pos.att two-paths[a]	−0.05 (0.05)	0.05 (0.04)
friend: pos.att-neg.att two-paths[a]	−0.04 (0.05)	0.03 (0.03)
friend: pos.att agreement	0.05 (0.02)*	0.02 (0.02)
friend: neg.att agreement	0.07 (0.03)*	0.06 (0.03)*
Overall max. t-ratio	0.19	0.14
Param. max. t-ratio	0.04	0.04
Iterations	8147	8147

Note.$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Dependent tie is dotted line. [a]Newly programmed effect. [b]Estimate and SE multiplied by 100

TABLE A.7: Full Stochastic Actor Oriented Model parameter estimates. (cont)

	Cohort 1	Cohort 2
Rate pos W2-W3	6.30 (0.38)***	7.32 (0.33)***
Rate neg W2-W3	4.79 (0.35)***	5.43 (0.27)***
Rate pos W3-W4	7.37 (0.46)***	7.12 (0.30)***
Rate neg W3-W4	4.97 (0.39)***	5.46 (0.27)***
Rate pos W4-W5	6.52 (0.43)***	6.73 (0.29)***
Rate neg W4-W5	3.55 (0.28)***	4.31 (0.22)***
pos.att: outdegree (density)	−1.53 (0.12)***	−1.47 (0.07)***
neg.att: outdegree (density)	−1.75 (0.15)***	−1.61 (0.08)***
pos.att: friend to agreement	0.06 (0.02)*	−0.01 (0.02)
neg.att: friend to agreement	0.06 (0.03)	−0.01 (0.03)
pos.att: friend and neg.att two-paths[a]	−0.11 (0.05)*	−0.02 (0.04)
neg.att: friend and pos.att two-paths[a]	0.08 (0.05)	−0.07 (0.04)
pos.att: 4-cycles (1)[b]	0.31 (0.03)***	0.13 (0.01)***
neg.att: 4-cycles (1)[b]	0.63 (0.07)***	0.26 (0.02)***
pos.att: shared neg.att (1) to agreement[b]	0.07 (0.05)	0.01 (0.02)
neg.att: shared pos.att (1) to agreement[b]	0.18 (0.05)***	0.07 (0.02)***
pos.att: indegree - popularity[b]	0.65 (0.20)**	0.19 (0.06)**
neg.att: indegree - popularity[b]	0.14 (0.37)	0.06 (0.09)
pos.att: indegree neg.att popularity[b]	−0.27 (0.31)	−0.57 (0.09)***
neg.att: indegree pos.att popularity[b]	−1.02 (0.40)**	−0.46 (0.09)***
pos.att: different neg.att (1) to disagreement type 2[ab]	0.57 (0.12)***	0.37 (0.04)***
neg.att: different pos.att (1) to disagreement type 2[ab]	0.47 (0.15)**	0.36 (0.05)***
Overall max. t-ratio	0.19	0.14
Param. max. t-ratio	0.04	0.04
Iterations	8147	8147

Note. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Dependent tie is dotted line. [a]Newly programmed effect. [b]Estimate and SE multiplied by 100

A.9 NETWORK DESCRIPTIVES

Friendship descriptives

TABLE A.8: Friendship network descriptives, Cohort 1

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
N	202	236	225	215	210
Structural missing	59	25	36	46	51
Density	0.01	0.02	0.02	0.02	0.02
Mean degree	0.92	2.39	2.30	2.44	1.92
SD indegree	1.26	2.40	2.37	2.59	2.10
SD outdegree	1.73	3.17	3.08	3.29	2.69
Jaccard (t vs. $t-1$)		0.37	0.66	0.68	0.75

Structural missing indicates individuals not part of the cohort at time of data collection.

TABLE A.9: Friendship network descriptives, Cohort 2

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
N	529	635	625	620	603
Structural missing	131	25	35	40	57
Density	0.005	0.01	0.01	0.01	0.01
Mean degree	1.54	1.93	2.22	2.55	2.15
SD indegree	1.65	1.90	2.16	2.22	2.13
SD outdegree	1.98	2.63	2.94	3.17	3.31
Jaccard (t vs. $t-1$)		0.55	0.68	0.70	0.76

Structural missing indicates individuals not part of the cohort at time of data collection.

Attitudinal descriptives

Table A.10: Political network descriptives, Cohort 1

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Density, pos	0.37	0.34	0.34	0.34	0.34
Density, neg	0.23	0.24	0.21	0.23	0.22
Mean indegree, pos	49.55	44.27	42.73	42.23	36.18
Mean indegree, neg	30.59	31.41	26.36	28.91	23.91
Mean outdegree, pos	4.18	3.73	3.60	3.56	3.05
Mean outdegree, neg	2.58	2.65	2.22	2.44	2.02
SD indegree, pos	33.69	24.94	25.82	26.39	23.20
SD indegree, neg	25.69	24.06	22.40	22.79	20.41
SD outdegree, pos	4.41	4.25	4.39	4.31	4.20
SD outdegree, neg	3.00	3.22	2.99	3.19	2.94
Jaccard, pos (t vs. $t-1$)		0.76	0.76	0.76	0.77
Jaccard, neg (t vs. $t-1$)		0.73	0.72	0.72	0.78

Table A.11: Political network descriptives, Cohort 2

	Wave 1	Wave 2	Wave 3	Wave 4	Wave 5
Density, pos	0.35	0.32	0.31	0.30	0.31
Density, neg	0.25	0.24	0.23	0.24	0.23
Mean indegree, pos	119.27	88	91.32	90.64	76.23
Mean indegree, neg	84.64	66.27	67.68	72.09	56.95
Mean outdegree, pos	3.98	2.93	3.04	3.02	2.54
Mean outdegree, neg	2.82	2.21	2.26	2.40	1.90
SD indegree, pos	81.51	55.40	58.57	56.53	52.26
SD indegree, neg	62.74	45.90	47.04	50.54	41.11
SD outdegree, pos	4.12	3.89	3.88	3.77	3.66
SD outdegree, neg	3.28	3.30	3.25	3.33	3.10
Jaccard, pos (t vs. $t-1$)		0.72	0.71	0.73	0.72
Jaccard, neg (t vs. $t-1$)		0.72	0.72	0.69	0.73

A.10 EXPECTATION AND OBSERVATION BY WAVE IN FOUR DIMENSIONS

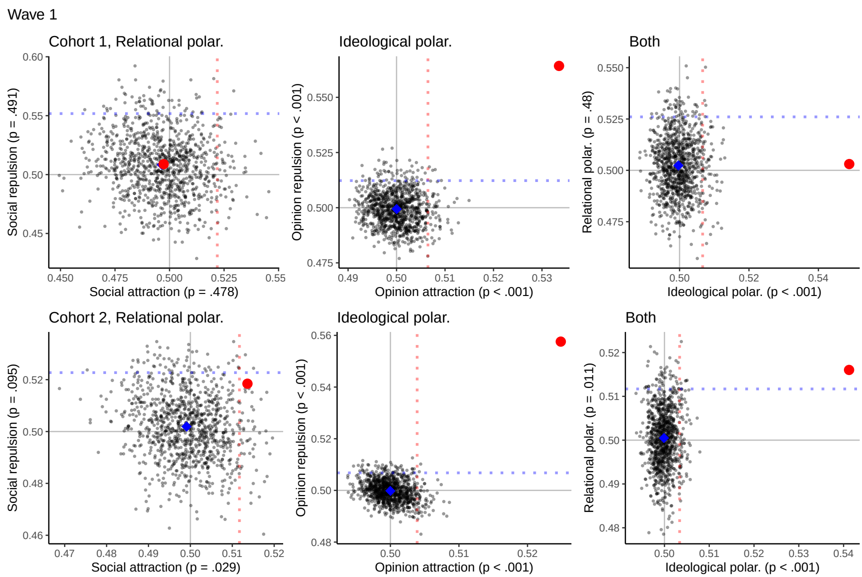
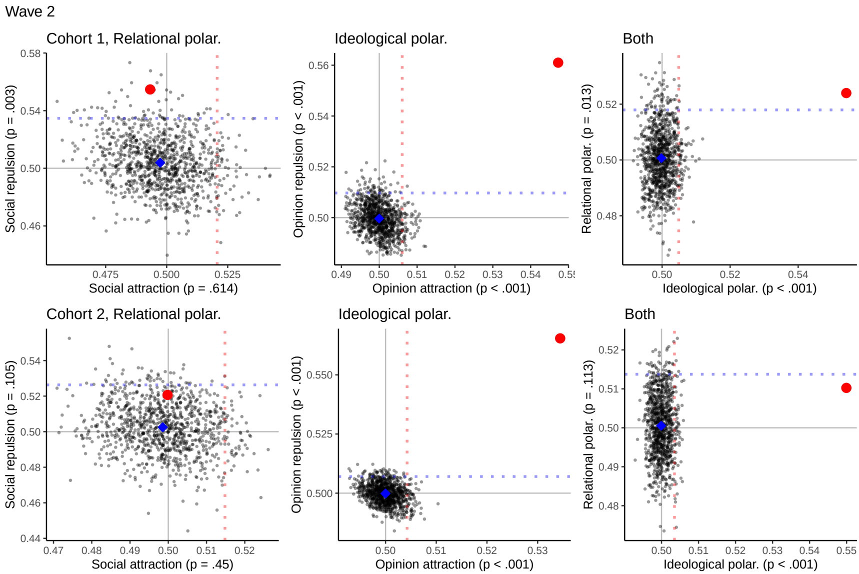
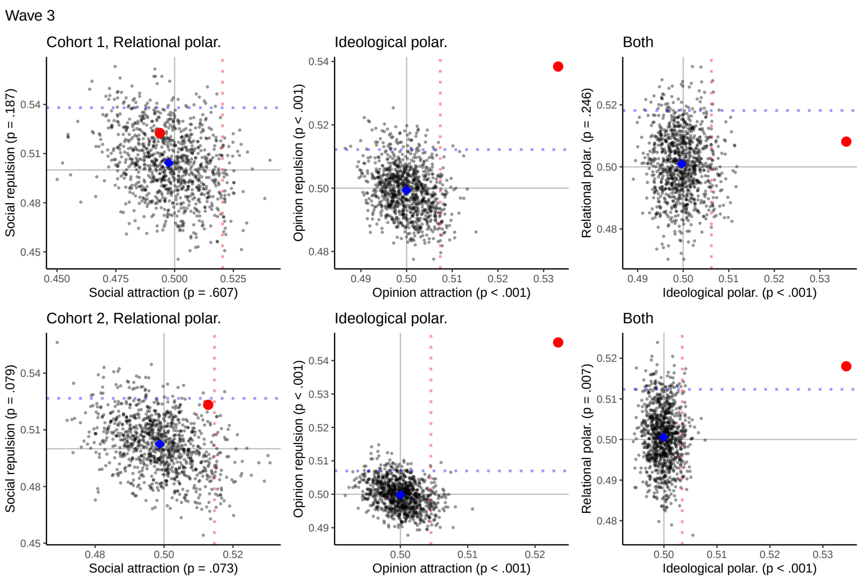
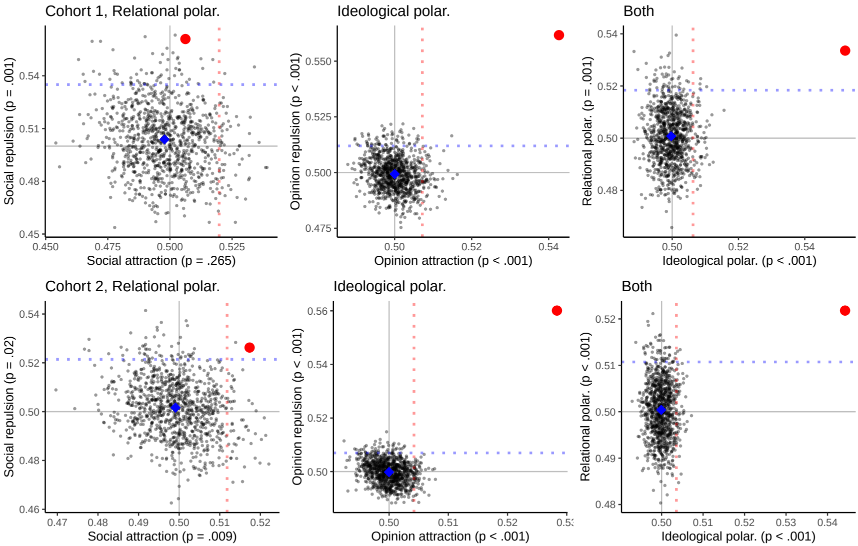


FIGURE A.2: Two-dimensional observed network polarization (red circle) relative to expectation in two cohorts (blue diamond), with simulated values (grey dots). Dotted blue line indicates boundary for vertical axis, red indicates boundary of variable on horizontal axis, both at $p < .05$ (one-sided).

