Beyond Generative Sufficiency

On Interactions, Heterogeneity & Middle-Range Dynamics

Martin Arvidsson

Linköping Studies in Arts and Sciences No. 836
Institute for Analytical Sociology Dissertation Series No. 4
Faculty of Arts and Sciences
Linköping 2022
At the Faculty of Arts and Sciences at Linköping University, research and doctoral studies are carried out within broad problem areas. Research is organized in interdisciplinary research environments and doctoral studies mainly in graduate schools. Jointly, they publish the series Linköping Studies in Arts and Sciences. This thesis comes from the Institute for Analytical Sociology at the Department of Management and Engineering.

Distributed by:
Department of Management and Engineering
Linköping University
SE-581 83 Linköping, Sweden

Martin Arvidsson
Beyond Generative Sufficiency
On Interactions, Heterogeneity & Middle-Range Dynamics

Edition 1:1
ISBN 978-91-7929-330-7 (print)
ISBN 978-91-7929-331-4 (PDF)
ISSN 0282-9800

© Martin Arvidsson, 2022
Printed by LiU-Tryck, Linköping 2022

This work is licensed under a Creative Commons “Attribution-NonCommercial 4.0 International” license.
ABSTRACT

Explaining how properties at the level of individuals translate into properties at the level of collectives is a core objective of sociology. Because the social world is characterized by complex webs of social interdependencies, establishing how micro and macro are related to one another requires a detailed understanding of how individuals are influenced by their social environments and the consequences that such influences have for the dynamics of the social process. However, until very recently, it has been difficult to conduct detailed empirical investigations of micro-macro linkages due to the lack of large-scale data containing information on how individuals interact with one another. In the absence of such data, substantive research has tended to (a) focus its attention elsewhere: studying how social factors influence individual outcomes, rather than how actors in interaction with one another bring about collective outcomes, or (b) propose models of micro-macro linkages that—for reasons of parsimony and tractability—often assume artificially high levels of homogeneity. Against this background, this thesis sets out to investigate, first, how the data and tools that have emerged from the digital and computational revolution can help sociologists construct empirically well-founded mappings from the micro to the macro level, and second, how the conclusions about the role of social interdependencies and networks change when the analysis is informed by real-world heterogeneities.

In the introductory chapter, a conceptual and analytical framework for studying micro-macro processes is proposed that integrates the theoretical principles of analytical sociology with the data and methods of computational social science. This framework constitutes the foundation of the thesis. It is used in Essays I-III, and it is methodologically built upon in Essay IV.

In Essay I, the role of social networks in labor-market segregation processes is examined. Scholarship on labor-market segregation commonly assume that social networks have a segregating effect because of homophilous selection tendencies in network-based recruitment. Using large-scale register data and focusing attention on individuals’ heterogeneous opportunities to form same-category ties in different workplaces, Essay I finds that opportunity structures often dominate homophilic preferences. In particular, a mechanism is identified which shows—in contradiction with the main tenet of previous research—that networks often reduce rather than increase segregation by triggering mobility events that counteract the impact of segregating mobility events.

Essay II examines the conditions under which social influence can decouple adoption behaviour from individual preferences and thereby bring about unexpected collective outcomes. Prior research has shown that such decoupling can occur, but conflicting evidence and implicit assumptions of strong homogeneity mean that we still know little about the conditions under which this is likely to occur in the real world. Addressing these limitations, this study uses fine-grained, real-world behavioural data from Spotify to estimate heterogeneous social influence effects conditional on properties of individuals’ social environments, and then examine their macro-implications in empirically calibrated simulations. It is found that partial overlap in preferences and strong social ties between the senders and receivers of social influence is needed for social influence to produce decoupling.
Essay III centers on the phenomenon of urban scaling and examines the relationship between within-city and between-city inequality. Previous urban scaling research has documented how cities’ total outputs increase more than proportionally with city size and has proposed theoretical models which demonstrate impressive predictive accuracy at aggregate levels. However, this research has overlooked the stark inequalities that exist within cities. Using microdata from multiple countries, it is found that between 36–80% of the previously reported scaling effects can be explained by differences in the distributional tails of cities. Providing explanatory depth to these findings, a cumulative advantage mechanism is identified which elucidates one important channel through which differences in the size of cities’ tails emerge.

In Essay IV, a method is proposed for inferring theoretically meaningful dimensions from complex high-dimensional data such as text. The results show that the method captures latent semantic concepts better than or on-par with the current state of the art. For the study of social interactions, the method constitutes a new and potentially important tool for inferring theoretically meaningful dimensions about individuals and their social environments, and in so doing, improves our ability to adjust for specific types of homophily and enables richer and more precise measures of heterogeneity in social interaction processes.
Acknowledgments

A central theme of this thesis is the often-surprising effects of social interaction. As it happens, the thesis itself is an unexpected result of long line of such interactions. My journey from statistics to sociology and the result, this dissertation, would not have been possible without the input and support from the many people I have had the privilege to interact with along the way.

First, I would like to thank my terrific supervisors, Peter Hedström and Marc Keuschnigg. Thank you for giving me the confidence, support, and inspiration to develop beyond my methodological interests. Thank you for all the insightful discussions. Thank you for the many rounds of excellent feedback that I have received over the years and thank you for our collaborations. Working and co-authoring with both of you has been an absolute pleasure, and I have learnt a tremendous amount. Your influence has certainly been of the wide kind!

I feel very lucky to have done my PhD at such a great workplace as the Institute for Analytical Sociology (IAS), surrounded by amazing colleagues and researchers. During these years, there has never been a shortage of interest in discussing ideas and providing feedback. Countless of such discussions, formal and informal, small and large, at seminars, in the lunchroom, in the hallways, at the pub, have undoubtedly shaped my thinking and made this thesis better than it otherwise would have been. A special thanks to Eduardo Tapia, Anders Hed, Benjamin Jarvis, Chanchal Balachandran, Sarah Valdez, Christian Steglich, Jacob Habinek, Karl Wennberg, Maria Brandén, Carl Nordlund, Flóra Samu, Károly Tákacs, Åsa Arnoldsson, and Madelene Töpfer, for your support over these years.

I am very fortunate to have been a part of a fantastic group of doctoral students. Àlex Giménez de la Prada, Niclas Lovsjö, and Selcan Mutgan, Miriam Hurtado Bodell and Emanuel Wittberg, Laura Fürsich and Rodrigo Martinez Peña, I treasure all the good times, all the small (and sometimes very long) coffee breaks, all the discussions, and all the support you have given me throughout. You have indeed made this journey a much smoother and more enjoyable one than it otherwise would have been!
Two essays of the thesis were co-authored with researchers other than my supervisors, and I would like to sincerely thank them for great collaborations. Miriam Hurtado Bodell and Måns Magnusson, I really enjoyed those intensive spring months of 2019 when we worked on Essay IV, and I learned a lot from both of you! I also fondly remember the several meetings we had in Stockholm. François Collet, our journey with Essay I was long, but it was a great pleasure to collaborate throughout! Thank you also for hosting my research visit at ESADE, Barcelona in the beginning of the project.

Throughout these years, I have been fortunate to have several excellent researchers reading and providing careful input on my work. I want to thank Christofer Edling, Eduardo Tapia, and Jacob Habiniec (half-time seminar), and Michael Mäs and Adel Daoud (pre-defense seminar), for their highly useful feedback. Your comments without a doubt helped me make the thesis better! Outside of formal seminars, I would like to also extend my thanks a number of fantastic scholars who have been so generous to read and comment on different parts of the thesis. Thank you, Benjamin Jarvis, Eduardo Tapia, Adel Daoud, and Christian Steglich (Essay I), Arnout van de Rijt and James Evans (Essay II), Scott Page and Scott Ortman (Essay III), Jacob Habiniec (Essay IV), and Sarah Valdez (Kappa).

Prior to arriving at IAS, several people at LiU’s Statistics department showed great support and encouragement. In this regard, I would like to especially thank Mattias Villani and Anders Nordgaard. Your support inspired me to push myself further and become a better student and researcher as a result. Mattias, I am also very grateful for linking me up with IAS.

To my parents Lars and Ann-Britt, and my sister Ida, I am forever grateful for your unconditional support. You have always believed in me and have always been there for me. Thank you for everything! I also want to thank my extended family who also have been so supportive, especially Ing-Britt and Elisabeth. Finally, I would like to thank my wonderful sambo, Lisa, for her endless love and support. You have been there through thick and thin, and I am so fortunate to have you in my life.

Martin Arvidsson
Linköping, Sweden
April, 2022
List of Essays

This thesis is based on the following essays, which are referred to in the text by their Roman numerals.

https://www.science.org/doi/full/10.1126/sciadv.abf6730 
Published under CC-BY-NC license


https://aclanthology.org/D19-1661/ 
Published under CC-BY license
# Contents

Abstract iii
Acknowledgments vi
List of Essays vii
Contents ix
Introduction 1
Appended Essays 39
Discussion and conclusions 45
Bibliography 53
Essay I 73
Essay II 97
Essay III 159
Essay IV 199
Introduction

Explaining how properties at the level of individuals (e.g., preferences, motivations, resources) translate into properties at the level of collectives (e.g., segregation, inequality, polarization) is a core objective of sociology, and has interested scholars ever since the dawn of the discipline (see e.g., Durkheim, 1895; Tarde, 1903). Because the social world is characterized by complex webs of social interdependencies (Simmel, 1955), establishing how the micro and macro levels are related to one another requires a detailed understanding of how individuals are influenced by their social environments and the consequences that such influences have for the dynamics of social processes (Hedström, 2005; Schelling, 1978). However, until very recently it has been difficult to conduct detailed empirical investigations of micro-macro linkages due to the lack of large-scale data containing information on how individuals interact with one another. In the absence of such data, substantive research has tended to either (a) focus its attention elsewhere: studying how social factors influence individual outcomes, rather than how actors in interaction with one another bring about collective outcomes (Macy & Willer, 2002; Sorensen, 1998), or (b) propose models of micro-macro linkages that—for reasons of parsimony and tractability—often assume artificially high levels of homogeneity. Against this background, this thesis sets out to investigate, first, how the data and tools that have emerged from the digital and computational revolution can help sociologists construct empirically well-founded micro-macro mappings,
and second, how conclusions about the role of social interdependencies and networks change when the analysis is informed by real-world heterogeneities.

This introductory chapter is organized as follows. First, I discuss the nature of collective properties and the complexity involved in going from one scale of the social to another. Second, I discuss the key source of this complexity, interdependence, and the challenges this presents for empirical work. Third, I consider the relationship between empirical realism and bottom-up explanations in analytical sociology, and fourth, I reflect on the promise of the digital and computational revolution for sociological theorizing and for the identification of micro-macro dynamics using observational data. Finally, I propose a conceptual and analytical framework that fuses the data and tools of the digital and computational revolution with the explanatory principles of analytical sociology.

**The nature of collective properties.** A ubiquitous observation in social life is that many properties of collective entities are characterized by high concentration, clustering, and interrelatedness—often resembling the type of patterns displayed in the right-hand (B) rather than the left-hand panels (A) of Figures 1–3. That is, rather than following a normal distribution (Figure 1A), the choices that individuals make, the activities they engage in, and the rewards that they reap, tend to be distributed in a highly unequal fashion, being concentrated to small numbers of alternatives/individuals, which results in so-called heavy-tailed distributions (Figure 1B). In cultural markets, for example, although there are millions of artists, actors, and authors at whom individuals could in principle direct their attention, when asked what music they listen to, what movies they watch, and what books they read, individuals’ replies are highly concentrated: the lion’s share of all listening, watching, and reading is concentrated to a very small fraction of artists, actors, and authors (Greco, 2013; Keuschnigg, 2015; Salganik et al., 2006). Similarly, where individuals live (Auerbach, 1913), which technologies they use (Brynjolfsson et al., 2011), and what websites they visit (Albert et al., 1999), are all heavily concentrated relative to the number of options available. Further examples of this phenomenon are found in the financial success of individuals and organizations (Alvaredo et al., 2017; Atkinson et al., 2011; Pareto, 1896; Tomaskovic-Devey et al., 2020), fame (van de Rijt et al., 2013), prestige (Cook, 1995; Erickson & Nosanchuk, 1984), public attention (Hilgartner & Bosk, 1988), academic citations (Eom & Fortunato, 2011; Fortunato et al., 2018; Price, 1976), and social ties (Barabási, 2009; Dunbar, 2016).