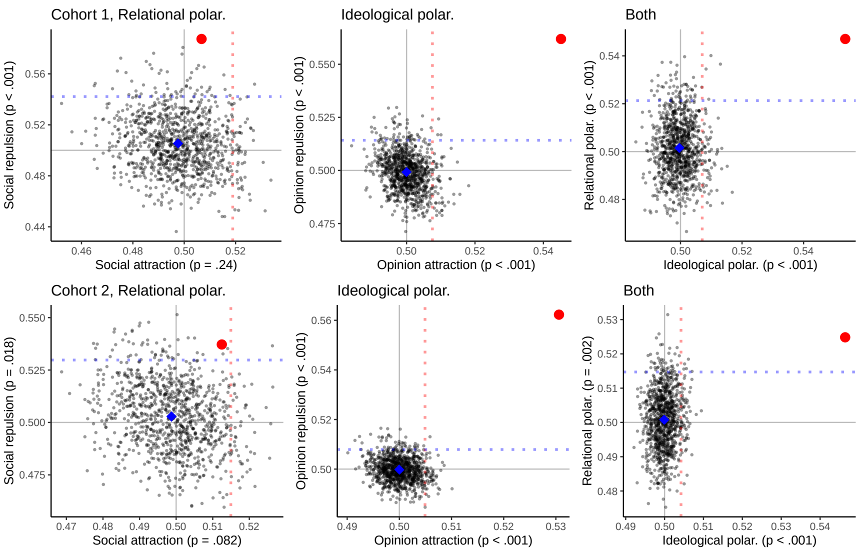

Wave 2
Cohort 1, Relational polar.
Social repulsion (p = .003)
Social attraction (p = .614)
Ideological polar.
Opinion repulsion (p < .001)
Opinion attraction (p < .001)
Both
Relational polar. (p = .013)
Ideological polar. (p < .001)
Cohort 2, Relational polar.
Social repulsion (p = .105)
Social attraction (p = .45)
Ideological polar.
Opinion repulsion (p < .001)
Opinion attraction (p < .001)
Both
Relational polar. (p = .113)
Ideological polar. (p < .001)


Wave 3
Cohort 1, Relational polar.
Social repulsion (p = .187)
Social attraction (p = .607)
Ideological polar.
Opinion repulsion (p < .001)
Opinion attraction (p < .001)
Both
Relational polar. (p = .246)
Ideological polar. (p < .001)
Cohort 2, Relational polar.
Social repulsion (p = .079)
Social attraction (p = .073)
Ideological polar.
Opinion repulsion (p < .001)
Opinion attraction (p < .001)
Both
Relational polar. (p = .007)
Ideological polar. (p < .001)

Wave 4


Cohort 1, Relational polar.
Social repulsion (p = .001)
Social attraction (p = .265)
Ideological polar.
Opinion repulsion (p < .001)
Opinion attraction (p < .001)
Both
Relational polar. (p = .001)
Ideological polar. (p < .001)
Cohort 2, Relational polar.
Social repulsion (p = .02)
Social attraction (p = .009)
Ideological polar.
Opinion repulsion (p < .001)
Opinion attraction (p < .001)
Both
Relational polar. (p < .001)
Ideological polar. (p < .001)

Wave 5


Cohort 1, Relational polar.
Social repulsion (p < .001)
Social attraction (p = .24)
Ideological polar.
Opinion repulsion (p < .001)
Opinion attraction (p < .001)
Both
Relational polar. (p < .001)
Ideological polar. (p < .001)
Cohort 2, Relational polar.
Social repulsion (p = .018)
Social attraction (p = .082)
Ideological polar.
Opinion repulsion (p < .001)
Opinion attraction (p < .001)
Both
Relational polar. (p = .002)
Ideological polar. (p < .001)

A.11 MAIN GOF STATISTICS AGGREGATED ACROSS WAVES

Figures A.3 and A.4 show the most relevant structures which the estimated model should reproduce. These are those that are also used in the calculation of network polarization. In these figures, the x-axis labels give a notational and graphical representation of all relevant tetradic and triadic multilevel network structures.

The violin plots here show the frequency distribution of the structure specified on the x-axes, in networks generated by simulating from each observation wave until the following observation 5000 times, using the estimated model. The red diamond represents the observed value relative to these simulations, aggregated across all simulations. Numbers are the true counts of the given statistic. A better fitting model has observed values closer to the center of the simulated distribution. Using the Mahalanobis distance of the given structures in the observed versus simulated data, we test for significant discrepancy between our model and our observation. If the p-value is low, observations in the data fall too far to the extremes of the distributions generated by model simulation and therefore it is unlikely that our observed data could have been drawn from the estimated model.

First, we examine the mixed four cycles of pairs of individuals and their attitudes, indicating the attitude clustering regardless of social connection, as given in Formulas 2.1 and 2.2. Figure A.3 shows that the model plausibly reproduces the observed data in terms of attitudes between individuals in both cohorts, although there is a tendency towards over-generation of perfect disagreement structures and mixed agreement-disagreement structures in Cohort 2. Overall, the model fits well to the focal features. Goodness of fit on other features of the network are reported in Appendix A.13.

Second, we examine mixed triads of individuals and political statements, connected by friendships and attitudes. As mentioned previously, these represent the clustering of individuals with shared (or opposing) attitudes, i.e. the social aspect of polarization given in Formulas 2.3 and 2.4. Figure A.4 shows the fit of the model to the observed census of mixed triads in the network. We see that the model does not significantly misfit on these structures in either cohort. However, it seems that in Cohort 1 there are slightly too many structures with reciprocal friendship ties generated by the model, and slightly too few with one-sided friendships. This bias is not apparent in Cohort 2.

In analyses not reported here, separation of the fit statistics by period (i.e. counting the structures at each sequential wave, instead of in aggregate) suggests the model significantly overestimates the tendency towards specific polarization structures at the earlier time point, but underestimates them at the later one, particularly in Cohort 2 (Appendix A.12). Given the high number of parameters included in the model, however, we opt to retain the uniform parameter estimates to reduce the risk of overfitting and type II errors.

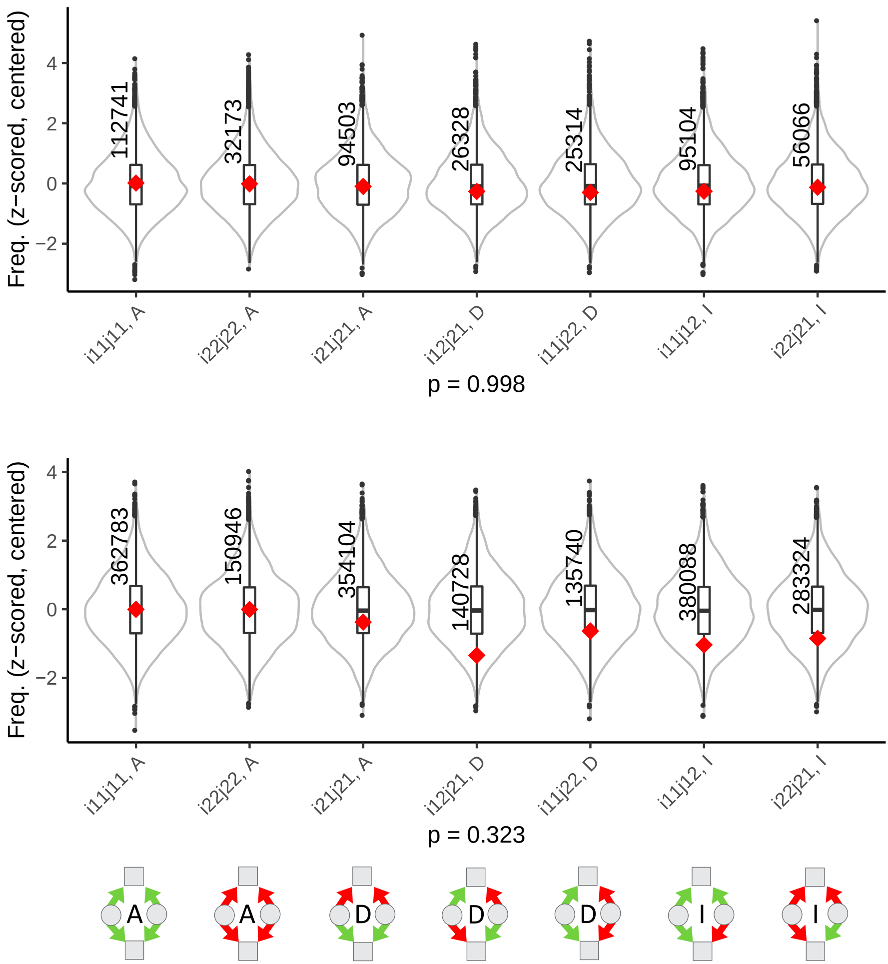


Figure A.3: Goodness-of-fit on four-cycle meso-structures. Upper row is Cohort 1, lower row is Cohort 2. Columns represent structures with corresponding notation and pictogram. *i* and *j* represent the individuals on the left and right, with the first digit representing the upper tie and the second digit the lower. Digits 1 and 2 indicate positive and negative network. *p*-values from Mahalanobis distance between all structures in simulation and observed data.