The choices that people make, the activities they engage in, and the rewards that they reap, are not only highly concentrated. In addition, rather than being
Properties of collectives often exhibit clustering. Uniformly concentrated across a population (Figure 2A), they tend to exhibit high levels of clustering (Figure 2B). For example, a person’s favorite artist, actor, or author may vary substantially depending on which group of people you ask. Some are into country music, others prefer jazz, while yet others like hip hop, and such preferences are not randomly distributed but highly clustered in different segments of the population—along dimensions such as age, gender, ethnicity, education, income, and geography (Bourdieu, 1987; Grinblatt et al., 2008; Mark, 1998; McPherson et al., 2001). Segregation provides another example of clustering. People live, work, and go to school with others who tend to be similar to themselves along several dimensions (Massey & Denton, 1988; Moody, 2001; Reskin, B., 1993).

Not only are likes and dislikes—and behaviors more generally—clustered in the population, they also tend to exhibit strong correlational patterns: if you like X this probably means that you also like Y, or if you hold opinion Z you are also likely to approve of W (Figure 3B). Enjoying jazz, for example, is positively correlated with also liking blues and classical music, while being
Introduction

into rap is positively correlated with liking pop and R&B (Ollivier et al., 2019; Rentfrow et al., 2011). Correlational patterns such as these are found not only within a given category (e.g., musical taste or car preferences), but also across categories. Attitudes towards homosexuality, for example, have been found to be substantially correlated with individuals’ musical tastes and reading habits. Based on surveys conducted in the US (Baldassarri & Gelman, 2008), DellaPosta et al. (2015) found that individuals who liked reggae and read fiction were far less likely to disapprove of homosexual intimacy. A more general result from the same study, which mirrors the findings from a long line of previous research (Chaney, 2002; Giddens, 1991; Miller, 1985), was that individuals’ lifestyles were substantially correlated with their political affiliations.

Figure 3: Properties of collectives often exhibit multiplex clustering.

The linear aggregation thesis. How can we explain macro-regularities such as those described in Figures 1–3? One possible hypothesis is that they simply reflect properties inherent to the individuals and items in question: that the quality of some artists, actors, and authors just happens to be so much better than that of the rest (explaining the existence of heavy tails), that people who belong to different sociodemographic categories or live in different areas just happen to have very different innate preferences for things like music, sports, and politics (explaining local clustering), and that innate personality traits cause individuals to have both certain political beliefs and particular musical tastes (explaining the cross-category correlational patterns). That is, in short, knowing the likes and dislikes of the individuals who are part of the collective, we can simply aggregate their preferences to arrive at the collective properties of the system.

While its simplicity and parsimony is appealing, this ‘linear aggregation thesis’ falls short in a number of ways. First, empirical research has found that
only a small fraction of the variability in behaviors and life outcomes can be explained by reference to individual properties alone (Garip, 2020; Salganik et al., 2020). Similarly, research also shows that predicting the success of specific cultural items (e.g., artists) or technologies based solely on their properties is very error-prone (Arthur et al., 1994; De Vany, 2003; Gardner, 2010; Janosov et al., 2020; van de Rijt et al., 2014). For example, rather than jumping at the opportunity to sign a deal with George Lucas (Star Wars) or J.K. Rowling (Harry Potter), the initial reaction of industry executives to these works was reportedly lackluster; one after the other rejected these (now) world-wide best sellers (Watts, 2007).

But perhaps these observations are simply the result of data limitations: if only we could measure more things about individuals, cultural items, and technologies, we would perhaps be able to predict these kinds of outcomes? This leads us to the second problem with the linear aggregation thesis: theoretical models have demonstrated that even perfect knowledge of the properties of micro-level entities may tell us little about the macro-level properties that will emerge, and vice versa (Anderson, 1972), even if the rules guiding the behavior at the lower scale are simple (Wolfram, 2002). In fact, the collective patterns produced by such models often seem to contradict the behavioral rules implemented at the micro-level. In his now classical model of segregation, for example, Schelling (1971, 1978) showed how individually tolerant agents can collectively produce extreme levels of segregation (see also Sakoda, 1971). Similarly, theoretical models of status hierarchies (Correll et al., 2017; Gould, 2002; Lynn et al., 2009; Manzo & Baldassarri, 2015) have demonstrated that stark inequality in status can emerge in groups of individuals with similar individual abilities, and that the connection between individuals’ ability and status can become substantially blurred. Another illuminating example of the seeming disconnect between micro and macro is provided by DellaPosta et al. (2015), who showed how arbitrary correlations between likes and dislikes can emerge within groups in such a way that lines of polarization between groups are also drawn arbitrarily.

A third category of evidence against the linear aggregation thesis comes from macro-sociological experiments (Hedström, 2006). These experiments—in which multiple experimental “worlds” are constructed and allowed to evolve separately from each other—have demonstrated that very different collective outcomes can emerge in different worlds, even though conditions are identical and the subjects are sampled from the same population. In their seminal MusicLab experiment, Salganik et al. (2006) used this kind of experimental setup and showed that different songs emerged as chart toppers in different worlds: the popularity
of a song in one world thus said rather little about its popularity in another. Using a similar design, Macy et al. (2019) corroborated the findings of DellaPosta et al. (2015), and showed that different cleavages between political affiliations emerged in different worlds.

These examples illustrate a seeming disconnect between different levels of the social: the outcomes that we observe at the macro-level are not uncommonly counterintuitive and surprising considering the properties of the individuals who make up the collective. Where does this leave us? If the linear aggregation thesis fails, what else might explain the type of macro-regularities displayed in Figures 1–3?

Interdependence and the micro-macro problem. The difficulty in mapping micro-level properties to macro-level outcomes is known as the micro-macro problem in sociology. The failure of the linear aggregation thesis demonstrates a central feature of this problem—that the whole is often different from the sum its parts (Anderson 1972), and that the whole cannot be understood by studying its parts in isolation. The micro-macro problem has been a core interest of sociologists throughout the history of the discipline. Emile Durkheim, for example, famously observed that:

“...society is not the mere sum of individuals, but the system formed by their association represents a specific reality which has its own characteristics. Undoubtedly no collective entity can be produced if there are no individual consciousnesses: this is a necessary but not a sufficient condition. In addition, these consciousnesses must be associated and combined, but combined in a certain way. It is from this combination that social life arises and consequently it is this combination which explains it.” (Durkheim, 1982 [1895], p. 129).

In other words, what is missing from the linear aggregation thesis, and what makes the whole different from the sum of its parts, is interdependence. The individuals that comprise a social system do not exist in separate vacuums. Instead, their behaviors depend not only on their own properties (e.g., preferences, resources, etc.) but are also a function of the past and present behaviors of other individuals. Such dependence may result from a direct type of interaction between individuals, as is the case when individuals learn about job opportunities through friends and acquaintances (Granovetter, 1974), but individuals may also be interdependent without any direct interaction, as when an individual leaves a particular job or housing unit and in doing so creates a vacancy (opportunity) for others to fill (White, 1970).
One should distinguish between two interrelated aspects of interdependence. The first concerns how, at the local level, any two individuals may be interdependent and influence each other, while the other pertains to how social relations are structured. With regard to the former, a large body of research in sociology and social psychology has investigated the sources of social influence (Cialdini & Goldstein, 2004), identifying a number of important ways in which individuals come to influence each other, such as social pressure (Asch, 1955; Homans et al., 1950; Sherif, 1935), persuasion (Myers, 1982), cognitive dissonance (Festinger, 1957), externalities (Coleman, 1990), and social learning (Akers et al., 1979; Bikhchandani et al., 1992; Hedström, 1998). Relatedly, this line of research has also investigated how different properties moderate such influence effects, identifying contextual factors such as uncertainty (Coleman et al., 1957; Deutsch & Gerard, 1955; Hedström, 1998), and relational properties such as similarity, status, trust, and credibility to be important determinants (Aral & Walker, 2012; Centola, 2018; Rogers, 2010).

Concerning structure, network research has documented a wide range of macroscopic properties about the structure of relations, including the average distance between individuals, showing that we live in a "small world" and are able to reach most other individuals in a population with just a few steps (Milgram, 1967; Travers & Milgram, 1969; Watts & Strogatz, 1998); clustering, showing that we nevertheless live in communities with many shared ties with those closest to us (Holland & Leinhardt, 1971; Luce & Perry, 1949); density and degree, showing that individuals have few ties relative to theoretical maximums, and that some have many more ties than others (Barabási & Albert, 1999; Friedkin, 1981; Newman et al., 2001); and assortative mixing, showing that individuals who are close to each other in the network also tend to be close to each other in terms of individual attributes (Newman, 2002). Complementing these macroscopic observations, research has also identified properties of individuals and local network structures that are associated with these macro patterns, including individuals’ tendencies to reciprocate ties (Gouldner, 1960; Newcomb, 1956; Schaefer et al., 2010); triadic closure, the tendency of a friend of a friend to also become a friend (Goodreau et al., 2009; Granovetter, 1973, 1985; Newcomb, 1961); homophily, the tendency to befriend others who are similar to oneself (Kossinets & Watts, 2009; Lazarsfeld & Merton, 1954; McPherson et al., 2001); and that cognitive limitations constrain the number of social relations any one individual can maintain (Dunbar, 1992, 2016).

1It is worth clarifying the relationship between the terms “interaction” and “influence”. Throughout this thesis, I use social influence to describe how individuals in interaction affect each others’ behaviors.
Considered jointly, dyadic influence and overlapping webs of social ties imply that individuals’ behaviors are influenced not only by immediate alters but also, indirectly, by the behaviors of even socially distant alters. For the relationship between micro and macro, these “chains” of interdependencies enable behaviors to cascade throughout the system and drive—often unexpected—collective outcomes. The existence of chains of interdependencies often imply path dependency. Small, possibly random differences in initial conditions, when accumulated over time, can matter greatly for the collective outcome that is eventually brought about.\(^2\) The MusicLab experiment provides a vivid demonstration of how chains of interdependencies can produce path dependencies and behavioral cascades that lead to unexpected collective outcomes (Salganik et al., 2006). MusicLab involved the creation of separate, digital music distribution platforms (“worlds”) that developed independently of one another. The songs available were the same in all worlds, but in some worlds, users could see how many others had downloaded each song. This form of interdependence had the consequence that songs which received more downloads early on—because of some combination of chance and the specific taste of the users who happened to be among the first to participate in the experiment—were then also more likely to be downloaded by subsequent users who relied on the “most downloaded list” to decide which music to listen to. This process, of (i) chance partly determining the first users’ downloads, and (ii) initial popularity being reinforced by subsequent users, resulted in different songs emerging as the most popular ones in different worlds. In yet other worlds, individuals were not able to observe each other’s music download behaviors—i.e., in these worlds there were no social interdependencies. As a consequence, there was no path dependency, and the collective outcome that emerged mirrored the aggregate preferences of the individual participants. MusicLab thus illustrates the complexities created by interdependency, shows why linear-aggregation type explanations do not work, and underscores the crucial importance of understanding the details of how individuals influence one another. Only when these details are known can we properly map how individual and collective properties are linked.\(^3\)

\(^2\)That path dependency often follows from social interactions is a recurring idea that is found in many central sociological theories, such as Berger and Luckmann’s (1966) explanation of the emergence of institutions and Bourdieu’s (1987) notion of habitus.