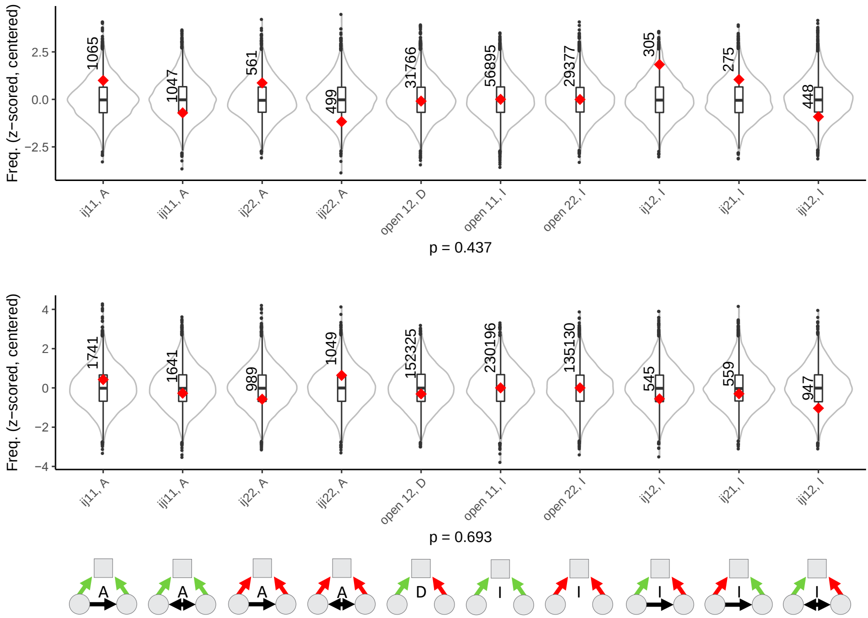
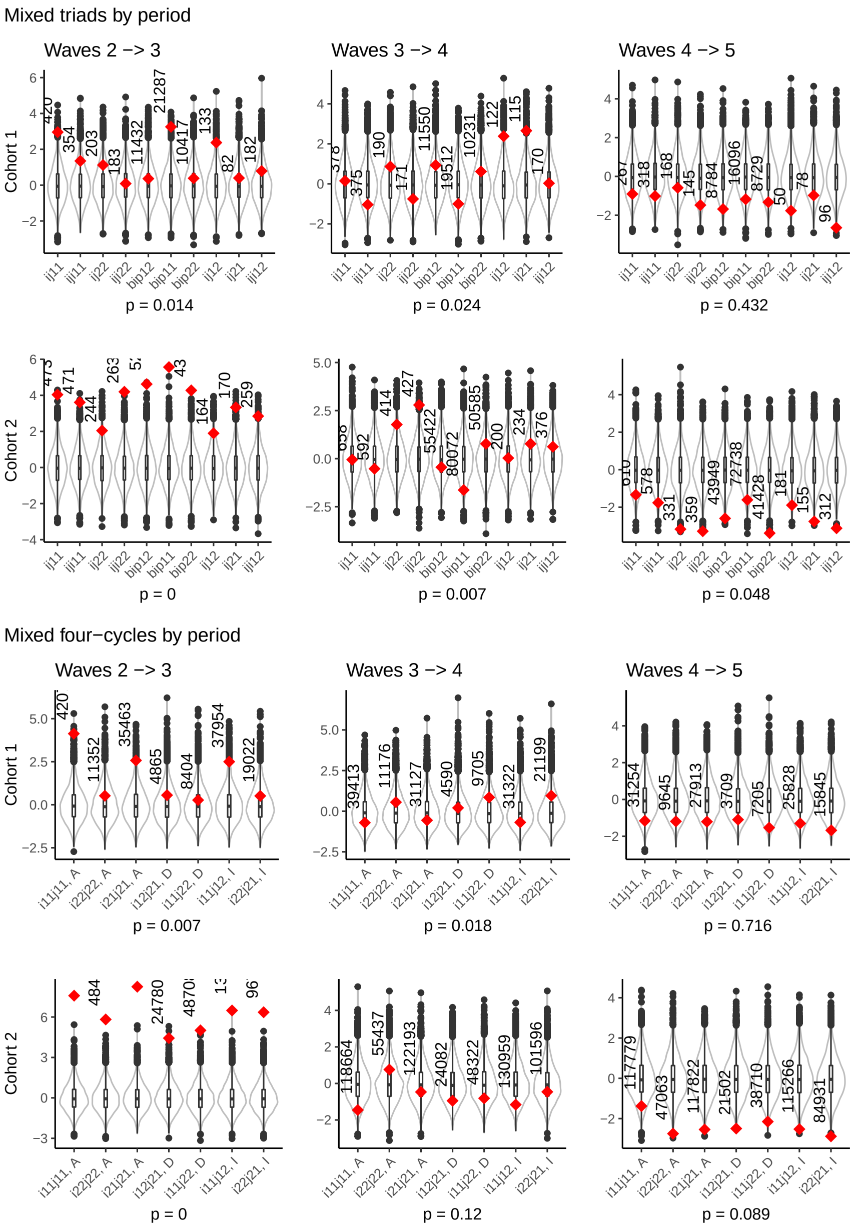
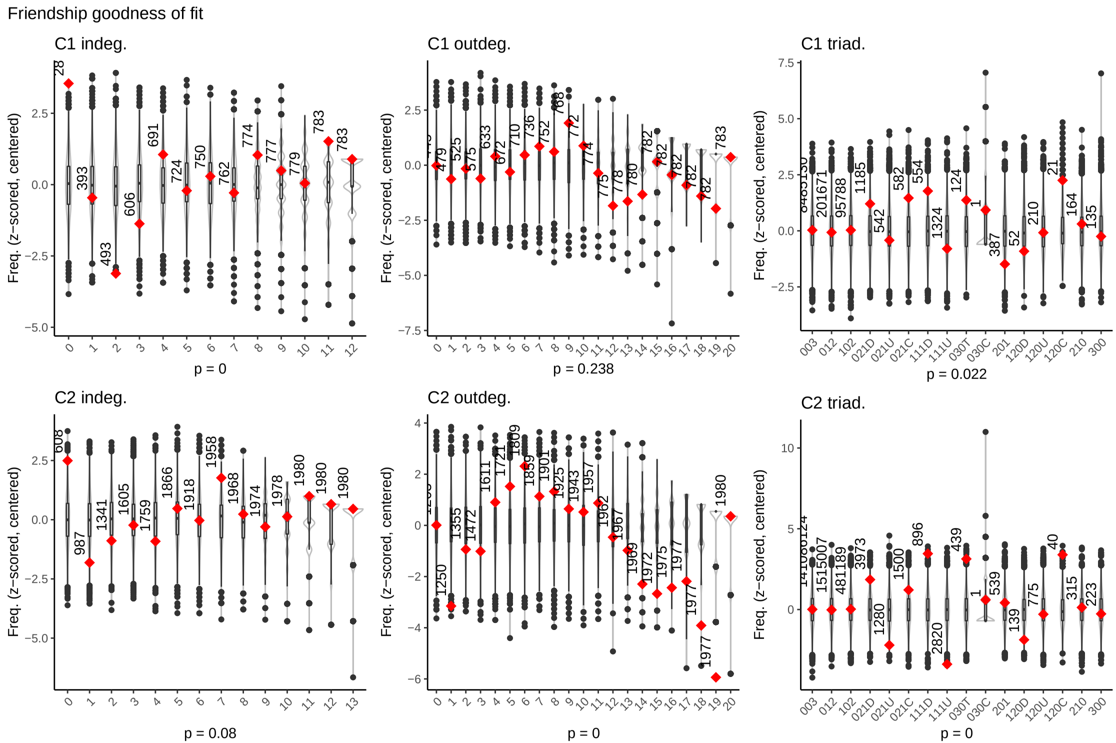


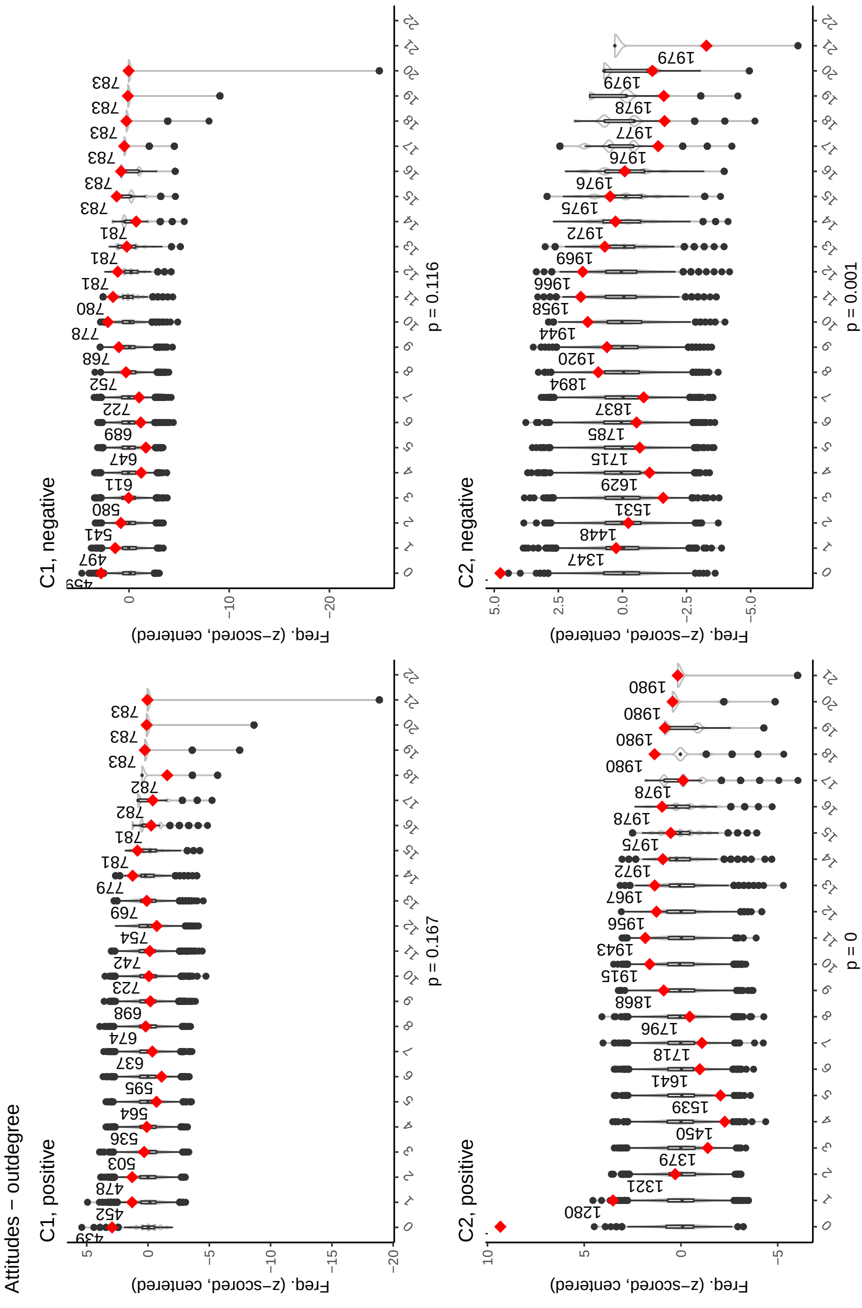
FIGURE A.4: Goodness-of-fit on triadic meso-structures. Upper row is Cohort 1, lower row is Cohort 2. Columns represent structures with corresponding notation and pictogram. *ij* and *iji* represent single and reciprocal social ties respectively. Digits 1 and 2 indicate positive and negative attitude network. First position indicates the tie of the individual on the left, the second, the individual on the right. *p*-values from Mahalanobis distance between all structures in simulation and observed data.

A.12 MAIN GOF STATISTICS, BY WAVE

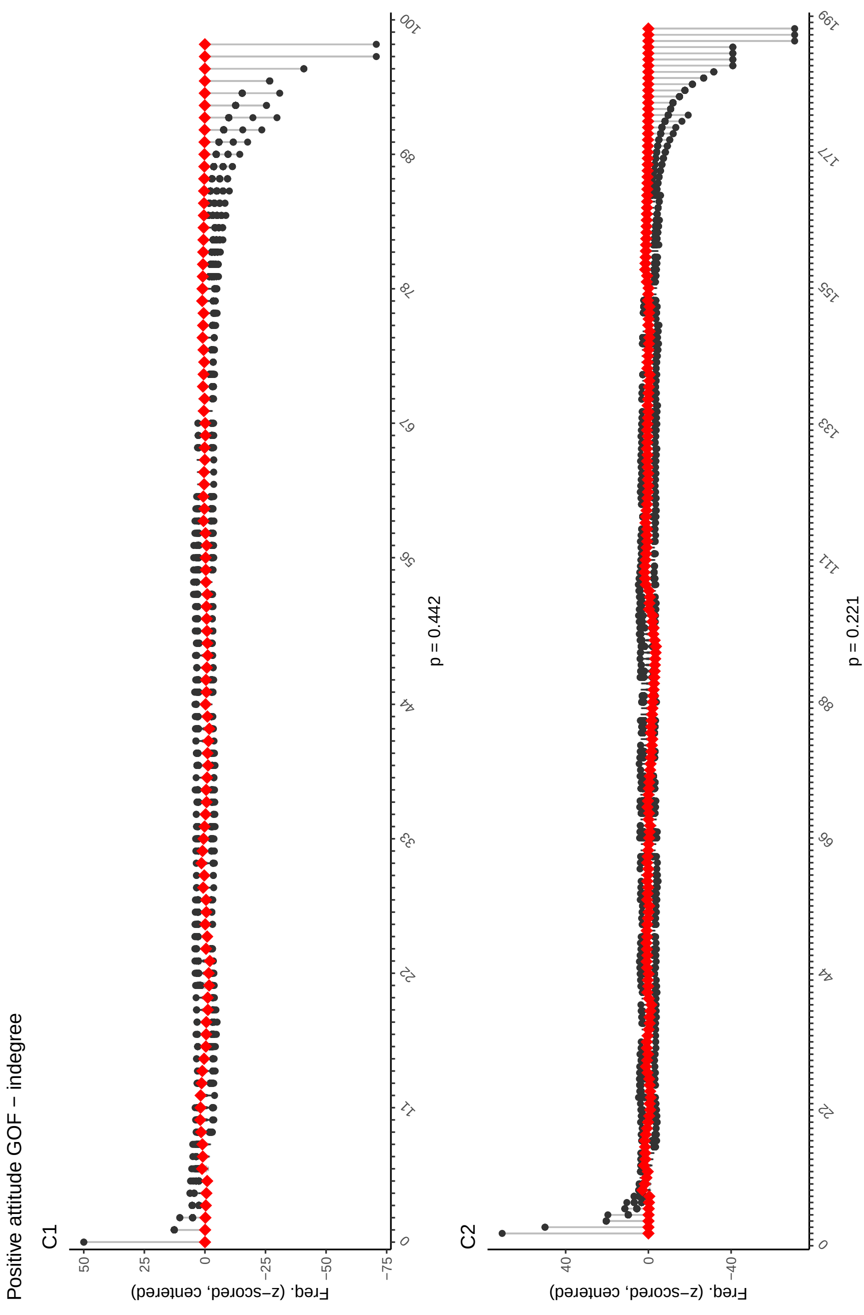


A.13 OTHER GOODNESS OF FIT STATISTICS FOR MAIN MODELS





Attitudes – outdegree
C1, positive
p = 0.167
C1, negative
p = 0.116
C2, positive
p = 0
C2, negative
p = 0.001
Freq. (z−scored, centered)