\(^3\)The fact that this mapping is possible exemplifies that properties of collectives are ultimately reducible to individuals and their interactions, and that there is no “mystery gap” between the micro and the macro level once the complexities of the social process are well-understood (Di Iorio & León-Medina, 2021). What makes this mapping difficult, and makes collective outcomes often seem counter-intuitive, are the interdependencies that exist. In this sense, the type of “emergence”
Figure 4 presents a stylized example of how “worlds” may develop differently depending on initial conditions when individuals are interdependent. In these worlds, we have two types of agents: blue and red. Without exposure, blue agents prefer X and red agents prefer Y. Under exposure, agents adopt the item which the majority of previous agents have adopted. Figure 4 shows how, depending on the color of the first two agents (world 1: blue, world 2: red), the world becomes completely dominated by either X or Y adoptions. This example constitutes a case of what the literature often refers to as an informational cascade (Bikhchandani et al., 1992). In contrast, no cascading is observed in the independent worlds: rearranging the order of the agents does not affect adoption rates at all.

Figure 4: Toy example contrasting the adoption patterns that emerge in two interdependent and two independent—but otherwise identical—fictive “worlds”. Individually, blue agents prefer X and red agents prefer Y. When interacting, agents adopt the item which the majority of previous agents have adopted.

In addition to emphasizing the central role of interdependence, these two examples (MusicLab and the toy example in Figure 4) also reveal that the types of data that are typically collected and analyzed by sociologists—i.e., random samples from populations of interest, ethnographic case studies, or aggregate statistics—would not easily allow us to arrive at a correct understanding of why e.g., certain songs happen to be more popular than others in certain groups, or why X came to dominate the market instead of Y. To capture the interdependency of individuals, we need to observe individuals’ behavior in the actual social environments in which they operate, at a scale, breadth, and temporal granularity that allows us to identify the causal effect of social influence and to account for relevant heterogeneities (Aral et al., 2009; Shalizi & Thomas, that occurs in the social sciences—between individuals and collectives—differs from that found in the natural sciences (Ramström, 2018).
Then, to capture the dynamics implied by interdependence and the cumulative effects of these dynamics, we need to observe the actual unfolding of the process—at the appropriate scale, breadth, and temporal granularity—over a sufficiently long period of time.

Hence, as noted by a number of scholars, although the micro-macro problem has been a core theoretical concern of sociologists since the inception of the discipline, the limitations imposed by the types of data and the types of statistical tools available meant that over the course of the 20th century, quantitative sociology came to increasingly abandon the quest to understand how the interdependent behaviors of actors produce different collective outcomes, and to instead focus on explaining individual outcomes in terms of different social factors (Abbott, 1988; Coleman, 1986; Hedström, 2005; Hedström & Swedberg, 1996; Macy & Willer, 2002; Sørensen, 1976, 1998).

Analytical sociology, mechanism-based explanations, and generative models. Analytical sociology (AS) emerged in part as a response to these developments, calling for a return to questions about interdependence and micro-macro relations, a tighter link between theory and empirical research, and an emphasis on certain explanatory principles (Hedström, 2005). According to AS, establishing the existence of macro-macro or micro-micro relationships between different variables—whether predictive or causal—is not sufficient to explain collective outcomes. Instead, AS postulates that explanations of collective outcomes must refer to the social mechanisms that brought them about (Hedström, 2005; Hedström & Bearman, 2009; Hedström & Ylikoski, 2010; León-Medina, 2017a; Manzo, 2010, 2014), where “[a] mechanism for a phenomenon consists of entities (or parts) whose activities and interactions are organized as to be responsible for the phenomena” (Glennan & Illari, 2017). Hence, the claim of AS is that explanations of collective outcomes should detail how the interdependent activities of individuals jointly produce the collective outcome in question.

Methodologically, analytical sociology’s quest for mechanism-based explanations of macro-outcomes has implied a reduced role for traditional, independence-assuming research designs and statistical analysis. The primary role of such analyses in the field of AS instead lies in establishing empirical patterns, which formal models then either incorporate as assumptions or seek to explain. Shifting the focus away from how social factors affect individual outcomes towards how interdependent actors bring about collective outcomes, tools for analyzing interactions has taken center stage in analytical sociology. Prominent examples of these tools include game theory (Coleman, 1990; Przepiorka, 2021; Swedberg, 2001), and agent-based simulation models, ABMs.
ABMs in particular—which offer substantial modeling flexibility, are able to deal with arbitrary levels of complexity, and which remain neutral to any specific epistemic assumptions about agents—have gained great traction among analytical sociologists. In brief, ABMs enable researchers to construct “digital worlds” in which the properties of agents, their behavioral rules, and the environments in which they are embedded, are designed by the researcher. Based on these digitally encoded assumptions, the social process of the digital world is then allowed to unfold through simulation. ABMs thus enable the systematic investigation of how micro-level behaviors, embedded in social structures, produce macro-level outcomes. Using counterfactual experiments, it becomes possible to study the implications of postulated mechanisms. Hence, there are deep connections between thinking in terms of mechanisms and building agent-based models. As Manzo (2010) put it:

“...when one is designing an agent-based model, one is building an artificial mechanism [and] when one simulates an agent-based model, one is activating in silico the process that the mechanism potentially contains.” (Manzo, 2010, p. 147).

Researchers typically specify the behavioral rules of ABMs by either (1) making more or less psychologically plausible assumptions, or (2) basing the assumptions on empirical research findings. The former approach is by far the most common, leading to tractable “toy models”. Research in this tradition has produced important insights about micro-macro linkages, showing how minimalistic assumptions about agents’ behavior and interdependence can produce complex macroscopic patterns observed in the real world (Bearman...
et al., 2004; Centola et al., 2005; DellaPosta et al., 2015; DiMaggio & Garip, 2011; Goldberg & Stein, 2018; Manzo & Baldassarri, 2015; Mark, 2003; Squazzoni & Gandelli, 2012). A prime example of this is found in the work of DellaPosta et al. (2015), who demonstrated that polarization in lifestyles between different groups can emerge arbitrarily as a result of a simple dynamic process of selection and influence, thus proposing a highly parsimonious and intuitive mechanism that could possibly lie behind the puzzling polarization in lifestyles that has been observed between supporters of the two political parties in the US.

While ABMs of this kind are very useful for exploring social theories, proposing novel mechanisms, and developing hypotheses, pitfalls arise when they are employed to explain concrete real-world social processes and collective outcomes. While a simple toy model can demonstrate that a set of assumptions \{A,B\} is able to reproduce the macro-outcome of interest, i.e., demonstrating generative sufficiency, there are potentially many other sets of assumptions \{A,C\}, \{B,D\}, etc., that might do the same. By using highly abstract, conventional ABMs, we may thus easily fall into the trap of telling as-if stories, mistaking how a collective phenomenon might have been produced for how it actually was produced. This risk is further amplified by the common reliance on common sense to determine the validity of micro-level assumptions (Watts, 2014). Demonstrating that the assumptions \{A,B\} in fact contributed to the macro-outcome of interest requires an empirical investigation of how the actual process unfolded.

To the extent that the second type of ABMs, i.e., those based on empirical calibrations, has been used, the data used for calibration have usually been survey data or aggregate statistics. This is potentially problematic, however. First, as emphasized earlier, to properly identify how individuals are influenced by their social environments, and to separate influence from homophily, rather special data properties are required, which are not typically found in traditional sociological research designs—because such designs often either ignore interdependence by design, or do not measure it at a sufficient level of detail and/or scale. Second, random samples of individuals or aggregate statistics do not allow for an investigation of the dynamics that characterizes the empirical process. As a result, the dynamics generated by a simulation model cannot be compared to the actual dynamics present in the real social process. This in turn means that, as in the case of the conventional non-empirically calibrated ABMs, these types of conventional sociological data do not allow us
to properly assess how well the simulated dynamics compare to the actual, real-world dynamics.  

From the point of view of the explanatory principles of AS, these approaches would seem not to be entirely satisfactory given that one of the central pillars of AS is *empirical realism*: as-if stories are not sufficient, only dynamics that actually operate are deemed to be acceptable constituents of explanations (Hedström, 2005; Jarvis et al., 2021). A reliance on plausible psychological assumptions (see e.g., Hedström and Ylikoski, 2014) and generative sufficiency (Epstein, 1999) would therefore not appear to be sufficient for explanations of concrete social phenomena: models based on these may provide compelling stories of how the collective phenomena in question *could plausibly have* been produced, without these stories having anything to do with how these phenomena actually were produced (Macy and Flache, 2009; León-Medina, 2017a; Hedström, 2021). To explain concrete social phenomena—e.g., the segregation of a particular city, the inequality between urban and rural areas of a particular country, the increased support for populist and far-right parties in a particular region, the polarization along some axis X on platform Y, etc.—considerable care and caution is required to ensure that the postulated dynamics do indeed apply to the case at hand (see Discussion and Conclusions for further a discussion of this point). As has been argued above, this typically necessitates data and research designs that are different from those included in the traditional sociological toolkit.

**The promise of the computational and digital revolution.** One important change that has occurred during the past couple of decades is the computational and digital revolution (Lazer et al., 2009). On the one hand, massive digitization efforts on the part of states and corporate bureaucracies are turning administrative records—such as employment, scholastic, residential, health, and criminal records—into digital data that are accessible to researchers. On the other, the increase in time spent online using various digital services, such as e-mail, web search, entertainment platforms, and social media also means that more of our social lives are becoming digitized. For analytical sociologists, and social scientists more generally, the digital revolution offers population-scale datasets with rich, high-dimensional information about individuals, their social environments, and their activities in those social environments. Substantial advances in methods for manipulating and analyzing these “big data” have complemented the digital revolution. Not surprisingly, these developments

---

6I use the term “dynamics” to refer to *types of (inter)action sequences that are defined in terms of higher-level properties, abstracting away details about specific actors, specific locations, and specific time-points. The cumulative advantage processes exhibited in the MusicLab experiment would be one example of such a dynamic.*
Introduction

have created considerable excitement across the social sciences, and have led to the formation of the interdisciplinary community of *computational social science*, CSS, bringing together scholars interested in making use of these developments (Lazer et al., 2009).

In light of the limitations discussed above, the computational and digital revolution would seem to have important ramifications for the practices of analytical sociology. Indeed, central figures within AS have suggested that these developments are just what is needed to close the gap between the field’s theoretical ambitions and its empirical research (Jarvis et al., 2021; Keuschnigg et al., 2018). Surprisingly, however, with the notable exception of large-scale online experiments (Centola, 2010; Guilleaulet et al., 2021; Macy et al., 2019; Salganik et al., 2006; van de Rijt, 2019), the CSS revolution has thus far had relatively little impact on the practices of AS. This is illustrated by the fact that no Merton award (the annual best paper award in analytical sociology) has yet been given for a paper using large-scale relational data from the digital revolution (Manzo, 2021). This may in part be due to its relative recency—although it now has been 13 years since Lazer et al.’s (2009) foundational *Science* article. Another possibility, not orthogonal to recency, is that there still remain important unanswered questions about how analytical sociologists might fruitfully make use of these new tools and data for the investigation of micro-macro processes.

While applications remain scarce among analytical sociologists, many can be found in the wider CSS community. This literature has provided important descriptive insights (e.g., Bettencourt et al., 2007; Flaxman et al., 2016; Sekara et al., 2016; Garg et al., 2018, Garcia et al., 2018, Fatehkia et al., 2018), and has demonstrated an impressive predictive ability in relation to a wide range of social phenomena (e.g., Madan et al., 2010; Bollen et al., 2011; Berk, 2012; Mestyan et al., 2013; Blumenstock et al., 2015; Jean et al., 2016). However, regarding the question of how to make sense of collective outcomes, the CSS literature tends to either (i) spend little attention to the details of local interaction dynamics—e.g., to separate influence from homophily and to specify the mechanisms involved—and instead use different micro or macro properties to predict the macro-pattern in question (e.g., Bakshy et al., 2011; Jenders et al., 2013; Cheng et al., 2014; Brady et al., 2017; Vosoughi et al., 2018), (ii) focus on the details of local interactions but ignore their macro implications (e.g., Aral and Nicolaides, 2017; Easley et al., 2020; Ternovski and Yasseri, 2020, Eckles and Bakshy, 2020), or (iii) rely on highly stylized mathematical or simulation models that are uninformed by detailed microdata (e.g., Barabási and Albert, 1999; Bettencourt, 2013; West, 2017). It would thus appear that, as in the AS community, there still exists a number of important questions in the CSS...
community regarding how these new tools and data might be used to establish mechanistic and empirically anchored micro-macro mappings.

Hence, the first objective of this thesis is to investigate how the methods and data emanating from the computational and digital revolution can be used to construct mechanism-based explanations of collective outcomes. This objective will be addressed first in the section “An analytical framework that bridges AS and CSS”, which introduces an analytical framework for dissecting micro-macro processes, and second in each of the four essays presented in the thesis, in which this framework is either applied or methodologically built upon.

**Conceptualizing micro-macro processes.** To devise empirical analyses that informs our understanding of micro-macro processes, we must first specify the elements that define such processes. That is, we must provide an abstract conceptualization of the process. Among analytical sociologists, the so-called “Coleman diagram” (Coleman, 1986) has emerged as a central tool for thinking about, conceptualizing, communicating, and criticizing proposed micro-macro processes (Ylikoski, 2016). It can be found in virtually all central expositions about AS, where it is used to demonstrate the general approach of AS, what social mechanisms are, how to conduct substantive research, and how to construct agent-based models (Hedström, 2005; Hedström & Bearman, 2009; Hedström & Swedberg, 1998; Hedström & Ylikoski, 2010; León-Medina, 2017a; Manzo, 2010, 2014). The Coleman diagram can be summarized as follows—adopting Ylikoski’s (2016) labeling-scheme of arrows and nodes (see Figure 5): Node A and D represent different macro-level facts/variables, with A representing the macro-level cause and B the ultimate macro-level outcome. The intermediate nodes B and C, on the other hand, are defined at the micro-level, with node B referring to the properties of agents and their situations, and node C to the behavior of agents. The arrows describe the causal relationships between these nodes. Arrow 1: how macro facts (A) influence individuals’ preferences and opportunities (B). Arrow 2: how such preferences and opportunities (B) affect the behaviors of individuals (C). Arrow 3: how such behaviors (C) combine to produce the macro-level outcome (D).

Coleman’s discussion of Max Weber’s analysis of Protestantism and capitalism provides a powerful example of how this conceptual framework can be used for both summarizing and identifying gaps in theories operating at multiple scales:

“For example, in Max Weber’s analysis of Protestantism and capitalism, he shows through illustration the effect of Protestant doctrine on individual values [arrow 1] and, again through illustration, the effect of
Introduction

*these values on individual orientations to economic behavior [arrow 2]. What he fails to show is how these individual orientations combined to produce the structure of economic organization that we call capitalism [arrow 3]." (Coleman, 1986, pp. 1322–1323).

![Figure 5: The Coleman diagram (Coleman, 1986).](image)

While the Coleman diagram captures the larger steps involved in mapping micro to macro and vice versa, it notably does not explicate one of the most central properties of social mechanisms, namely *the interdependencies that exist between individuals and the implied dynamics of such interdependencies* (León-Medina, 2017b; Más, 2018; Ylikoski, 2016). As a result, it cannot easily be used to explicate the core matter of some of the most celebrated mechanisms and models in the field of AS, e.g., Merton’s notion of Matthew effects (Merton, 1968) and self-fulfilling prophecies (Merton, 1948), Schelling’s segregation model (Schelling, 1971), and Granovetter’s threshold model (Granovetter, 1978). The fact that the dynamics implied by interdependence are not detailed also means that the Coleman diagram cannot easily communicate assumptions or facilitate empirical investigation about such dynamics. Thus, one aspect of addressing the first objective of this thesis—which concerns how the methods and data emanating from the digital and computational revolution can be used to develop realistic mechanism-based explanations—will entail providing a conceptualization of micro-macro processes that details the interaction.

---

7See Ylikoski (2016) for a recent extension of the Coleman diagram that do account for interdependencies in the form of a feedback loop from node C to B. As will become clear in the next section, however, even with this added arrow, the type of dynamics that can be described remains restricted.
Heterogeneity and its implications for micro-macro processes. Modern societies are characterized by considerable heterogeneity in terms of both the type of social environments in which individuals are embedded, and how individuals respond to various events that occur in these environments (Blau & Schwartz, 2018; Bourdieu, 1987; Simmel, 1950). Does this type of heterogeneity matter for how micro-properties map to macro-outcomes? In his classic threshold model, Granovetter (1978) provided a clear answer to this question: yes, it does. Granovetter showed that even small differences in the distribution of susceptibility to social influence across a population can have dramatic consequences for the extent to which a behavior is adopted in the population in question: whether a behavior is widely adopted or not may crucially depend on just a few individuals who are more willing than others to adopt the behavior at an early stage, and thus set in motion a process of contagion.

Although highly celebrated, when it comes to actual research and explanations of concrete macro-outcomes, this insight by Granovetter has often been overlooked for various reasons. On the one hand, empirical analysis of heterogeneity and its effects on the dynamics of social processes requires a very particular type of data (large-scale, granular, networked, etc.) which has typically not been possible to obtain using traditional sociological research designs. On the other hand, the combination of (a) an emphasis on the desirability of parsimony and highly simplified micro-level models (Lindenberg, 1992; Macy & Willer, 2002), and (b) the practice of empirically assessing these models using the criteria of generative sufficiency (Epstein, 1999), has encouraged the construction of models that assume homogenous agents. That is, if a given macro-outcome can be generated (explained) by a simpler model—which e.g., does not have to specify a particular kind of variation in behavior across agents or social environments—then all else equal, this model is to be preferred. In the absence of detailed micro-level data to investigate the effects of heterogeneity, this is hard to dispute. Even so, this practice appears to come with a non-negligible risk: that we systematically construct models and explanations with artificially low heterogeneity, and, as a result, limit their real world relevance.

However, with the rise of the computational and digital revolution, and the increasing availability of population-scale data, it is becoming increasingly possible to interrogate the empirical validity of these homogeneity assumptions,
and to move beyond them if this is motivated, by incorporating the essential heterogeneities observed in real social environments. Thus, complementing the first objective of this thesis—which, as noted above, is to investigate how the methods and data emanating from the digital and computational revolution can be used to develop realistic mechanism-based explanations—the second objective of the thesis is to investigate the implications that (explicit or implicit) homogeneity assumptions have for our understanding of how networks affect real-world micro-macro processes. More specifically, the second objective of this thesis is to investigate how homogeneity assumptions influence our understanding of how networks affect the processes of workplace segregation (Essay I), status-quality decoupling (Essay II), and urban scaling (Essay III).

The remainder of this thesis introduction is organized as follows. In the next section, I introduce the conceptual and analytical framework that will be used to investigate the research questions examined in the thesis, which is followed by summaries of the different thesis essays. Then, in the final section, a discussion of the results and conclusions of the thesis is presented.
An analytical framework that bridges AS and CSS

The previous section argued that, on the one hand, AS faces considerable challenges in realizing its theoretical ambitions as a result of the principle of empirical realism and the very particular nature of the data that is required to study interdependent populations, while on the other hand, the data and methods emanating from the CSS revolution hold considerable potential for dealing with these challenges. However, despite this potential, work that combines the explanatory principles of AS with the methods and data of the CSS revolution remains scarce within the AS community. The purpose of this section is to begin addressing the first objective of this thesis by introducing a conceptual and analytical framework that integrates AS and CSS as a means of improving our ability to empirically study micro-macro processes.

Conceptual framework

As discussed in the previous section, while the Coleman diagram has considerable utility in many situations, it is not sufficiently detailed when it comes to the interaction dynamics that lie at the heart of AS. Consequently, this section begins by introducing an alternative—but ultimately complementary—way of conceptualizing micro–macro processes that is centered on interaction dynamics.

How might we arrive at such a conceptualization? One approach would be to consider the features that unify some of the most celebrated mechanisms (and models) in the field of AS. Following this approach, I would suggest that these celebrated examples have three core aspects in common (with these aspects also incidentally reflecting different elements of common definitions of social mechanisms, e.g., Hedström and Ylikoski, 2010; Glennan and Illari, 2017):

1. A specification of some sort of interdependence;
2. A specification of the dynamics implied by such interdependencies;
3. A derivation of the qualitative outcomes that are expected when such dynamics unfold over time.

In order to facilitate systematic description and empirical translation of these core aspects, steps 1–2 can be formalized. Let $S_{it}$ and $A_{it}$ represent

---

8These examples include classics such as Schelling’s segregation model (Schelling, 1971), Merton’s notion of Matthew effects (Merton, 1968) and self-fulfilling prophecies (Merton, 1948), Granovetter’s threshold model (Granovetter, 1978), but also more recent examples like van de Rijt’s self-correcting influence mechanism (van de Rijt, 2019).
the social environment\(^9\) and the action of actor \(i\) at time-point \(t\), respectively. Further, let \(M_t\) represent the macro-outcome of interest at time-point \(t\). With these three variables as its key elements, the diagram in Figure 6 constitutes an abstract representation of local micro-macro dynamics.\(^{10}\) Arrows 1 and 4 (linking \(\Delta S \rightarrow A\)) capture how changes in individuals’ social environments affect their probability of engaging in a certain action. These arrows reflect step 1 above (interdependence) and are closely related to what have elsewhere been referred to as influence-response functions (Lopez-Pintado & Watts, 2008). The motivation for describing interaction at this level—exempt of action theory—is discussed below. The dashed arrows (3 and 6) that link \(A \rightarrow \Delta M\) represent the direct effect of action \(A\) on the macro-outcome of interest \(M\), while arrows 2 and 4 (linking \(A \rightarrow \Delta S\)) represent the direct effect of the actor’s action, \(A\), on the social environments of other actors. Considered jointly, arrows 1–6 reflect the second step above (the implied dynamics of interdependence).

![Figure 6: A dynamics-centered conceptualization of micro-macro processes.](image-url)

To illustrate these three steps more concretely, let us consider a familiar process, the Schelling segregation model (see Table 1 for further examples). Schelling’s model showed that even individuals who were willing to live in neighborhoods with a minority of their own kind tended to end up living in

\(^9\)I use the term “social environment” to refer to other actors who can in some way influence the focal actor. This includes direct influence, e.g., where the focal actor observes a friend making a particular consumption decision and mimics this behavior. It also includes indirect influences, such as unknown others dropping litter in the focal actor’s physical environment (Keizer et al., 2008; Keuschnigg & Wolbring, 2015; Wilson & Kelling, 1982)

\(^{10}\)Parts of this schema are elaborated in more detail in Arvidsson and Gimenez (forthcoming).
neighborhoods with a strong majority of their own kind. Why did this happen? A detailed examination of the interdependencies involved, and their implications for the dynamics of the process, provides the answer:

1. **Interdependence**: Individuals are sensitive to the attributes of their neighbors and hence to their neighbors’ residential moves (arrow 1,4).

2. **Implied dynamics of interdependence**: When an individual moves (e.g., node $A_i$), the individual’s kind becomes (a) less common in the individual’s previous neighborhood and (b) more common in her new neighborhood. This has the first-order effect of increasing the segregation level (arrow 3) and the second-order interdependency effect of making the previous neighborhood more attractive to individuals of the other kind and the new neighborhood more attractive to those of the same kind as the individual in question (arrow 2).

3. **Accumulation of dynamics**: Thus, every move on the one hand increases segregation\(^{11}\) (arrow 3,6), and on the other, increases the chance of additional moves of the same type (arrow 2,5). At the beginning of the process, when individuals are randomly allocated to neighborhoods, few live as members of the minority in their neighborhood. But, when some in the minority group leave, they change the neighborhood composition for others who were previously content, such that they also leave, and so on, creating a self-reinforcing chain of segregating moves.

As this dissection of the Schelling model illustrates, the objective is not to delineate all possible action sequences but to consider particular kinds of actions ($A$), particular kinds of changes in individuals’ social environments ($\Delta S$), and particular kinds of changes in macro-outcomes ($\Delta M$), and to link them together sequentially. Being clear and explicit about these three steps brings to the fore our assumptions about the dynamics of the interaction processes. This will help us to identify the core empirical questions that must be answered, and to assess their expected macro implications.