Positive attitude GOF – indegree
C1
Freq. (z–scored, centered)
p = 0.442
C2
Freq. (z–scored, centered)
p = 0.221

B

APPENDIX FOR CHAPTER 4

B.1 ISSUE TEXTS

Energy law

Below you will find the Article 1 of the said Energy Law: “1. This law shall contribute to an adequate, diversified, secure, economic and environmentally sound energy supply.2. It aims to:a. ensure the economic and environmentally sound provision and distribution of energy;b. ensure the economical and efficient use of energy; c. promote the transition to an energy supply based more on the use of renewable energies, in particular indigenous renewable energies."

Proponents of the Energy Act argue that its implementation will strengthen the Swiss economy, reduce foreign dependence, increase security of supply and improve the environmental balance.

Opponents argue that the Energy Law will make energy unaffordable, reduce security of supply, jeopardize jobs and prosperity, increase foreign dependence, and lead to more bureaucracy, bans, and landscape blight.

Pension reform

The reform of the old-age pension 2020 can only come into force if the additional financing by increasing the value added tax also comes into force.

The reform is intended to safeguard pensions and adapt old-age provision to social developments. Thanks to measures in occupational pension provision and an increase in new AHV old-age pensions of CHF 70 per month, the level of old-age pensions is to be maintained. The retirement age for women will be gradually raised from today’s 64 to 65. The reform allows flexible retirement between 62 and 70. This is to be financed by an increase in value-added tax of between 0.3 and 1 percent.

Proponents of the referenda argue that the level of pensions will be maintained, that the financing of the AHV will be secured, that unfair redistribution will be greatly reduced, that it is an important social advance, and that it is an adaptation to societal changes.

Opponents argue that the retirement age for women will be unacceptably increased, that the reform will not bring any improvement to current retirees, and that the reform in no way guarantees future retirees the level of their pensions.

TV license fee

The radio and television fee (Billag) must be paid by every Swiss household and company with radio and/or television access. The fees amount to CHF 450 per year and household and are mainly used to finance SRF (Swiss Radio and Television).

In favor of the acceptance of the referendum to abolish the radio and television fees it is argued that individual freedom of decision is promoted regarding how much money is spent on which media, companies are financially relieved, and more individual funds are available for the promotion of the national economy.

Against the acceptance of the referendum it is argued that the political and private independence of the SRF is endangered and that the reporting in marginal and smaller language regions of Switzerland remains not guaranteed.

Full money initative

The full money initiative proposes that only the Swiss National Bank (SNB) should be allowed to create money, while commercial banks would no longer be allowed to do so. In addition, the SNB should put money into circulation "debt-free", i.e. without any backing, by distributing it directly to the federal government, the cantons or the population.

The arguments in favor of accepting the referendum on the full money initiative are that money in payment accounts will become as secure as cash, that financial bubbles will be better prevented, that the state will no longer have to bail out the banks with billions in taxes, and that the SNB will be able to pay out additional money production proceeds to the population.

The arguments against accepting the referendum are that a full money system is a risky experiment, that financial services will become more expensive, that the initiative forces the National Bank to put new money into circulation without backing which puts it under political pressure, and that full money could not have prevented the earlier financial crisis.

Gambling act

The new Gambling Act, against which the current referendum is being held, implements the constitutional article on gambling that was approved by the people and the cantons on March 11, 2012. Online gambling games such as poker, blackjack or roulette will be permitted if offered by casinos based in Switzerland. Foreign providers will be blocked. Small poker tournaments outside casinos are now permitted with permission. Gambling winnings of up to CHF 1 million are no longer subject to tax.

In favor of the adoption of the Gambling Act is that it ensures that in the future all online gambling providers will pay taxes for the common good, that prevention and protection against gambling addiction will be strengthened, that the fight against money laundering and manipulation in sports betting will be intensified, and that it makes most gambling winnings tax-free.

Against the acceptance of the gambling law speaks that it introduces an internet censorship through internet blocks, that it pushes online gamblers to the black market, that the tax reduction in the gambling sector increases the risk of addiction for gamblers, that it increases tax losses, and that it offers an insufficient prevention of gambling addiction.

Food sovereignty

The referendum for food sovereignty demands that imports of non-sustainably produced food be subject to additional tariffs, that "fair prices" be set, and it demands measures so that more people can be employed in agriculture. In addition, genetically modified organisms are to be banned.

Proponents of the food sovereignty referendum argue that it will help strengthen domestic production, strengthen the rural agricultural sector, create a transparent and profitable domestic market, increase the value of agricultural workers, create a fair international market, and increase the number of agricultural workers.

Opponents of the referendum argue that new import tariffs are not against the law, that they would result in higher costs and prices, and that because of more employees, the added value must be distributed among more people.

Fair food initiative

The Fair Food Initiative wants the federal government to strengthen the supply of food that is of good quality and that has been produced in an environmentally and resource-friendly manner, animal-friendly and under fair working conditions. The federal government sets requirements for production and processing, favors imported products from fair trade and land-based farms, and ensures that negative environmental impacts of food transportation and storage are reduced.

Supporters of adopting the Fair Food Initiative argue that it improves animal welfare, supports fair trade, protects biodiversity and the climate, and curbs food waste.

Opponents of the Fair Food Initiative argue that it brings in too much bureaucracy, endangers professions in the food industry, makes food more expensive, restricts freedom of choice, breaks international commitments, and is redundant because of existing domestic laws.

Bike initiative

The Velo-Initiative aims to achieve that more bicycle paths are created and operated. To this end, the constitutional article on footpaths and hiking trails is to be expanded to include the term "cycle paths".

Proponents of the Velo-Initiative argue that this will create more safety in car, bicycle and pedestrian traffic, less congestion and more space for public transport, and that tourism will be supported. In addition, money will be saved in the planning, construction, maintenance, and operation of the pedestrian, hiking, and bicycle networks. Furthermore, the Velo-Initiative promotes cycling.

Opponents of the Velo-Initiative argue that the financial expenditures for this project would be too high (and would see these funds better invested elsewhere).

B.2 FRIENDSHIP NETWORK DESCRIPTIVES

TABLE B.1: Raw friendship network descriptives by (post-referendum) wave and cohort

Cohort	Obs. month	N^a	N ties	N isolated	Mean outdegree	Max. outdegree	Reciprocity	Cluster coeff.	Jaccard t to $t-1$	N miss.
C1	May '17	194	556	93	4.31	15	0.48	0.39	0.58	65
	Dec '17	133	296	149	4.11	10	0.52	0.46	0.73	61
	March '18	135	267	156	3.76	10	0.56	0.43	0.58	64
	Aug '18	132	224	159	3.67	10	0.47	0.44	0.70	71
	Dec '18	142	166	172	3.25	10	0.60	0.40	0.57	91
C2	March '18	215	637	81	4.52	20	0.50	0.37	0.51	74
	Aug '18	210	411	105	4.11	11	0.49	0.35	0.61	110
	Dec '18	137	251	159	3.59	11	0.60	0.37	0.56	67
C3	March '18	620	1,684	137	4.73	17	0.49	0.38	0.53	264
	Aug '18	603	1,045	221	4.98	20	0.48	0.31	0.62	393
	Dec '18	339	604	386	4.00	17	0.55	0.42	0.61	188

Note. [a]Estimated number of valid cohort members based on university records.

B.3 FULL MODELS WITH FRIENDSHIP NETWORKS

TABLE B.3: Logit regression of friend alter choice to vote on ego choice to vote

	Model 1a	Model 1b
Intercept	-1.18***	-1.76***
Prop. voting alters[a]	0.82**	0.71*
Number of alters	0.08**	0.07**
Ego choice to vote[a]	–	1.24***
Ego knowledge[a]	0.41***	0.35***
Political interest	0.84***	0.7***
Cohort 2 (ref 1)	0.27***	0.15***
Cohort 3 (ref 1)	-0.5	-0.61
Accuracy	0.78	0.8
AIC	763.72	703.16
McFadden's R^2	0.19	0.21
Valid cases	825	792
F1 score	0.86	0.88
Fraction observed to vote	0.87	0.87
Fraction predicted to vote	0.75	0.76
Permutations	5000	5000

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values [a]Variable constructed from individual data at t-1.

TABLE B.4: Logit regression of friend alter support on ego choice of vote

	Model estimate							
	2a	2b	2c	2d	2e	2f	2g	2h
Intercept	0.67***	0.43***	0.58	0.21**	0.59**	0.17**	0.89	0.29
Mean alter issue support[a]	0.4***	0.31**	0.62*	0.69**	0.36**	0.62*	0.7	0.8*
Ego support[a]	–	1.2***	–	1.2***	–	1.19***	–	1.19***
Ego knowledge[a]	–	–	-0.02	0.06	–	0.07	0	0.05
Ego undecided[a]	–	–	–	–	0.02	-0.01	-0.85	-0.62
Ego undec.*mean alter support	–	–	–	–	0.06	0.16	0.43	0.34
Mean alter support*ego knowl.	–	–	-0.07	-0.12	–	-0.11	-0.05	-0.06
Political interest	–	–	–	–	–	–	-0.18	-0.03
Ego undec.*political interest	–	–	–	–	–	–	0.51	0.36
Mean alter support*political interest	–	–	–	–	–	–	-0.09	-0.2
Mean alter support*ego undec.*political interest	–	–	–	–	–	–	-0.28	-0.1
Tax for pension	-0.12***	-0.58***	-0.13**	-0.6***	-0.15**	-0.78***	-0.16*	-0.77***
TV license	-2.38**	-2.33*	-2.24**	-2.32*	-2.29*	-2.27*	-2.27*	-2.25*
Energy law	1.44***	2***	1.75**	1.99***	1.61**	1.88***	1.66*	1.87***
Food sovereignty	-2.13**	-2.24	-2.03**	-2.18	-2.1	-2.15	-2.01	-2.1
Fair food initiative	-1.56	-2.14***	-1.49	-2.13***	-1.54	-2.12**	-1.48	-2.09**
Gambling act	-0.74***	-0.56**	-0.54**	-0.58*	-0.6	-0.56	-0.56	-0.51*
Bike initiative	0.49***	0.4***	0.8**	0.43***	0.76	0.43***	0.83	0.48***
Full money initative	-2.31***	-1.52***	-2.13**	-1.53***	-2.19*	-1.49***	-2.16*	-1.46***
Cohort 2 (ref 1)	-0.13***	0.03***	-0.2	0.02***	-0.2	0.02***	-0.19	0.01***
Cohort 3 (ref 1)	0.29**	0.13***	0.21	0.11***	0.2*	0.11***	0.19	0.08***
Accuracy	0.76	0.85	0.78	0.86	0.78	0.85	0.78	0.86
AIC	1223.73	688.76	1025.76	689.89	1009.54	678.62	1015.24	683.73
McFadden's R^2	0.24	0.51	0.26	0.51	0.25	0.51	0.26	0.51
Valid cases	1192	1016	1021	1016	1000	996	1000	996
F1 score	0.65	0.8	0.67	0.8	0.66	0.8	0.67	0.8
Fraction observed to vote	0.30	0.36	0.31	0.36	0.30	0.36	0.30	0.36
Fraction predicted to vote	0.38	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Permutations	5000	5000	5000	5000	5000	5000	5000	5000

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values.