While the proposed diagram (Figure 6) and the Coleman diagram (Figure 5) have different objectives, they are nonetheless related to each other in meaningful ways, and it is therefore worth comparing and contrasting the two. First, they both attribute change in the states of macro variables to micro

\(^{11}\)It is worth noting that there are versions of Schelling’s segregation model in which this is not the case, with agents instead moving to a random square. Here, I depart from Schelling’s original formulation of the model, which does not assume such random mobility (Schelling, 1971, 1978).
behavior, and they both sketch macro→micro and micro→macro arrows (albeit of slightly different kinds). Second, they differ in that:

(a) They operate at different levels of abstraction: the Coleman diagram summarizes a hypothesized process at a higher level, referencing an initial macro state and a final macro state, with one macro→micro arrow, one micro→macro arrow, and only one kind of action. By contrast, the schema in Figure 6 sketches intricate details about the micro dynamics, with potentially multiple types of actions, and with the direct effect of a given action on the macro-level outcome being considered at each step.

(b) As noted earlier, the Coleman diagram does not highlight interaction dynamics, which by contrast is the key feature of the diagram in Figure 6. See Ylikoski (2016) for an extension of the Coleman diagram with a feedback loop.

(c) The diagram in Figure 6 collapses the macro→micro and micro→micro arrows of the Coleman diagram into a single $\Delta S \rightarrow A$ arrow. A discussion of the motivation for, and the meaning and implications of this difference follows below.

(d) The diagram in Figure 6 differentiates between two types of macro-level variables, $M$ and $S$. The first of these is also present in the Coleman diagram, while the second is not; the second is actor-specific and describes the social environment of a specific actor or type of actor.

Perhaps not surprisingly given the basis on which Figure 6 was derived, but it should nonetheless be noted that the figure is consistent with the core principles of analytical sociology, as expressed in Hedström (2005): (1) methodological individualism (Udehn, 2002), i.e., that it is through the actions of individuals that societies change, and (2) social mechanisms are defined as consisting of entities $(i, j, k)$ and their activities $(A)$, that are organized $(\Delta S \rightarrow A \rightarrow \Delta S \rightarrow A \ldots)$ so as to bring about change in the system $(\Delta M)$. Additional notable properties of the diagram in Figure 6 are that (i) it naturally encodes heterogeneity by referencing actor types and social-environment types, (ii) it describes both direct and indirect effects of actions, and therefore allows for descriptions of more complex and changing feedback processes (Ylikoski, 2016), sequential concatenation of mechanisms (Gambetta, 1998), and for the differentiation of self-reinforcing and self-correcting processes (Merton, 1968; van de Rijt, 2019), and (iii) it facilitates the identification of generalizable dynamics and thus promotes the ambitions of middle-range theories (Merton, 1949).
A few remarks should be made concerning (c), i.e., the choice to collapse the macro→micro and micro→micro arrows of the Coleman boat into a single $\Delta S \rightarrow A$ arrow. First, this is not meant to suggest that action theory is irrelevant. On the contrary, action theories may play an important role in motivating or explaining the presence of a given $\Delta S \rightarrow A$ arrow. The main reason for collapsing the two is that doing so provides a description of the process that remains at the level of observables, which facilitates formulations of micro-macro dynamics that are empirically testable. Second, as other scholars have noted (e.g., Lopez-Pintado and Watts, 2008; Mäs, 2021; Hedström, 2021), if the objective is to study how individual behaviors aggregate, then it is not strictly necessary to specify assumptions at the level of the mind, since it is sufficient to know how a given input (influence) results in a given output (response), or what Lopez-Pintado and Watts (2008) have labelled *influence-response functions*. Third, specifying interactions at this level comes with additional benefits. In particular, because different action theories may correspond to the same influence-response function, the specification of influence-response functions (a) facilitates the transportability of identified micro-macro dynamics to other domains, and (b) if the micro-macro dynamics for a given influence-response function are already well understood, then this information can be imported—thus reducing the independent rediscovery of findings that are already known, and promoting cumulative science (Lopez-Pintado & Watts, 2008; Mäs, 2021). More generally, by decomposing interaction dynamics into generic—sequentially ordered—elements at the level of influence-response functions, this (c) facilitates the identification of *structural similarities* between disparate social processes. As an example, consider the models listed in Table 1. Although Schelling’s segregation model, Granovetter’s threshold model, and Merton’s notion of self-fulfilling prophecies describe rather different processes, their dynamics share some fundamental structural similarities. By decomposing them following the logic of the proposed conceptual framework, it is easy to recognize, for example, that in each of them, every individual’s action (i) has a well-defined directional effect on the macro-outcome in question (increases segregation, increases size of riot, decreases solvency of bank), and (ii) increases the chance of more actions of the same kind. That is, all three describe self-reinforcing social influence processes. Decomposing social processes in this way also helps in identifying important *differences*. Considering again the examples described in Table 1, one quickly realizes that there are (iii) differences in the types of social environments involved: individuals on the adjacent squares of a grid (Schelling), everyone in the population in question (Granovetter), and individuals’ friendship networks (Merton), and (iv) differences in the nature
of the influence-response functions: a general threshold (Schelling), personal thresholds (Granovetter), and a general threshold of 1 (Merton).\footnote{Note that Merton did not specify the social environment or the social-influence function concretely when discussing the bank run example, but this is arguably a plausible interpretation.}
Examples of central mechanisms and models in the field of AS

<table>
<thead>
<tr>
<th>Model</th>
<th>$\Delta S_{it} \rightarrow A_{it}$</th>
<th>$A_{it} \rightarrow \Delta S_{jt+1}$</th>
<th>$A_{it} \rightarrow \Delta M_{t+1}$</th>
<th>Dynamics and outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schelling’s (1971) segregation model</td>
<td>Focal individual $i$ is sensitive to the composition of neighboring individuals $X_i$. $X_i$ changes as a function of residential moves by others, and if $X_i$ falls below a given threshold $\tau$, $i$ moves to a neighborhood satisfying $\tau$.</td>
<td>Individual $i$’s move, in turn, changes the composition for others ($j$), potentially making it so that their threshold no longer is satisfied.</td>
<td>Per definition, since $i$’s origin did not satisfy $\tau$, while the destination did, this implies that a move between the two increases segregation (James and Tauber 1985).</td>
<td>Every move, on the one hand, increases segregation, and on the other, increases the chance of additional moves of the same type, creating a self-reinforcing chain of segregating moves, thus enabling segregation to increase even under a tolerant $\tau$.</td>
</tr>
<tr>
<td>Granovetter’s (1978) threshold model</td>
<td>Focal individual $i$ is sensitive to the number of others who previously have adopted a behavior. When the number reaches a threshold $\tau$, $i$ also adopts.</td>
<td>Individual $i$’s adoption increases the adoption count for others ($j$), potentially making it so that their threshold ($\tau_j$) is satisfied too, making them adopt as well.</td>
<td>Every adoption increases the total adoption count, which defines the macro outcome of interest (e.g., the size of a riot).</td>
<td>Every move, on the one hand, increases the adoption count, and on the other, increases the chance of additional adoptions, creating a self-reinforcing chain of adoptions. Depending on the distribution of $\tau$, just a few or the whole population may adopt the behavior.</td>
</tr>
<tr>
<td>Merton’s (1948) Self-fulfilling prophecies (bank runs)</td>
<td>Individuals are socially influenced by their friends’ interactions with shared banks. If a friend of $i$, believing that the bank $X$ is soon to become insolvent, retracts all her money, $i$ follows suit and does so too.</td>
<td>Individual $i$ retracting all her money from bank $X$, in turn, social influences $i$’s friends ($j$), increasing the chance that they too will retract their money.</td>
<td>Every retraction decreases the solvency of the bank, which defines the macro outcome of interest.</td>
<td>Every retraction, on the one hand, reduces the solvency of the bank, and on the other, increases the chance of additional retractions, creating a self-reinforcing chain of retractions. Depending on the nature of social influence, such a dynamic could result in bank $X$ eventually becoming insolvent, thus making true what initially only was falsely believed.</td>
</tr>
</tbody>
</table>

Table 1: Illustration of how celebrated mechanisms/models in AS can be conceptualized following the diagram in Figure 6.
Introduction

Analytical framework

The conceptual framework introduced in the previous section demonstrates how micro-macro processes can be disassembled into a set of generic, sequentially ordered elements. The objective of this section is to translate this conceptual framework into an analytical framework that can be used in empirical research. In the pages that follow, I first consider the central methodological challenge associated with turning this into a framework that is applicable to empirical analysis of observational data, which involves the identification of social influence effects. Having considered how this challenge might be addressed, I outline a general analytical strategy that details the practical steps to be taken in order to empirically study micro-macro dynamics. Lastly, I present a practical demonstration of this analytical strategy using simulated data.

Identification of social influence effects. As is underscored by the examples presented in Table 1 (e.g., Schelling’s segregation model), in order to comprehend how micro-behaviors generate macro-outcomes we need to understand how individuals interact and influence one another. In terms of the conceptual diagram presented in Figure 6, this corresponds to empirically identifying and estimating the $\Delta S \rightarrow A$ arrows. This, however, is far from easy in observational settings. When we observe a group of friends acting in a similar manner, this might be due to them influencing each other and coming to align their behaviors as a result of this influence. Another equally plausible explanation, however, is that they became friends because of their shared behavior (or other traits correlated with this behavior), i.e., because of homophily (McPherson et al., 2001). Distinguishing between genuine social influence and homophily is important, because although they may leave similar data traces (i.e., correlated behaviors among interconnected individuals), the dynamics and the potential for social change differ radically between the two (Aral et al., 2009; Steglich et al., 2010). Similarly, common exposure to an external source of influence (e.g., policy changes or algorithmic recommendations) often leaves data traces that are hard to distinguish from those of social influence. To distinguish between genuine social influence and the effects of homophilic selection and common exposure in observational settings, we need rather specific types of data. First, a sufficiently granular, longitudinal dataset is required to identify the order of events; did an individual change her behavior before or after seeing her friends doing so? Second, the relations through which a given type of influence flows must be observed. Third, all influences on both the likelihood of Ego adopting the behavior and the likelihood of Ego being exposed
to the behavior through Alters (the “treatment”) need to be accounted for (Shalizi & Thomas, 2011).

In order to explicate the challenges, assumptions, and possible solutions for the identification of social influence effects based on observational data, let us formalize this problem. Adopting the basic setup described by Shalizi and Thomas (2011), let network ties—through which influence flows—be represented by a binary variable $A_{i,j}$, which is equal to 1 if there is an edge from $i$ to $j$, and 0 otherwise. Further, let $X_i$ and $Z_i$ be observed and unobserved traits of $i$, respectively, and finally, let $Y_{i,t}$ denote the behavior of interest—that which is to be diffused—for $i$ at time $t$.

Figure 7 illustrates why the identification of social influence is often difficult in empirical settings. It presents a so-called causal directed acyclic graph (DAG), that encodes causal relationships between the observed (gray nodes) and the unobserved (white nodes) variables described in the previous paragraph. The social influence effect is captured here by the arrow between $Y_{j,t}$ ($j$’s outcome at $t$) and $Y_{i,t+1}$ ($i$’s outcome at the subsequent time point, $t + 1$). In order to identify this effect, one must adjust for factors that influence both the exposure ($A_{i,j}, Y_{j,t}$) and the outcome ($Y_{i,t+1}$). In the case of Figure 7, both $X$ and $Z$ have this type of influence. However, while $X_i$ and $X_j$ are observed and can therefore be adjusted for, $Z_i$ and $Z_j$ are unobserved. Applying Judea Pearl’s do-calculus (Pearl, 2009) it becomes clear that, as a result of $Z_i$ and $Z_j$ being unobserved, the social influence effect is not identifiable. This is because there exists a so-called open back-door path between the exposure ($Y_{j,t}$) and the outcome ($Y_{i,t+1}$). Intuitively, $j$’s outcome at time $t$ ($Y_{j,t}$), i.e., the exposure, is informative about $j$’s unobserved characteristics $Z_j$, which in turn is informative about $i$’s unobserved characteristics ($Z_i$) when $i$ and $j$ are friends. Finally, $Z_i$ is informative about the probability of $i$’s outcome at $t + 1$. This path is highlighted in red in Figure 7.