[a]Variable constructed from individual data at t-1

B.4 FULL MODELS WITH DISCUSSION NETWORKS

TABLE B.5: Full logit regression of discussant choice to vote on ego choice to vote

	Model 1a	Model 1b
Intercept	-0.7***	-1.36***
Prop. voting alters[a]	0.29	0.2
Number of alters	0.15*	0.05
Ego choice to vote[a]	–	1.67***
Ego knowledge[a]	0.39***	0.32***
Political interest	0.72***	0.64***
Cohort 2 (ref 1)	0.28***	0.03***
Cohort 3 (ref 1)	-0.62	-0.89***
Accuracy	0.82	0.84
AIC	533.21	473.93
McFadden's R^2	0.17	0.22
Valid cases	629	608
F1 score	0.9	0.91
Fraction observed to vote	0.92	0.91
Fraction predicted to vote	0.80	0.81
Permutations	5000	5000

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. [a]Variable constructed from individual data at t-1 [b]Here, 96.5% of individuals are predicted to vote by the model, while 80.8% did so.

TABLE B.6: Full logit regression of discussant alter support on ego choice of vote

	Model estimate							
	2a	2b	2c	2d	2e	2f	2g	2h
Intercept	0.85**	0.72**	0.73	0.59	0.79**	0.64	0.94	0.36
Mean alter issue support[a]	0.25**	0.16	0.05	-0.09	0.32**	-0.17	0.08	-0.19
Ego support[a]	–	1.35***	–	1.36***	–	1.36***	–	1.38***
Ego knowledge[a]	–	–	0.01	0.06	–	0.09	0	0.01
Ego undecided[a]	–	–	–	–	0.01	0.24	-0.82	-0.1
Ego undec.*mean alter support	–	–	–	–	-0.14	0.14	0.39	0.48
Mean alter support*ego knowl.	–	–	0.06	0.09*	–	0.1*	0.08	0.13*
Political interest	–	–	–	–	–	–	-0.06	0.33
Ego undec.*political interest	–	–	–	–	–	–	0.51	0.17
Mean alter support*political interest	–	–	–	–	–	–	-0.04	-0.07
Mean alter support*ego undec.*political interest	–	–	–	–	–	–	-0.26	-0.2
Tax for pension	-0.1***	-0.52***	-0.09	-0.52***	-0.18**	-0.79***	-0.16	-0.8***
TV license	-2.56**	-2.8	-2.38**	-2.85	-2.41**	-2.99	-2.41*	-2.98
Energy law	1.19***	1.46	1.5	1.37*	1.42**	1.2**	1.4	1.24*
Food sovereignty	-2.02**	-2.43***	-1.96**	-2.41***	-2.04	-2.71*	-2.01	-2.66*
Fair food initiative	-1.49*	-2.42***	-1.47*	-2.39***	-1.55	-2.69***	-1.51	-2.66***
Gambling act	-0.76**	-0.65	-0.49**	-0.64	-0.59*	-0.92	-0.53*	-0.82
Bike initiative	0.47**	0.39***	0.86	0.41***	0.77	0.1***	0.82	0.15***
Full money initative	-2.51**	-1.74***	-2.21**	-1.76***	-2.28*	-2.03***	-2.28*	-2.04**
Cohort 2 (ref 1)	-0.31**	-0.1***	-0.42*	-0.14***	-0.41*	-0.13***	-0.46*	-0.15***
Cohort 3 (ref 1)	0.12**	-0.11***	0**	-0.18***	0.02**	-0.18***	-0.03*	-0.21***
Accuracy	0.77	0.87	0.79	0.87	0.78	0.86	0.79	0.86
AIC	1010.26	529.53	850.57	531.08	835.07	521.85	841.15	524.31
McFadden's R^2	0.24	0.55	0.27	0.55	0.26	0.55	0.27	0.56
Valid cases	986	846	847	846	831	830	831	830
F1 score	0.67	0.82	0.69	0.82	0.68	0.81	0.69	0.81
Fraction observed to vote	0.31	0.37	0.32	0.36	0.31	0.37	0.31	0.36
Fraction predicted to vote	0.38	0.37	0.37	0.37	0.37	0.37	0.37	0.37
Permutations	5000	5000	5000	5000	5000	5000	5000	5000

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. [a]Variable constructed from individual data at t-1

B.5 CHOICE TO VOTE MODEL WITH DEMOGRAPHICS

TABLE B.7: Logit regression of alter support on ego choice of vote, with demographics

	Model 1a	Model 1b
Intercept	-0.53***	-1.2***
Prop. voting alters[a]	0.12	0.04
Number of alters	0.15*	0.07
Ego choice to vote[a]	–	1.62***
Ego knowledge[a]	0.42***	0.34***
Political interest	0.72***	0.65***
Female (ref male)	0.53*	0.55*
Econ. status	-0.4	-0.35
French canton (ref. German)	0.85	0.95
Italian canton	-0.44	-0.32
Rumantsch canton	-0.67*	-0.65*
Accuracy	0.83	0.86
AIC	526.44	470.94
McFadden's R^2	0.19	0.24
Valid cases	628	607

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. Fixed effects for cohort excluded. [a]Variable constructed from individual data at t-1

B.6 MODELS WITH DISCUSSION GROUPS

Table B.8: Logit regression of group discussant alter support on ego choice of vote

	Model 1a	Model 1b
Intercept	-1.09***	-1.61***
Prop. voting alters[a]	0.32	0.18
Number of alters	0.07*	0.06
Ego choice to vote[a]	–	0.97*
Ego knowledge[a]	0.44***	0.41***
Political interest	1***	0.88***
Accuracy	0.81	0.8
AIC	314.79	304.02
McFadden's R^2	0.22	0.23
Valid cases	361	353

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. Fixed effects for cohort excluded. [a]Variable constructed from individual data at t-1

TABLE B.9: Logit regression of group discussant alter support on ego choice to vote

	Model estimate							
	2a	2b	2c	2d	2e	2f	2g	2h
Intercept	-2.2	-2.05*	-2.31	-2.41	-2.15	-2.26	-2.16	-2.55
Mean alter issue support[a]	0.24	0.31	0.19	0.4	0.39*	0.48	-0.02	0.42
Ego support[a]	–	1.21***	–	1.2***	–	1.2***	–	1.22***
Ego knowledge[a]	–	–	0.03	0.09	–	0.07	-0.06	-0.04
Ego undecided[a]	–	–	–	–	-0.31	-0.24	-1.53*	-1.19
Ego undec.*mean alter support	–	–	–	–	-0.13	-0.06	1.11	0.64
Mean alter support*ego knowl.	–	–	0.04	-0.02	–	-0.04	-0.02	-0.04
Political interest	–	–	–	–	–	–	0.14	0.43
Ego undec.*political interest	–	–	–	–	–	–	0.65	0.47
Mean alter support*political interest	–	–	–	–	–	–	0.29	0.05
Mean alter support*ego undec.*political interest	–	–	–	–	–	–	-0.71	-0.41
Accuracy	0.78	0.86	0.79	0.85	0.79	0.86	0.79	0.86
AIC	592.69	317.35	479.59	320.32	477.46	323.04	481.68	322.68
McFadden's R^2	0.23	0.51	0.25	0.51	0.25	0.51	0.27	0.53
Valid cases	586	483	484	483	481	480	481	480

Note. Significance determined from observed statistics against percentiles of 5000 permutations, using absolute, centered values. Fixed effects for cohort excluded. [a]Variable constructed from individual data at t-1

BIBLIOGRAPHY

Abelson, R. P. (1964). Mathematical models of the distribution of attitudes under controversy. In N. Frederiksen & H. Gulliksen (Eds.), *Contributions to mathematical psychology* (pp. 142–160). Holt, Reinehart; Winston, Inc.

Abramowitz, A. I., & Saunders, K. L. (2008). Is polarization a myth? *The Journal of Politics*, *70*(2), 542.

Adams, J., Green, J., & Milazzo, C. (2012). Who moves? elite and mass-level depolarization in britain, 1987–2001. *Electoral Studies*, *31*(4), 643.

Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, *74*(1), 47.

Albi, G., Pareschi, L., & Zanella, M. (2016). On the optimal control of opinion dynamics on evolving networks. In L. Bociu, J.-A. Désidéri, & A. Habbal (Eds.), *System modeling and optimization* (pp. 58–67). Springer International Publishing.

Alford, J. R., Hatemi, P. K., Hibbing, J. R., Martin, N. G., & Eaves, L. J. (2011). The politics of mate choice. *The Journal of Politics*, *73*(2), 362.

Algan, Y., Dalvit, N., Do, Q.-A., Le Chapelain, A., & Zenou, Y. (2020). Friendship networks and political opinions: A natural experiment among future french politicians, 63.

An, J., Kwak, H., Posegga, O., & Jungherr, A. (2019). Political discussions in homogeneous and cross-cutting communication spaces. *Proceedings of the International AAAI Conference on Web and Social Media*, *13*(1), 68.

Antheunis, M. L., Valkenburg, P. M., & Peter, J. (2012). The quality of online, offline, and mixed-mode friendships among users of a social networking site. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *6*(3).

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, *41*(2), 203.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216.

Baker, A., Ames, B., & Renno, L. R. (2006). Social context and campaign volatility in new democracies: Networks and neighborhoods in brazil's 2002 elections. *American Journal of Political Science*, *50*(2), 382.

Baldassarri, D., & Bearman, P. (2007). Dynamics of political polarization. *American Sociological Review*, *72*(5), 784.

Baldassarri, D., & Gelman, A. (2008). Partisans without constraint: Political polarization and trends in american public opinion. *American Journal of Sociology*, *114*(2), 408.

Baldassarri, D., & Page, S. E. (2021). The emergence and perils of polarization. *Proceedings of the National Academy of Sciences*, *118*(50), e2116863118.

Baldassarri, D., & Park, B. (2020). Was there a culture war? partisan polarization and secular trends in US public opinion. *The Journal of Politics*, *82*(3), 809.

Bello, J. (2012). The dark side of disagreement? revisiting the effect of disagreement on political participation. *Electoral Studies*, *31*(4), 782.

Bello, J., & Rolfe, M. (2014). Is influence mightier than selection? forging agreement in political discussion networks during a campaign. *Social Networks*, *36*, 134.

Bernard, H. R., & Killworth, P. D. (1977). Informant accuracy in social network data II. *Human communication research*, *4*(1), 3.

Bhatti, Y., & Hansen, K. M. (2012). Leaving the nest and the social act of voting: Turnout among first-time voters. *Journal of Elections, Public Opinion and Parties*, *22*(4), 380.

Bishop, B., & Cushing, R. G. (2008). *The big sort: Why the clustering of like-minded america is tearing us apart*. Houghton Mifflin.

Block, P. (2015). Reciprocity, transitivity, and the mysterious three-cycle. *Social Networks*, *40*, 163.

Block, P. (2018). Network evolution and social situations. *Sociological Science*, *5*, 402.

Bochsler, D., Gerber, M., & Zumbach, D. (2016). The 2015 national elections in switzerland: Renewed polarization and shift to the right. *Regional & Federal Studies*, *26*(1), 95.