The scenario depicted in Figure 7 reflects the typical case using data from traditional sociological research designs—which often contain a number of sociodemographic variables ($X$), but leave a lot of the fundamental drivers of selection unobserved ($Z$). The computational and digital revolution has improved our ability to identify social-influence effects in at least three ways. First, instead of relying on just a few sociodemographic control variables, large-scale relational datasets that contain rich information about individuals’ past behaviors and the social environments in which they are embedded make it possible to adjust for the central preferential forces driving selection (Easley et al., 2020; Eckles & Bakshy, 2020; Livneh et al., 2020). For example, in a

---

13To be clear, the variable $A$ here is not related to the variable $A$ in Figure 6.
Figure 7: Causal DAG with latent homophily making social influence unidentifiable.

recent study by Eckles and Bakshy (2020), the authors used Facebook data to examine the effect of being exposed to a friend sharing a particular URL on a focal individual’s probability of subsequently doing the same. By adjusting for high-dimensional measures of past behavior on Facebook, they found that peer effects estimated using observational data were statistically indistinguishable from the “true” effects obtained from a randomized field experiment. In other words, past behavior can serve as a useful proxy for individuals’ preferences, allowing researchers to net out social influence effects from homophily effects.

Second, in addition to past choices, digital data often contain detailed network information that allows us to adjust for latent homophily using subtle measures of relatedness between individuals and the social positions they occupy. To obtain such measures, recent developments in machine learning for inferring latent network locations (e.g., Grover and Leskovec, 2016; Gollini and Murphy, 2016) can be leveraged. In a nutshell, the idea here is that by knowing individuals’ positions in the network, we implicitly have information about the traits on which the selection occurred that led to the individuals finding themselves in the positions they occupy. Thus, by adjusting for network locations, we adjust for proxies of unobserved factors that drive tie-formation. In other words, in common to these two approaches is that although the underlying confounders themselves may not be observed, they can be proxied, either by high-dimensional measures of past behavior or by inferred network locations. Figure 8 displays a DAG corresponding to this scenario, with an additional variable $Q$ that is generated from the unobserved confounder $Z$. To the extent that $Q$ provides a good proxy of $Z$, $Q$ closes the back-door path
between $Y_{j,t}$ and $Y_{i,t+1}$, and the social-influence effect then becomes possible to identify (McFowland III & Shalizi, 2021).

**Figure 8**: Causal DAG containing (a) observed proxy variable $Q$ of unobserved confounder $Z$, and (b) instrumental variable $W$, both separately making social influence identifiable.

Third, and finally, the wide proliferation of digital trace data has increased the possibility of using creative instrumental variable and regression discontinuity designs to identify social influence effects. For example, Aral and Nicolaides (2017) employed an instrumental variable approach to identify social influence on exercise behaviors within an online social network. Local weather conditions serve as a source of exogenous variation in exercise behaviors that then spill over to friends in other cities. Figure 8 also considers this design, containing the additional variable $W$, which is an instrument (individual $j$’s local weather). To the extent that $W$ provides a good instrument, it becomes possible to identify the social-influence effect (Angrist et al., 1996).

Beyond the identification and estimation of average treatment effects, computational techniques and novel data sources have also advanced our capacity to estimate heterogeneous social influence effects. First, the high-dimensional nature that characterizes many digital datasets allow for the measurement of novel individual, relational, and contextual variations that may condition social influence. Second, the large-scale nature of these datasets means that large sample sizes can be obtained even in smaller subsets, thus also enabling the robust estimation of conditional social influence effects. Third, advances at the intersection of econometrics and machine learning have enabled us to search inductively for heterogeneous effects in an efficient and statistically sound manner (Athey & Imbens, 2016; Brand et al., 2021; Daoud & Johansson, 2019; Wager & Athey, 2018). Applied to large-scale networked data, this has the
potential to allow us to inductively seek out how social influence varies across local social contexts, and to identify the most important sources of heterogeneity. For example, a recent study by Yuan et al. (2021) adapted tree-based methods introduced by Athey and Imbens (2016) to identify heterogenous effects of social influence depending on the types of local network structures in which individuals were embedded.

**Analytical framework for empirical studies of micro-macro processes.**

On the basis of the advances described in association with Figure 8, I here provide a practical translation of the conceptual framework in Figure 6, developing it into a general analytical strategy for identifying the elements of social-interaction processes that are of crucial importance for collective outcomes in large N/P/T/I data, where \( N \) = the number of observations, \( P \) = the number of variables, \( T \) = the number of time-points, and \( I \) = the number of interactions. A visual example demonstrating the steps described below is provided in Figure 9:

1. **Isolate** actors in *time* and *social space*, e.g., the egocentric network of individuals or their residential neighborhoods at different points in time.

2. **Measure changes** in actors’ *social environments*, e.g., that one of the focal individual’s peers adopts a new, particular type of behavior at time-point \( t \) (\( \Delta S_{it} \)).

3. Use the methodological developments described in association with Figure 8 above to **contrast** the observed behavior among those exposed to the change in the social environment with the estimated counterfactual behavior\(^{14}\) that would occur without exposure; e.g., compare exposed individual’s behavior at time \( t + 1 \) with a very similar (but unexposed) individual’s behavior at time \( t + 1 \). The difference between the two provides an estimate of the causal effect of social influence, i.e., of how the change in the focal individual’s social environment affected her subsequent behavior (\( \Delta S_{it} \rightarrow A_{it+1} \)).\(^{15}\)

4. To **link micro and macro**, measure the direct effect of the exposed individual’s action on the macro outcome of interest, e.g., the direct effect

\(^{14}\)Although here formulated in terms of a matching design, step 3 can also be implemented using other causal inference strategies, e.g. regression adjustment, instrumental variables, etc.

\(^{15}\)It is worth underscoring that although what is emphasized here is a pure social influence process—where a relational structure is implicitly assumed—the general strategy is generalizable to also identify selection dynamics (Holme & Saramäki, 2012; Overgoor et al., 2019; Snijders et al., 2010; Stadfeld & Amati, 2021). See Arvidsson et al. (2023) for a further discussion on this topic.
of individual $i$’s behavior on the segregation or inequality of the market ($A_{it+1} \rightarrow \Delta M_{t+1}$).\textsuperscript{16}

5. **Capture the dynamics**: consider how the behavior of the exposed individual affects the social environment of other actors ($A_{it+1} \rightarrow \Delta S_{jt+1}$), and how this in turn affects their behavior ($\Delta S_{jt+1} \rightarrow A_{jt+2}$), and so on.

6. **Assess macro implications**: Use the findings from 1-5 to specify an agent-based simulation model (Macy & Willer, 2002), and use this model to perform virtual experiments examining how the macro-outcome is affected by interventions on the identified dynamics, e.g., by letting the social-influence effects be present or absent.

---

\textsuperscript{16}In the case of segregation, for example, the principle of transfer (James & Taeuber, 1985) implies that any move between two workplaces/neighborhoods/schools that increases the difference in the sociodemographic composition between origin and target also increases the segregation of the market, and vice versa. In order to properly establish this effect, a sufficient granularity in the data is required, such that compositions can be computed prior to the move—yet again underscoring the importance of large N/P/T/I data.
the properties of the identified dynamic, its generalizability, and its scope conditions. The second type of simulation exercise studies the implications of the identified dynamic in the context of the actual process in which it was observed, and is thus useful for assessing its role and explanatory power for the empirical case under consideration. These two versions should not be viewed as substitutes for one another, but rather as complementary. A second remark about step 6 is that calibrating steps 1–5 means that we do not satisfy ourselves with establishing generative sufficiency, but that we also ensure that the collective outcome was brought about in the right way, i.e., in the same way in the simulation model as in the real world, or at a minimum, that we make sure that there is a match between the most central aspects of the simulated and observed processes (León-Medina, 2017a; Macy & Flache, 2009).

A remark should also be made about step 5. One way of performing this step is to trace observed sequences of action in the data and to quantify the expected dynamics. Another way is to derive the dynamics that are possible in principle based on the results of step 1–4. For example, having observed that there is a threshold of 2 for the adoption of some new behavior, we can derive the expected dynamics in a given local social structure and identify whether or not a process of social contagion can be sustained (see Figure 9).

Challenges associated with large N/P/T/I data. While the availability of large N/P/T/I data is the fundamental reason that the analytical strategy outlined above is plausible, this type of data also implies considerable practical and statistical challenges. In particular, the extraction of theoretically meaningful and robust measures from complex, high-dimensional, and noisy data, such as digital trace data, remains a fundamental challenge in computational social science (Sen et al., 2021). Methods developed by computer scientists for dealing with these emerging types of datasets are often optimized for objectives that are different from those pursued by social scientists. Instead of maximizing interpretability, they often maximize black-box predictability. With regard to the analytical framework that I have outlined, this can crucially inhibit our ability to extract meaningful measures of all the key components \((A, S, \text{ and } M)\). For example, central to the study of micro-macro dynamics is the understanding of how social influence effects vary by differences in contextual and relational features. And when the data consist of digital traces, establishing these features is not always straightforward, since this often involves inferring them from complex high-dimensional spaces. In high-dimensional settings, one cannot easily identify all dimensions of potential relevance by hand. Even if this was practically feasible, however, the theoretical relevance of a specific dimension (e.g., a word in a document–word matrix) is often hard to judge
divorced from its context. In addition, in many high-dimensional datasets, such as textual data, individual dimensions (words) are sparse and noisy. For these reasons, researchers often infer lower-dimensional representations of higher-dimensional spaces. However, ensuring that the distilled dimensions describe theoretically meaningful aspects of interest remains a considerable challenge, often limiting the applicability of off-the-shelf text analysis tools in social science research.

Being able to infer theoretically meaningful features of individuals and their social environments is not only relevant for the conditioning of social influence effects, but often also for their identification. For instance, a researcher may expect some particular abstract dimension of language (e.g., tonality, sentiment, gender) to be a confounder that affects both the probability of exposure and the focal individual's probability of adoption. Although it may be possible to control for black-box features of language that implicitly capture these aspects, adjusting for the full space can be inefficient and may even affect the identification negatively (e.g., by controlling for collider variables). Having interpretable dimensions further fosters causal inference, for example by enabling balance (in the case of matching) to be evaluated specifically for these dimensions.

Thus, as a part of addressing the first objective of this thesis—which concerns how CSS can be utilized by AS—essays II and IV propose new ways of using natural language processing techniques to construct theoretically meaningful measures of individuals and their social environments.

Demonstration of the analytical framework using simulated data. To illustrate how the framework can be used in practice, this section presents a small case study on the spread of behaviors in a fictive population based on simulated data. In a nutshell, the idea is to create a toy world in which we know the real social influence mechanisms that operate (i.e., the “ground truth”), and then examine whether the framework can help us identify the key dynamics of this process and their macro implications.

To create this toy world, I use the DAG in Figure 7 as a point of departure. Starting with $X$ and $Z$, the two confounders, each agent $i \in 1, ..., N$ is assigned two random values drawn from two independent normal distributions. Then, on the basis of $X$ and $Z$, I let the tie-formation process be governed by an Exponential Random Graph Model (ERGM) with basic structural terms and homophilic preferences along $X$ and $Z$. The result of this process is an $N \times N$ matrix, $A$, which specifies who is connected to whom. With the structure of relations specified, I let the adoption process that is to unfold on top of the network

---

17 More specifically, $X_i \sim N(\mu_X = 100, \sigma_X = 10)$, $Z_i \sim N(\mu_Z = 100, \sigma_Z = 10)$. 

33
be defined by agents selecting between two items, $I_1$ and $I_2$, at every time-point $t$. This process is assumed to be governed by a simple binary discrete choice model, with $X$ and $Z$ as well as past adoptions of alters dictating the probability of adoption. To incorporate heterogeneity into agents' preferences and responses to social influence, I let each agent $i$ be randomly assigned to one of two categories—the green category or the red category, $C_i \in \{R,G\}$, with a 3:1 representation of $R$ relative to $G$—and assume that both social influence and preferences vary systematically by agent type in accordance with Tables 2–3. The result of this adoption process is an $N \times T$ matrix, $Y$, which stores all the adoptions across agents and across time.

<table>
<thead>
<tr>
<th>Alter</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ego</td>
<td>R</td>
<td>G</td>
</tr>
<tr>
<td>R</td>
<td>0.25</td>
<td>-0.25</td>
</tr>
<tr>
<td>G</td>
<td>0.00</td>
<td>0.25</td>
</tr>
</tbody>
</table>

Table 2: Heterogeneous social influence effects

<table>
<thead>
<tr>
<th>Item</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ego</td>
<td>$I_1$</td>
<td>$I_2$</td>
</tr>
<tr>
<td>R</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>G</td>
<td>0.8</td>
<td>0.2</td>
</tr>
</tbody>
</table>

Table 3: Heterogeneous preferences

With the true data-generating mechanism specified in this way, the idea now is to generate finite and partial ($Z$ being unobserved) samples from this process, and to use the proposed analytical framework in an attempt to recover the key dynamics of the process. Following steps 1–2 of the framework, I begin by isolating agents in time and social space—carving out their respective ego-centric networks at each time-point $t$—and identify the composition of items that each agent was exposed to at $t - 1$. Then, for step 3, in order to estimate the effect of such exposures in the presence of latent homophily, I use a recent extension of Hoff et al.’s (2002) seminal latent space model (Gollini & Murphy, 2016) to infer observable proxies, $W$, for the unobserved variable $Z$. Presenting the results for step 3, Figure 10 shows that causal inference which relies solely on observed confounders (i.e., $X$) produces severely biased social influence effects in the presence of latent homophily. By additionally adjusting for the inferred proxies $W$, however, this bias is almost completely wiped out, and the estimates closely approximate the true social influence effects.\(^{18}\)

Steps 4–5 of the framework involves identifying the implied dynamics of the estimated social influence effects, and ask us to formulate expectations about\(^{18}\)

\(^{18}\)It is worth noting that while I here pre-partitioned the data in order to estimate heterogeneous social influence effects, it is also possible to inductively discover the key sources of heterogeneity using e.g., the class of methods proposed by Susan Athey and colleagues (Athey & Imbens, 2016; Wager & Athey, 2018).
the way events typically unfold in local social contexts. This entails asking how particular types of actions by particular types of agents are expected to influence, in a direct way, (a) particular others, and (b) the macro-outcome of interest, and doing so sequentially. In the current case study, we have two types of actions (adopting $I_1$ or $I_2$) and two types of agents (green and red). Further, suppose our macro-outcome of interest is the relative market share of the two items, with the ambition being to explain why a certain item is more/less likely to end up on top. With these things in mind, let us reconsider Tables 2–3 and Figure 10. Two central dynamics are then brought to the fore. First, the positive social influence that is observed within agent types (see e.g., the primary diagonal of Table 2) suggests a self-reinforcing type of dynamic: If an actor of type $X$ adopts an item of type $Y$, then this has the first-order effect of increasing the market share of item $Y$ and the second-order effect of increasing the probability of further adoptions of the same type by those belonging to the same group of actors (for an illustration, see bottom panel of Figure 11).

Second, the negative influence that green agents exert on red agents (see e.g., the secondary diagonal of Table 2) implies a particular type of counteracting

---

Figure 10: Estimated social influence effects in the presence of latent homophily.
dynamic: When green agents adopt item $I_1$—which they are likely to do given the makeup of their preferences and susceptibility to influence—then this has the first-order effect of increasing the market share of item $I_1$. However, and importantly, such $I_1$ adoptions have the second-order effect of reducing the probability that red-agent peers will adopt $I_1$, instead pushing them to adopt $I_2$. From the perspective of status–quality decoupling—which concerns how aggregate independent preferences (quality) map onto realized outcomes (status) (see e.g., Lynn et al., 2009; van de Rijt, 2019)—the green agents’ quality-reinforcing adoptions of $I_1$ have the adverse indirect effect of making red agents more likely to adopt $I_2$ items, thus setting in motion chains of status–quality decoupling events (for an illustration, see top panel of Figure 11).

For the 6th and final step of the analytical framework, in which the objective is to examine the cumulative effects of the identified dynamics from steps 1–5, I develop a simple ABM based on the estimated social influence effects of

---

20Item $I_1$ is here considered to have the higher quality because it preferred on average.
To elucidate the effect of social influence, and the role played by the identified dynamics, a comparison will first be made with a scenario in which no influence is operating, with this contrast providing a direct measure of the overall effect of social influence on the macro-outcome. Second, and importantly, a comparison will also be made with a counterfactual scenario in which the counteracting dynamic is surgically blocked, thus enabling an assessment of the role played by this dynamic for the macro-outcome. Figure 12 presents the results of this simulation exercise, showing the resulting market shares of the two items $I_1$ and $I_2$ under different scenarios.

![Figure 12: Agent-based simulations based on the social influence effects of Figure 10.](image)

To begin with, we can observe that, without any social influence, item $I_1$ obtains the largest market share. This result can be attributed to the fact that green agents have strong preferences for $I_1$, while red agents have neutral preferences. In contrast, when social influence is activated—and the simulation is governed by either the ground-truth model or the model with inferred proxy variables—$I_2$ ends up having the largest market share. In other words, social

---

21 The ABM is additionally calibrated structurally by using the same, observed network structure.
22 The bar chart reflect averages across 100 separate runs, with each run consisting of 20 iterations.
**Introduction**

Influence has a decoupling effect whereby the resulting macro-outcome differs from what would be expected based purely on agents’ aggregated preferences. What role does the identified counteracting dynamic play in this decoupling process? To examine this, a counterfactual scenario is considered in which this dynamic is blocked (by cancelling the negative influence that flows from green to red agents). Figure 12 shows that under such an intervention (condition = “CD blocked”), the decoupling effect of social influence disappears, and \( I_1 \) obtains the largest market share.

In addition, something should also be said about the result obtained when adoptions are simulated on the basis of the *biased* social influence model (which only adjusts for observed confounders, \( X \)). In this simulation, item \( I_1 \) completely dominates the market, with a ~95% market share. The reason for this is that, in contrast to both the ground-truth model and the model adjusting for proxy variables, the biased model contains strong positive social influence effects also *between* agent types. This result thus demonstrates the crucial importance of the careful identification of social influence effects, and by extension the relevance of the CSS revolution. Bias at this level can—as exemplified here—lead us to draw the wrong conclusions about the dynamics of a social process, their cumulative effects, and thus the micro-macro link.

In sum, this case study shows that, on the one hand, latent homophily posits considerable challenges for empirical research that attempts to establish micro-macro linkages, and that traditional research designs—which rely only on weakly correlated control variables such as demographic variables—are likely to fall short. On the other hand, and more optimistically, the case study also shows that by following the steps of the proposed analytical framework—which combines the explanatory principles of AS with the types of data and methods that have emerged from the CSS revolution—such biases can be addressed, the true dynamics can be identified, and a correct mapping can be formulated.
Appended Essays

This section provides a brief summary of the essays included in the thesis. Essays I-III all rely in one way or another on the framework proposed in the previous section. Essay IV does not use this framework but rather builds upon it. More concretely, Essay IV develops a method that improves the description of $S$ and facilitates the identification of $\Delta S \rightarrow A$ arrows.

Essay I

The scholarship on labor market segregation commonly assumes that networks and network-based recruitment lead to a more segregated market (Bielby & Baron, 1986; Elliott, 2001; Reskin, B. F. et al., 1999; Trimble & Kmec, 2011). This conclusion is derived from the well-documented observation that individuals’ contact networks tend to be homophilous along demographic lines (McPherson et al., 2001). The homophily thesis suggests that if an individual with property $X$ joins an organization, then the probability of additional individuals with property $X$ joining the organization increases, and this kind of self-reinforcing process is expected to generate a more segregated market. What this line of research has overlooked, however, is that opportunities to form same-category ties can vary substantially from context to context and can place constraints on homophilic preferences.

Focusing on gender segregation and the heterogeneous mixing constraints that employees face in tie formation within organizations, Francois Collet, Peter
Hedström and I investigate the effect that networks created by individuals moving between organizations—which we refer to as mobility networks—have on the segregation of the market. In particular, following the logic of the conceptual model proposed in the Introduction, we draw on the seminal work of Peter Blau and colleagues (Blau, 1977; Blau & Schwartz, 2018) and identify a new mechanism, the Trojan-horse mechanism, which shows how networks can counteract the impact of segregating mobility events. We refer to this mechanism as the Trojan-horse mechanism because it shows, just as in the case of the soldiers in the hollow Trojan horse, how the recruitment of a new employee belonging to one specific category (e.g., a male) can open the gates to those of another category (e.g., females), thereby changing the composition of the organization in unexpected ways.

To test the predictions of the Trojan-horse mechanism, we adopt the analytical framework proposed in the Introduction and use a large-scale longitudinal register dataset with rich demographic and socioeconomic information, as well as detailed mobility records, for every individual and every organization that resided in the greater Stockholm Metropolitan area during the years 2000–2017. Our analyses first provides strong empirical evidence for the presence of the Trojan-horse mechanism, and counterfactual simulations then reveal that if this mechanism is commonly observed in a market, network-based mobility is likely to desegregate rather than segregate the market.

Essay II

Existing research on status hierarchies points to social influence as one of the central mechanisms enabling the emergence of a discrepancy between quality and status (Gould, 2002; Lynn et al., 2009; Manzo & Baldassarri, 2015; Salganik et al., 2006; van de Rijt, 2019). However, conflicting evidence and limited external validity mean that we still know little about the conditions under which social influence can lead to this form of decoupling in the real world. In particular, neither theoretical models nor experiments have allowed for selection (influence) to be driven by (conditioned on) homophily, and have therefore implicitly assumed artificially homogenized social environments and influence weights. It is therefore open to question whether van de Rijt (2019)'s conclusion that social influence is in general not strong enough to decouple individuals’ behavior from their preferences is likely to hold in most real-world settings.
To address these limitations, Peter Hedström, Marc Keuschnigg and I use the analytical framework proposed in the Introduction and confront the decoupling thesis using fine-grained, real-world behavioral data from Spotify that allow us to estimate heterogeneous social influence effects conditional on properties of individuals’ social environments, and thereby establish under what conditions decoupling is likely to occur. In particular, we propose a distinction between what we call “wide” and “narrow” social influence: While the narrow type of social influence either does not expose individuals to novel behaviors, or does not sufficiently motivate them to adopt these behaviors, and therefore only generates expected outcomes, the wide type of social influence is capable of bringing about the unexpected by both informing individuals about new behavioral alternatives and at the same time motivating them to adopt these novel behaviors. We furthermore postulate that the overlap between Ego’s and Alter’s preferences (homophily) is key to the operation of wide social influence: Some degree of preference compatibility is required for Ego to be susceptible to Alter’s influence, and some degree of preference incompatibility is required for Alter to expose Ego to alternatives outside her behavioral repertoire.

In order to test these hypotheses, we collected a large-scale digital trace dataset containing the adoption histories of 1.5M interconnected Spotify users via Spotify’s public API. Using a combination of high-dimensional matching (to estimate conditional social influence effects) and empirically calibrated agent-based modeling (to simulate counterfactual scenarios in which we either block or amplify wide social influence), we find strong support for the narrow/wide hypothesis, as well as for the central moderating role of homophily. We thus conclude that mere strength of social influence is not sufficient for decoupling to occur. Instead, for this outcome to occur, social influence needs to apply in the wide sense; to be effective also for alternatives outside individuals’ typical behavioral repertoires.

**Essay III**

Over the past decade, the literature on urban scaling has documented striking and seemingly universal relationships between city size and various urban quantities, showing how cities’ total outputs systematically increase more than proportionately with city size (Bettencourt et al., 2007; West, 2017). In addition, this literature has developed elegant mathematical models to explain such patterns (Arbesman et al., 2009; Bettencourt, 2013; Pan et al., 2013; West, 2017). While these models have demonstrated impressive predictive accuracy at aggregate levels, they have all overlooked the stark inequality
and heterogeneity that exist within cities. Similarly, empirical studies have relied on sums and means to quantify agglomeration effects, targeting the "average" resident in their interpretations (Bettencourt & West, 2010; Gomez-Lievano et al., 2016; Schläpfer et al., 2014; West, 2017). In reality, both human networking and productivity exhibit heavy-tailed distributions with some individuals contributing disproportionately to city totals (Alvaredo et al., 2017; Barabási & Albert, 1999; Newman, 2005; Pareto, 1896).