Boero, R., & Squazzoni, F. (2005). Does empirical embeddedness matter? methodological issues on agent-based models for analytical social science. *Journal of Artificial Societies and Social Simulation*, *8*(4), 6.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, *489*(7415), 295.

Bornschier, S. (2015). The new cultural conflict, polarization, and representation in the swiss party system, 1975–2011. *Swiss Political Science Review*, *21*(4), 680.

Bougher, L. D. (2017). The correlates of discord: Identity, issue alignment, and political hostility in polarized america. *Political Behavior*, *39*(3), 731.

Boutyline, A., & Vaisey, S. (2017). Belief network analysis: A relational approach to understanding the structure of attitudes. *American Journal of Sociology*, *122*(5), 1371.

Boutyline, A., & Willer, R. (2017). The social structure of political echo chambers: Variation in ideological homophily in online networks. *Political Psychology*, *38*(3), 551.

Boxell, L., Gentzkow, M., & Shapiro, J. M. (2017). Greater internet use is not associated with faster growth in political polarization among US demographic groups. *Proceedings of the National Academy of Sciences*, *114*(40), 10612.

Bramson, A., Grim, P., Singer, D. J., Fisher, S., Berger, W., Sack, G., & Flocken, C. (2016). Disambiguation of social polarization concepts and measures. *The Journal of Mathematical Sociology*, *40*(2), 80.

Brechwald, W. A., & Prinstein, M. J. (2011). Beyond homophily: A decade of advances in understanding peer influence processes. *Journal of Research on Adolescence*, *21*(1), 166.

Brewer, M. B. (2017). Intergroup discrimination: Ingroup love or outgroup hate? In *The cambridge handbook of the psychology of prejudice* (pp. 90–110). Cambridge University Press.

Bundeszentrale fur politische Bildung. (2017). *Interaktives wahltool wahl-o-mat*. Retrieved June 1, 2022, from http://www.bpb.de/politik/wahlen/wahl-o-mat.

Butts, C. T. (2008). A relational event framework for social action. *Sociological Methodology*, *38*(1), 155.

Butts, C. T. (2009). Revisiting the foundations of network analysis. *Science*, *325*(5939), 414.

Campos, C. F. S., Hargreaves Heap, S., & Leite Lopez de Leon, F. (2017). The political influence of peer groups: Experimental evidence in the classroom. *Oxford Economic Papers*, *69*(4), 963.

Carley, K. (1991). A theory of group stability. *American Sociological Review*, *56*(3), 331.

Cartwright, D., & Harary, F. (1956). Structural balance: A generalization of heider's theory. *Psychological Review*, *63*, 277.

Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, *81*(2), 591.

Chattoe-Brown, E. (2014). Using agent based modelling to integrate data on attitude change. *Sociological Research Online*, *19*(1), 16.

Chu, O. J., Donges, J. F., Robertson, G. B., & Pop-Eleches, G. (2021). The microdynamics of spatial polarization: A model and an application to survey data from ukraine. *Proceedings of the National Academy of Sciences*, *118*(50), e2104194118.

Coleman, J. S. (1990). *Foundations of social theory*. Belknap Press of Harvard University Press.

Colleoni, E., Rozza, A., & Arvidsson, A. (2014). Echo chamber or public sphere? predicting political orientation and measuring political homophily in twitter using big data. *Journal of Communication*, *64*(2), 317.

Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., & Flammini, A. (2011). Political polarization on twitter. *Proceedings of the international aaai conference on web and social media*, *5*, 89.

Converse, P. E. (1964). The nature of belief systems in mass publics. In A. David E. (Ed.), *Ideology and discontent* (pp. 206–261). The Free Press of Glencoe.

Cotton, D., & Alcock, I. (2013). Commitment to environmental sustainability in the UK student population. *Studies in Higher Education*, *38*(10), 1457.

Cowan, S. K., & Baldassarri, D. (2018). "it could turn ugly": Selective disclosure of attitudes in political discussion networks. *Social Networks*, *52*, 1.

Dalege, J., Borsboom, D., Van Harreveld, F., Van den Berg, H., Conner, M., & Van der Maas, H. L. (2016). Toward a formalized account of attitudes: The causal attitude network (CAN) model. *Psychological review*, *123*(1), 2.

Davis, A., Gardner, B. B., & Gardner, M. R. (1941). *Deep south: A social anthropological study of caste and class*. University of Chicago Press.

Davis, J. A. (1970). Clustering and hierarchy in interpersonal relations: Testing two graph theoretical models on 742 sociomatrices. *American Sociological Review*, *35*(5), 843.

Deffuant, G., Neau, D., Amblard, F., & Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, *03*(1), 87.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, *69*(345), 118.

Dekker, D., Krackhardt, D., & Snijders, T. A. B. (2007). Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. *Psychometrika*, *72*(4), 563.

DellaPosta, D. (2020). Pluralistic collapse: The "oil spill" model of mass opinion polarization. *American Sociological Review*, 000312242092298.

DellaPosta, D., & Macy, M. W. (2015). The center cannot hold. In *Order on the edge of chaos* (p. 86). Cambridge University Press.

DellaPosta, D., Shi, Y., & Macy, M. (2015). Why do liberals drink lattes? *American Journal of Sociology*, *120*(5), 1473.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, *51*(3), 629.

Dias, N., & Lelkes, Y. (2022). The nature of affective polarization: Disentangling policy disagreement from partisan identity. *American Journal of Political Science*, *66*(3), 775.

DiMaggio, P., Evans, J., & Bryson, B. (1996). Have american's social attitudes become more polarized? *American Journal of Sociology*, *102*(3), 690.

Dinkelberg, A., O'Reilly, C., MacCarron, P., Maher, P. J., & Quayle, M. (2021). Multidimensional polarization dynamics in US election data in the long term (2012–2020) and in the 2020 election cycle. *Analyses of Social Issues and Public Policy*, *21*(1), 284.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, *65*(2), 135.

Druckman, J. N., Klar, S., Krupnikov, Y., Levendusky, M., & Ryan, J. B. (2020). Affective polarization, local contexts and public opinion in america. *Nature Human Behaviour*.

Druckman, J. N., & Levendusky, M. S. (2019). What do we measure when we measure affective polarization? *Public Opinion Quarterly*, *83*(1), 114.

Druckman, J. N., Peterson, E., & Slothuus, R. (2013). How elite partisan polarization affects public opinion formation. *American Political Science Review*, *107*(1), 57.

Dubois, E., & Blank, G. (2018). The echo chamber is overstated: The moderating effect of political interest and diverse media. *Information, Communication & Society*, *21*(5), 729.

Duggins, P. (2017). A psychologically-motivated model of opinion change with applications to american politics. *Journal of Artificial Societies and Social Simulation*, *20*(1), 13.

Durrett, R., Gleeson, J. P., Lloyd, A. L., Mucha, P. J., Shi, F., Sivakoff, D., Socolar, J. E. S., & Varghese, C. (2012). Graph fission in an evolving voter model. *Proceedings of the National Academy of Sciences*, *109*(10), 3682.

Duxbury, S. W. (2021). The problem of scaling in exponential random graph models. *Sociological Methods & Research*, 0049124120986178.

Eagly, A., & Chaiken, S. (1998). Attitude structure and function. *Handbook of social psychology*, *1*, 269.

Elmer, T., Mepham, K., Hoffman, M., Boda, Z., Raabe, I. J., Vörös, A., & Stadtfeld, C. (2022, Feb. 18). The swiss student life study: Codebook.

Elmer, T., & Stadtfeld, C. (2020). Depressive symptoms are associated with social isolation in face-to-face interaction networks. *Scientific Reports*, *10*(1), 1444.

Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, *4*(5), 41.

European Social Survey ERIC. (2020). ESS-8 2016 documentation report.

Eveland Jr, W. P., & Hively, M. H. (2009). Political discussion frequency, network size, and "heterogeneity" of discussion as predictors of political knowledge and participation. *Journal of Communication*, *59*(2), 205.

Federal Statistical Office. (2017). *Votes: Year 2017*. Retrieved Aug. 29, 2022, from https://www.bfs.admin.ch/bfs/en/home/statistiken/politik/abstimmungen/jahr-2017.html

Federal Statistical Office. (2018). *Votes: Year 2018*. Retrieved Aug. 29, 2022, from https://www.bfs.admin.ch/bfs/en/home/statistiken/politik/abstimmungen/jahr-2018.html

Federal Statistical Office. (2021). *Votes: Year 2021*. Retrieved Aug. 29, 2022, from https://www.bfs.admin.ch/bfs/en/home/statistics/politics/popular-votes/year-2021.html

Festinger, L. (1950). Informal social communication. *Psychological Review*, *57*, 271.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Fiorina, M. P., & Abrams, S. J. (2008). Political polarization in the american public. *Annual Review of Political Science*, *11*(1), 563.

Fischer, C. S. (1982). What do we mean by 'friend'? an inductive study. *Social Networks*, *3*(4), 287.

Flache, A., & Macy, M. W. (2011). Small worlds and cultural polarization. *The Journal of Mathematical Sociology*, *35*(1), 146.

Flache, A., Mäs, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., & Lorenz, J. (2017). Models of social influence: Towards the next frontiers. *Journal of Artificial Societies and Social Simulation*, *20*(4), 2.

French, J. R. (1956). Formal theory of social power. *Psychological Review*, *63*(3), 181.

Friedkin, N. E. (1999). Choice shift and group polarization. *American Sociological Review*, 856.

Friese, M., Smith, C. T., Plischke, T., Bluemke, M., & Nosek, B. A. (2012). Do implicit attitudes predict actual voting behavior particularly for undecided voters? *PLOS ONE*, *7*(8), e44130.

Garimella, V. R. K., & Weber, I. (2017). A long-term analysis of polarization on twitter. *Eleventh International AAAI Conference on Web and Social Media*.

Gestefeld, M., Lorenz, J., Boehnke, K., & Henschel, N. T. (2021, Mar. 30). *Decomposing attitude distributions to characterize attitude polarization in europe* (preprint). SocArXiv.

Giardini, F., Vilone, D., et al. (2021). Opinion dynamics and collective risk perception: An agent-based model of institutional and media communication about disasters. *Journal of Artificial Societies and Social Simulation*, *24*(1), 1.

Gidron, N., Adams, J., & Horne, W. (2020). American affective polarization in comparative perspective. In F. E. Lee (Ed.), *Elements in american politics*. Cambridge University Press.

Goel, S., Mason, W., & Watts, D. J. (2010). Real and perceived attitude agreement in social networks. *Journal of Personality and Social Psychology*, *99*(4), 611.

Goldberg, A., & Stein, S. K. (2018). Beyond social contagion: Associative diffusion and the emergence of cultural variation. *American Sociological Review*, *83*(5), 897.

Gonthier, F., & Guerra, T. (2022). How party polarization shapes the structuring of policy preferences in europe. *Party Politics*, 13540688211064606.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360.

Guerra, P., Jr, W. M., Cardie, C., & Kleinberg, R. (2013). A measure of polarization on social media networks based on community boundaries. *Proceedings of the International AAAI Conference on Web and Social Media*, *7*(1), 215.

Guess, A. M. (2021). (almost) everything in moderation: New evidence on americans' online media diets. *American Journal of Political Science*, *65*(4), 1007.

Hänggli, R., & Häusermann, S. (2015). Consensus lost? disenchanted democracy in switzerland. *Swiss Political Science Review*, *21*(4), 475.

Harrell, F. E. (2001). *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis* (Vol. 608). Springer.

Harrigan, N. M., Labianca, G., & Agneessens, F. (2020). Negative ties and signed graphs research: Stimulating research on dissociative forces in social networks. *Social Networks*, *60*, 1.

Hedström, P., & Swedberg, R. (1998). *Social mechanisms: An analytical approach to social theory*. Cambridge University Press.

Hegselmann, R., Krause, U., et al. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, *5*(3).

Heider, F. (1946). Attitudes and cognitive organization. *The Journal of psychology*, *21*(1), 107.

Hetherington, M. J. (2009). Putting polarization in perspective. *British Journal of Political Science*, 413.

Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology*, *53*(1), 575.

Hill, R. A., & Dunbar, R. I. (2003). Social network size in humans. *Human nature*, *14*(1), 53.

Himelboim, I., McCreery, S., & Smith, M. (2013). Birds of a feather tweet together: Integrating network and content analyses to examine cross-ideology exposure on twitter. *Journal of Computer-mediated Communication*, *18*(2), 154.

Hobolt, S. B., Leeper, T. J., & Tilley, J. (2021). Divided by the vote: Affective polarization in the wake of the brexit referendum. *British Journal of Political Science*, *51*(4), 1476.

Holland, P. W., & Leinhardt, S. (1970). A method for detecting structure in sociometric data. *American Journal of Sociology*, *76*(3), 492.

Hollway, J., Lomi, A., Pallotti, F., & Stadtfeld, C. (2017). Multilevel social spaces: The network dynamics of organizational fields. *Network Science*, *5*(2), 187.

Hopmann, D. N. (2012). The consequences of political disagreement in interpersonal communication: New insights from a comparative perspective. *European Journal of Political Research*, *51*(2), 265.

Hopmann, D. N., Matthes, J., & Nir, L. (2015). Informal political conversation across time and space: Setting the research agenda. *International Journal of Public Opinion Research*, *27*(4), 448.

Hopmann, D. N., Vliegenthart, R., De Vreese, C., & Albæk, E. (2010). Effects of election news coverage: How visibility and tone influence party choice. *Political Communication*, *27*(4), 389.

Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion*. Yale University Press.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness*. *Public Opinion Quarterly*, *15*(4), 635.

Huber, G. A., & Malhotra, N. (2016). Political homophily in social relationships: Evidence from online dating behavior. *The Journal of Politics*, *79*(1), 269.

Huckfeldt, R., & Mendez, J. M. (2008). Moths, flames, and political engagement: Managing disagreement within communication networks. *The Journal of Politics*, *70*(1), 83.

Huckfeldt, R., Mendez, J. M., & Osborn, T. (2004). Disagreement, ambivalence, and engagement: The political consequences of heterogeneous networks. *Political Psychology*, *25*(1), 65.

Huckfeldt, R., & Sprague, J. (1987). Networks in context: The social flow of political information. *The American Political Science Review*, *81*(4), 1197.

Huckfeldt, R., & Sprague, J. (1995). *Citizens, politics, and social communication: Information and influence in an election campaign*. Cambridge University Press.

Huddy, L. (2001). From social to political identity: A critical examination of social identity theory. *Political Psychology*, *22*(1), 127.

Hunter, J. D. (1991). *Culture wars: The struggle to define america*. BasicBooks.

Huston, T. L., & Levinger, G. (1978). Interpersonal attraction and relationships. *Annual review of psychology*, *29*(1), 115.

International Social Survey Program Research Group. (2016). International social survey programme: Citizenship - ISSP 2014. *Cologne: GESIS Data Archive*.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the united states. *Annual Review of Political Science*, *22*(1), 129.

Iyengar, S., Sood, G., & Lelkes, Y. (2012). Affect, not ideology: A social identity perspective on polarization. *Public Opinion Quarterly*, *76*(3), 405.

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, *59*(3), 690.

Jones, J. J., Bond, R. M., Bakshy, E., Eckles, D., & Fowler, J. H. (2017). Social influence and political mobilization: Further evidence from a randomized experiment in the 2012 u.s. presidential election. *PLOS ONE*, *12*(4), e0173851.

Jost, J. T., Baldassarri, D. S., & Druckman, J. N. (2022). Cognitive–motivational mechanisms of political polarization in social-communicative contexts. *Nature Reviews Psychology*, *1*(10), 560.

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology*, *60*(1), 307.

Jost, J. T., Ledgerwood, A., & Hardin, C. D. (2008). Shared reality, system justification, and the relational basis of ideological beliefs. *Social and Personality Psychology Compass*, *2*(1), 171.

Jussim, L., & Osgood, D. W. (1989). Influence and similarity among friends: An integrative model applied to incarcerated adolescents. *Social Psychology Quarterly*, *52*(2), 98.

Kandel, D. B. (1978). Homophily, selection, and socialization in adolescent friendships. *American Journal of Sociology*, *84*(2), 427.

Kashy, D. A., & Kenny, D. A. (1990). Do you know whom you were with a week ago friday? a re-analysis of the bernard, killworth, and sailer studies. *Social Psychology Quarterly*, *53*(1), 55.

Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence: The part played by people in the flow of mass communications*. Free Press.

Kazee, T. A. (1981). Television exposure and attitude change: The impact of political interest. *Public Opinion Quarterly*, *45*(4), 507.

Killworth, P., & Bernard, H. (1976). Informant accuracy in social network data. *Human Organization*, *35*(3), 269.

Kitts, J. A. (2003). Egocentric bias or information management? selective disclosure and the social roots of norm misperception. *Social Psychology Quarterly*, *66*(3), 222.

Klofstad, C. A. (2007). Talk leads to recruitment: How discussions about politics and current events increase civic participation. *Political Research Quarterly*, *60*(2), 180.

Klofstad, C. A. (2015). Exposure to political discussion in college is associated with higher rates of political participation over time. *Political Communication*, *32*(2), 292.

Klofstad, C. A., McDermott, R., & Hatemi, P. K. (2013). The dating preferences of liberals and conservatives. *Political Behavior*, *35*(3), 519.

Klofstad, C. A., Sokhey, A. E., & McClurg, S. D. (2013). Disagreeing about disagreement: How conflict in social networks affects political behavior. *American Journal of Political Science*, *57*(1), 120.

Kozlowski, A. C., & Murphy, J. P. (2021). Issue alignment and partisanship in the american public: Revisiting the 'partisans without constraint' thesis. *Social Science Research*, *94*, 102498.

Krackhardt, D. (1988). Predicting with networks: Nonparametric multiple regression analysis of dyadic data. *Social Networks*, *10*(4), 359.

Krackhardt, D., Nohria, N., & Eccles, B. (2003). The strength of strong ties. *Networks in the knowledge economy*, *82*.

Krause, R. W., Huisman, M., Steglich, C., & Snijders, T. (2020). Missing data in cross-sectional networks – an extensive comparison of missing data treatment methods. *Social Networks*, *62*, 99.

Lau, R. R., & Redlawsk, D. P. (2001). Advantages and disadvantages of cognitive heuristics in political decision making. *American Journal of Political Science*, *45*(4), 951.

Laumann, E. O. (1969). Friends of urban men: An assessment of accuracy in reporting their socioeconomic attributes, mutual choice, and attitude agreement. *Sociometry*, *32*(1), 54.

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice: How the voter makes up his mind in a presidential campaign.* (1st ed.). Columbia University Press.

Lazega, E., & Snijders, T. A. (2015). *Multilevel network analysis for the social sciences: Theory, methods and applications* (Vol. 12). Springer.

Lazer, D., Rubineau, B., Chetkovich, C., Katz, N., & Neblo, M. (2010). The coevolution of networks and political attitudes. *Political Communication*, *27*(3), 248.

LeDuc, L. (2007). Opinion formation and change in referendum campaigns. In C. de Vreese (Ed.), *The dynamics of referendum campaigns* (pp. 21–44). Springer.

Lelkes, Y. (2016). Mass polarization: Manifestations and measurements. *Public Opinion Quarterly*, *80*, 392.

Levendusky, M. S. (2009). The microfoundations of mass polarization. *Political Analysis*, *17*(2), 162.

Levitan, L. C., & Visser, P. S. (2009). Social network composition and attitude strength: Exploring the dynamics within newly formed social networks. *Journal of Experimental Social Psychology*, *45*(5), 1057.

Lewis, K., & Kaufman, J. (2018). The conversion of cultural tastes into social network ties. *American Journal of Sociology*, *123*(6), 1684.

Lizardo, O. (2006). How cultural tastes shape personal networks. *American Sociological Review*, *71*(5), 778.

Lomi, A., & Stadtfeld, C. (2014). Social networks and social settings: Developing a coevolutionary view. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie*, *66*(1), 395.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, *37*, 2098.

Lorenz, J. (2021). Data-driven agent-based modeling in computational social science. In U. Engel, A. Quan-Haase, S. Liu, & L. E. Lyberg (Eds.), *Handbook of computational social science, volume 1* (pp. 150–167). Routledge.

Lospinoso, J., & Snijders, T. A. (2019). Goodness of fit for stochastic actor-oriented models. *Methodological Innovations*, *12*(3), 2059799119884282.

Lupu, N. (2015). Party polarization and mass partisanship: A comparative perspective. *Political Behavior*, *37*(2), 331.

Macy, M. W., & Willer, R. (2002). From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology*, *28*(1), 143.

Manzo, G. (2007). Variables, Mechanisms, and Simulations: Can the Three Methods Be Synthesized? A Critical Analysis of the Literature. *Revue française de sociologie*, *48*(5), 35.

Marsden, P. V. (1987). Core discussion networks of americans. *American Sociological Review*, *52*(1), 122.

Mäs, M. (2019). Challenges to simulation validation in the social sciences. a critical rationalist perspective. In C. Beisbart & N. J. Saam (Eds.), *Computer simulation validation: Fundamental concepts, methodological frameworks, and philosophical perspectives* (pp. 857–879). Springer International Publishing.

Mäs, M., Flache, A., Takács, K., & Jehn, K. A. (2013). In the short term we divide, in the long term we unite: Demographic crisscrossing and the effects of faultlines on subgroup polarization. *Organization Science*, *24*(3), 716.

Matthes, J., Knoll, J., Valenzuela, S., Hopmann, D. N., & Von Sikorski, C. (2019). A meta-analysis of the effects of cross-cutting exposure on political participation. *Political Communication*, *36*(4), 523.

McClurg, S. D. (2003). Social networks and political participation: The role of social interaction in explaining political participation. *Political Research Quarterly*, *56*(4), 449.

McClurg, S. D. (2006). The electoral relevance of political talk: Examining disagreement and expertise effects in social networks on political participation. *American Journal of Political Science*, *50*(3), 737.

McPherson, J. M., & Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review*, *52*(3), 370.

McPherson, J. M., Smith-Lovin, L., & Cook, J. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415.

Mepham, K., Vörös, A., & Stadtfeld, C. (2022). Network polarization: The study of political attitudes and social ties as dynamic multilevel networks.

Mesch, G. S., & Talmud, I. (2007). Similarity and the quality of online and offline social relationships among adolescents in israel. *Journal of Research on Adolescence*, *17*(2), 455.

Minozzi, W., Song, H., Lazer, D. M. J., Neblo, M. A., & Ognyanova, K. (2020). The incidental pundit: Who talks politics with whom, and why? *American Journal of Political Science*, *64*(1), 135.

Morey, A. C., Eveland, W. P., & Hutchens, M. J. (2012). The "who" matters: Types of interpersonal relationships and avoidance of political disagreement. *Political Communication*, *29*(1), 86.

Munzert, S., & Bauer, P. C. (2013). Political depolarization in german public opinion, 1980–2010*. *Political Science Research and Methods*, *1*(1), 67.

Mutz, D. C. (2002). The consequences of cross-cutting networks for political participation. *American Journal of Political Science*, *46*(4), 838.

Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press.

Newcomb, T. M. (1961). The acquaintance process as a prototype of human interaction. In *The acquaintance process.* (pp. 259–261). Holt, Rinehart & Winston.

Nickerson, D. W. (2008). Is voting contagious? evidence from two field experiments. *American Political Science Review*, *102*(1), 49.

Nieuwbeerta, P., & Flap, H. (2000). Crosscutting social circles and political choice: Effects of personal network composition on voting behavior in the netherlands. *Social Networks*, *22*(4), 313.

Nir, L. (2011). Disagreement and opposition in social networks: Does disagreement discourage turnout? *Political Studies*, *59*(3), 674.

Nocaj, A., Ortmann, M., & Brandes, U. (2015). Untangling the hairballs of multi-centered, small-world online social media networks. *Journal of Graph Algorithms and Applications*, *19*(2), 595.

Noelle-Neumann, E. (1974). The spiral of silence: A theory of public opinion. *Journal of Communication*, *24*(2), 43.

Ohme, J., de Vreese, C. H., & Albaek, E. (2018). The uncertain first-time voter: Effects of political media exposure on young citizens' formation of vote choice in a digital media environment. *New Media & Society*, *20*(9), 3243.

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. New York, NY: Penguin.

Parker, S. L., Parker, G. R., & McCann, J. A. (2008). Opinion taking within friendship networks. *American Journal of Political Science*, *52*(2), 412.

Pattie, C. J., & Johnston, R. J. (2009). Conversation, disagreement and political participation. *Political Behavior*, *31*(2), 261.

Perrett, S. (2021). A divided kingdom? variation in polarization, sorting, and dimensional alignment among the british public, 1986–2018. *The British Journal of Sociology*, *72*(4), 992.

Pettigrew, T. F. (1998). Intergroup contact theory. *Annual Review of Psychology*, *49*(1), 65.

Pew Research Center. (2014, June). *Political polarization in the american public*. Pew Research Center.

Praet, S., Guess, A. M., Tucker, J. A., Bonneau, R., & Nagler, J. (2021). What's not to like? facebook page likes reveal limited polarization in lifestyle preferences. *Political Communication*, 1.

Price, D. D. S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, *27*(5), 292.

Raabe, I. J., Boda, Z., & Stadtfeld, C. (2019). The social pipeline: How friend influence and peer exposure widen the STEM gender gap. *Sociology of Education*, *92*(2), 105.

Raccuia, L. (2016). Single-target implicit association tests (ST-IAT) predict voting behavior of decided and undecided voters in swiss referendums. *PLOS ONE*, *11*(10), e0163872.

Rashotte, L. (2007). Social influence. In G. Ritzer (Ed.), *The blackwell encyclopedia of sociology* (1st ed., pp. 4434–4437). Blackwell Publishing.

Rawlings, C. M. (2020). Cognitive authority and the constraint of attitude change in groups. *American Sociological Review*, *85*(6), 992.

Rawlings, C. M. (2022). Becoming an ideologue: Social sorting and the microfoundations of polarization. *Sociological Science*, *9*, 313.

Reiljan, A. (2020). 'fear and loathing across party lines' (also) in europe: Affective polarisation in european party systems. *European Journal of Political Research*, *59*(2), 376.

Richey, S. (2008). The social basis of voting correctly. *Political Communication*, *25*(4), 366.

Ripley, R. M., Snijders, T. A., Boda, Z., Vörös, A., & Preciado, P. (2019, Jan. 16). *Manual for SIENA version 4.0*. University of Oxford, Department of Statistics; Nuffield College.

Robins, G. (2015). *Doing social network research: Network-based research design for social scientists*. Sage Publications Ltd.

Robins, G., Snijders, T., Wang, P., Handcock, M., & Pattison, P. (2007). Recent developments in exponential random graph (p*) models for social networks. *Social Networks*, *29*(2), 192.

Rogowski, J. C., & Sutherland, J. L. (2016). How ideology fuels affective polarization. *Political Behavior*, *38*(2), 485.

Rolfe, M. (2014). Social networks and agent-based modelling. In *Analytical sociology* (pp. 233–260). John Wiley & Sons, Ltd.

Rolfe, M., & Chan, S. (2017, Apr. 5). Voting and political participation. In J. N. Victor, A. H. Montgomery, & M. Lubell (Eds.), *Oxford handbook of political networks*. Oxford University Press.

Ryan, T. J. (2017). How do indifferent voters decide? the political importance of implicit attitudes. *American Journal of Political Science*, *61*(4), 892.

Salvy, S.-J., de la Haye, K., Bowker, J. C., & Hermans, R. C. (2012). Influence of peers and friends on children's and adolescents' eating and activity behaviors. *Physiology & behavior*, *106*(3), 369.

Santoro, L. R., & Beck, P. A. (2016, Sept. 1). Social networks and vote choice. In J. N. Victor, A. H. Montgomery, & M. Lubell (Eds.), *Oxford handbook of political networks*. Oxford University Press.

Schaefer, D. R., Kornienko, O., & Fox, A. M. (2011). Misery does not love company: Network selection mechanisms and depression homophily. *American Sociological Review*, *76*(5), 764.

Scheufele, D. A., Nisbet, M. C., Brossard, D., & Nisbet, E. C. (2004). Social structure and citizenship: Examining the impacts of social setting, network heterogeneity, and informational variables on political participation. *Political Communication*, *21*(3), 315.

Schmitt-Beck, R. (2004). Political communication effects; the impact of mass media and personal conversations on voting. In B. Fetsch & F. Esser (Eds.), *Comparing political communication: Theories, cases, and challenges* (pp. 293–324). Cambridge University Press.

Schweighofer, S., Schweitzer, F., & Garcia, D. (2020). A weighted balance model of opinion hyperpolarization. *Journal of Artificial Societies and Social Simulation*, *23*(3), 5.

Sciarini, P., & Tresch, A. (2011). Campaign effects in direct-democratic votes in switzerland. *Journal of Elections, Public Opinion and Parties*, *21*(3), 333.

Shalizi, C. R., & Thomas, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, *40*(2), 211.

Shi, Y. (2016). Cross-cutting messages and voter turnout: Evidence from a same-sex marriage amendment. *Political Communication*, *33*(3), 433.

Silva, B. C. (2018). Populist radical right parties and mass polarization in the netherlands. *European Political Science Review*, *10*(2), 219.

Simas, E. N., Clifford, S., & Kirkland, J. H. (2020). How empathic concern fuels political polarization. *American Political Science Review*, *114*(1), 258.

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology*, *31*(1), 361.

Snijders, T. A. B., & Steglich, C. E. G. (2015). Representing micro–macro linkages by actor-based dynamic network models. *Sociological Methods & Research*, *44*(2), 222.

Snijders, T. A. B., van de Bunt, G. G., & Steglich, C. E. G. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, *32*(1), 44.

Snijders, T. A., Lomi, A., & Torló, V. J. (2013). A model for the multiplex dynamics of two-mode and one-mode networks, with an application to employment preference, friendship, and advice. *Social Networks*, *35*(2), 265.

Sobkowicz, P. (2009). Modelling opinion formation with physics tools: Call for closer link with reality. *Journal of Artificial Societies and Social Simulation*, *12*(1), 14.

Sokhey, A. E., & Djupe, P. A. (2011). Interpersonal networks and democratic politics. *PS: Political Science & Politics*, *44*(1), 55.

Song, H. (2015). Uncovering the structural underpinnings of political discussion networks: Evidence from an exponential random graph model. *Journal of Communication*, *65*(1), 146.

Stadtfeld, C. (2018). The micro–macro link in social networks. In R. A. Scott, M. Buchmann, & S. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences* (pp. 1–15). American Cancer Society.

Stadtfeld, C., & Amati, V. (2021). Network mechanisms and network models. In G. Manzo (Ed.), *Research handbook on analytical sociology* (pp. 432–452).

Stadtfeld, C., & Block, P. (2017). Interactions, actors, and time: Dynamic network actor models for relational events. *Sociological Science*, *4*, 318.

Stadtfeld, C., Snijders, T. A. B., Steglich, C., & van Duijn, M. (2020). Statistical power in longitudinal network studies. *Sociological Methods & Research*, *49*(4), 1103.

Stadtfeld, C., Takács, K., & Vörös, A. (2020). The emergence and stability of groups in social networks. *Social Networks*, *60*, 129.

Steglich, C., & Snijders, T. A. B. (2022). Stochastic network modelling as generative social science. Edward Elgar Publishing.

Steglich, C., Snijders, T. A. B., & Pearson, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociological Methodology*, *40*(1), 329.

Strother, L., Piston, S., Golberstein, E., Gollust, S. E., & Eisenberg, D. (2020). College roommates have a modest but significant influence on each other's political ideology. *Proceedings of the National Academy of Sciences*.

Su, J., Liu, B., Li, Q., & Ma, H. (2014). Coevolution of opinions and directed adaptive networks in a social group. *Journal of Artificial Societies and Social Simulation*, *17*(2), 4.

Sunstein, C. R. (2007). *Republic.com 2.0* (2. print., and 1. paperback print). Princeton Univ. Press.

Swiss Federal Chancellery. (2017a, Aug.). *Referenden*. https://www.bk.admin.ch/bk/de/home/politische-rechte/referenden.html

Swiss Federal Chancellery. (2017b, Aug.). *Volksinitiativen*. https://www.bk.admin.ch/bk/de/home/politische-rechte/volksinitiativen.html

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, *50*(3), 755.

Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology*. Cambridge University Press.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–37). Monterey, CA: Brooks/Cole.

Takács, K., Flache, A., & Mäs, M. (2016). Discrepancy and disliking do not induce negative opinion shifts. *PLOS ONE*, *11*(6), e0157948.

Tedin, K. L. (1980). Assessing peer and parent influence on adolescent political attitudes. *American Journal of Political Science*, *24*(1), 136.

Trechsel, A. H., & Sciarini, P. (1998). Direct democracy in switzerland: Do elites matter? *European Journal of Political Research*, *33*(1), 99.

Turner, J. C. (1991). *Social influence*. Thomson Brooks/Cole Publishing Co.

Vörös, A., Block, P., & Boda, Z. (2019). Limits to inferring status from friendship relations. *Social Networks*, *59*, 77.

Vörös, A., Boda, Z., Elmer, T., Hoffman, M., Mepham, K., Raabe, I. J., & Stadtfeld, C. (2021). The swiss StudentLife study: Investigating the emergence of an undergraduate community through dynamic, multidimensional social network data. *Social Networks*, *65*, 71.

Vörös, A., & Snijders, T. A. B. (2017). Cluster analysis of multiplex networks: Defining composite network measures. *Social Networks*, *49*, 93.

VOTO. (2022). *Publikationen und daten*. Retrieved July 7, 2022, from https://www.voto.swiss/etudes-et-donnees/

Vu, T. M., Brennan, A., Probst, C., Strong, M., Epstein, J. M., & Purshouse, R. C. (2019). Toward inverse generative social science using multi-objective genetic programming. *GECCO 2019 - Proceedings of the 2019 Genetic and Evolutionary Computation Conference*, 1356.

Wagner, M. (2021). Affective polarization in multiparty systems. *Electoral Studies*, *69*, 102199.

Wang, C., Lizardo, O., & Hachen, D. S. (2020). Neither influence nor selection: Examining co-evolution of political orientation and social networks in the NetSense and NetHealth studies. *PLOS ONE*, *15*(5), e0233458.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, UK: Cambridge University Press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, *393*(6684), 440.

Webster, S. W., & Abramowitz, A. I. (2017). The ideological foundations of affective polarization in the u.s. electorate: *American Politics Research*.

Weeks, B. E., & Gil de Zúñiga, H. (2019). What's next? six observations for the future of political misinformation research. *American Behavioral Scientist*, 0002764219878236.

Weeks, B. E., & Gil de Zúñiga, H. (2021). What's next? six observations for the future of political misinformation research. *American Behavioral Scientist*, *65*(2), 277.

Whitt, S., Yanus, A. B., McDonald, B., Graeber, J., Setzler, M., Ballingrud, G., & Kifer, M. (2021). Tribalism in america: Behavioral experiments on affective polarization in the trump era. *Journal of Experimental Political Science*, *8*(3), 247.

Zaller, J. (1992). *The nature and origins of mass opinion*. Cambridge University Press.

Zhang, K. (2019). Encountering dissimilar views in deliberation: Political knowledge, attitude strength, and opinion change. *Political Psychology*, *40*(2), 315.

Zuckerman, A. S., & Kroh, M. (2006). The social logic of bounded partisanship in germany: A comparison of west germans, east germans, and immigrants. *Comparative European Politics*, *4*(1), 65.