Using population-scale microdata from Sweden, Russia, and the United States—which provides detailed information on within-city distributions of urban indicators of interconnectivity, productivity, and innovation—Niclas Lovsjö, Marc Keuschnigg and I investigate the relative contribution of cities' tails to observed urban scaling relations. Furthermore, following the logic of the conceptual and analytical framework proposed in the Introduction, we postulate and test a new mechanism which we dub the size-dependent cumulative advantage mechanism, according to which large cities provide novel but heterogeneously distributed interaction opportunities needed for sustained growth at the micro level, and which predicts tail differences by city size at the macro level.

Acknowledging heterogeneity within cities, our findings first reveal that cities' tails—and their growth by city size—account for 36–80% of previously reported scaling effects and 56–87% of the variance in scaling between indicators of different economic complexity. Second, in support of the hypothesized size-dependent cumulative advantage mechanism, we find that those who were successful early on in their careers in large cities flourished to a greater extent—thereby increasingly distancing themselves from both the typical individual within their own city and from the tail individuals of smaller cities—while the typical individuals from both smaller and larger cities experienced almost identical trajectories. Third, and finally, using agent-based simulations we demonstrate that this mechanism is able to reproduce our key empirical findings, and we examine the theoretical conditions under which this occurs.

**Essay IV**

The increased availability of large digitized corpora and significant developments in natural language processing (NLP) have sparked a growing interest in the use of computational methods for textual data in the field of computational social science (CSS) (DiMaggio et al., 2013; Grimmer, 2010; Jockers & Mimno, 2013; Laver et al., 2003; Tsur et al., 2015). Word embeddings, a family of unsupervised methods for representing words as dense vectors (Mikolov, Sutskever, et al., 2013; Pennington et al., 2014), constitute one
such development. Although word embeddings have demonstrated strong performance on NLP tasks (Mikolov, Chen, et al., 2013; Mikolov, Yih, et al., 2013), limited interpretability and the unsupervised nature of word embeddings have restricted their use in CSS. Beginning to address these limitations, recent work has derived interpretable dimensions via post-processing in the form of antonym-pair vector algebra (Garg et al., 2018; Kozlowski et al., 2018) and ideal-point anchoring of antonym word-pairs (Lauretig, 2019).

A promising recent development is the formulation of word embeddings as probabilistic models (Barkan, 2017; Havrylov, Titov, et al., 2018; Rudolph et al., 2016; Vilnis & McCallum, 2015). Probabilistic models naturally enable the incorporation of domain knowledge through priors, and have been used to study language differences by time and by group. In this paper, Miriam Hurtado Bodell, Måns Magnusson and I add to the literature on interpretable word embeddings, proposing a novel use of informative priors to create predefined interpretable dimensions—thus leveraging the expressiveness and generalizability of the probabilistic framework.

The results show that sensible priors can capture latent semantic concepts better than or on-par with the current state of the art, while retaining the simplicity and generalizability associated with the use of priors. For the study of social interactions, these priors provide a new tool for inferring theoretically meaningful dimensions about individuals and their social environments, and in so doing, enabling richer descriptions of social interaction processes. In turn, such improved descriptions facilitate both better identification of social influence effects and the ability to capture relevant sources of heterogeneity.

23It should be noted that since the writing and publication of this article, a great deal of development has taken place. Today, word embeddings have become substantially more integrated into the CSS toolkit.
Discussion and conclusions

There is considerable excitement about what the computational and digital revolution has to offer sociology, and analytical sociology in particular (Arvidsson et al., 2023; Jarvis et al., 2021; Keuschnigg et al., 2018; Lazer et al., 2009; Salganik, 2019). Thus far, however, the application of tools and data emanating from this revolution for the purpose of mechanism-based explanations of collective outcomes remains scarce. Addressing this gap, this thesis has proposed an analytical framework that combines the explanatory principles of analytical sociology with the strengths of the data and methods of the CSS revolution in order to dissect the dynamics of micro-macro processes (Introduction). This framework has then been used in three different applications (Essays I-III) each of which has sought to elucidate the role of social interdependencies in a specific real-world social process, and which have demonstrated how the framework permits the empirical identification of micro-macro mechanisms. A fourth study, building further on the framework, has proposed a new method that facilitates such applications by improving the description of interaction dynamics in complex high-dimensional data (Essay IV).

Why could this analytical framework not simply have been formulated on the basis of traditional sociological methods and data sources, e.g., surveys, aggregate statistics, and ethnographic case studies? The reason is that in order to translate the conceptual diagram presented in Figure 6—which lays the theoretical foundation for the identification of micro-macro dynamics—into an
analytical framework amenable to empirical analysis, each of its arrows (1–6) need to be empirically identifiable. And in order for this to be case, we need to be able to observe individuals’ behavior in the social environments in which they operate, at a scale, breadth, and temporal granularity that allows us to identify the causal effect of social influence and to measure the direct effect of each action on others as well as on the macro-outcome in question. Such requirements are rarely met by traditional sociological research designs. In contrast, these properties (large N/P/T/I) are characteristic features of the new “digital” datasets that are emerging from the CSS revolution.

In order to meet the requirements of empirical identifiability, Essays I–III used different combinations of large-scale data and computational methods. In terms of data, Essay I relied on large-scale administrative register data, Essay II on digital trace data from Spotify, and Essay III on a mix of administrative register data, digitized historical data, and digital trace data. In terms of methods, Essay I used a node-embedding technique (Grover & Leskovec, 2016) to infer proxies of unobserved drivers of mobility. Essay II used topic models (Blei et al., 2003) to provide a proxy for the unobserved central confounder, musical taste. This use of topic models, in addition to facilitating causal inference, also enabled novel measurements of dyadic similarities between agents and the compatibility of musical artists with individuals’ tastes, which in turn enabled the empirical testing of the proposed mechanism. In Essay IV, a method was developed to improve the applicability of word embedding models (Mikolov, Sutskever, et al., 2013) for social scientific purposes.

More specifically, Essay IV proposed a method that facilitates the inference of theoretically meaningful dimensions from complex high-dimensional data such as text. In so doing, this method helps to advance our ability to describe the social environments in which individuals interact (e.g., online forums and social medias), as well as their behavior in these contexts. As such, these inferred dimensions may be of considerable use when studying the dynamics of social interaction processes. For example, they may be used to construct measures of theoretically meaningful relations between individuals, or, to infer theoretically expected sources of confounding (Sridhar & Getoor, 2019) and thus to improve the identification and estimation of social influence effects.

By empirically anchoring the dynamics of postulated mechanisms, the proposed framework offers a potential way of bridging two central but to date hard-to-reconcile aspects of AS: empirical realism and bottom-up explanations. In the absence of micro-level data on interaction dynamics, AS has tended to opt for psychologically plausible micro-level assumptions that satisfy generative sufficiency (e.g., Hedström and Ylikoski, 2014). However, as has been elaborated
in detail above, the risk with using such an approach to explain concrete social phenomena—and not just to explore mechanisms in the abstract—is that one may fall into the trap of as-if stories, explaining how an outcome might plausibly have been brought about, but not how it actually was. A risk which is amplified by the common reliance on common sense to determine the validity of micro-level assumptions (Watts, 2014). Relatedly, recent discussions have questioned the merit of grounding explanations of collective outcomes in the intentionality of actors and the comprehensibility of their behaviors—in the verstehen sense—given that we can empirically never truly interrogate the internal states of individuals’ minds (Hedström, 2021). The framework proposed here is consistent with this point of view in the sense that its focus is directed at observable dynamics. However, this does not imply that the framework is inconsistent with the alternative perspective. First, as was emphasized earlier, action theory can fulfil an important role in motivating the presence of a given influence-response function. Second, the conceptual side of the framework can also be used in the absence of any empirical data. Doing so has some concrete benefits. By translating one’s action theory into an influence-response function, and by spelling out the implied local dynamics, one derives a clear description about what one expects to observe in empirical settings, which in turn helps to clarify (a) which parts of one’s postulated mechanism are empirically testable, and (b) which other candidate mechanisms it can and cannot be distinguished from using empirical data. Additionally, as argued by Lopez-Pintado and Watts (2008) and Mäs, 2021, this makes it possible to see structural similarities between substantively very different micro-macro processes.

The proposed framework may not only be of relevance for AS, but potentially also for parts of the wider CSS community. As it currently stands, there would seem to be a disconnect between three substantively related strands of CSS research. The first strand estimates local social influence effects but pays little or no attention to their macro implications (e.g., Aral and Nicolaides, 2017; Eckles and Bakshy, 2020; Ternovski and Yasseri, 2020; Easley et al., 2020). The second strand predicts macroscopic patterns—e.g., cascade sizes—but do not consider the details of micro-level interaction dynamics (e.g., Bakshy et al., 2011; Jenders et al., 2013; Cheng et al., 2014; Brady et al., 2017; Vosoughi et al., 2018). The third strand focuses on the link between micro-behaviors and macro-level patterns but uses highly stylized models that are not founded upon detailed, micro-level empirical research (e.g., Barabási and Albert, 1999; Bettencourt, 2013; West, 2017). The analytical framework proposed here provides a possible link between these strands of research by focusing attention on the empirical identification of the social dynamics underlying micro-macro processes.
Several decades ago, Mark Granovetter demonstrated that heterogeneity in social influence can matter greatly for the type of collective outcomes that a group of individuals are likely to bring about (Granovetter, 1978). However, for a variety of reasons—of which a lack of appropriate data, and in its absence, a strong emphasis on simplistic micro-models and generative sufficiency, are likely candidates—this insight has often been overlooked in the construction of bottom-up explanations/models. However, the computational and digital revolution and the emergence of large-scale networked datasets is now enabling the identification of key heterogeneities to an unprecedented degree. Adopting the proposed analytical framework, this thesis thus set out to investigate the implications that homogeneity assumptions have for our understanding of how networks and social influence affect three different social phenomena: workplace segregation, status-quality decoupling, and urban scaling. The general finding, as reported in Essays I-III, is that acknowledging heterogeneity can on the one hand dramatically alter our conclusions about the role of networks, and on the other illuminate important dynamics that operate in the real world.

In the workplace segregation literature, the main tenet holds that networks and network-based recruitment contribute to increased levels of segregation (Bielby & Baron, 1986; Elliott, 2001; Marsden & Gorman, 2001; Reskin, B. F. et al., 1999; Trimble & Kmec, 2011). However, in Essay I, by acknowledging heterogeneity in individuals’ opportunities to form same-category ties across different workplaces, it became clear that opportunity structures often trump homophilic preferences and that this happens, in particular, in the most segregated workplaces (those with the greatest imbalance in e.g., gender representation). This insight led to the formulation and identification of the Trojan-horse mechanism, which, in contradiction with the homophily thesis, shows that segregation is often reduced rather than increased by network recruitment.

In the status-quality decoupling literature, recent studies have presented conflicting evidence and alternative interpretations in relation to earlier work (Manzo & Baldassarri, 2015; van de Rijt, 2019), suggesting that the ability of social influence to produce decoupling may have been overstated. However, across these studies, social environments have in one way or another been assumed to be highly homogenous, limiting external validity. Acknowledging heterogeneity in both individuals’ exposure and susceptibility to novel behaviors, and the important moderating role of homophily, Essay II specified the conditions—partial overlap in preferences and strong social ties between the senders and receivers of social influence—under which social influence can
lead to decoupling. These are conditions which neither of the experimental studies in the status-quality decoupling literature meet. More concretely, our findings suggest that a likely contributing factor behind the self-correcting process observed in van de Rijt (2019) re-analysis of the MusicLab experiments (Salganik et al., 2006) is that the type of social influence in operation was of the "narrow" kind; individuals interacted with others with whom they were not familiar at all, and consequently, exposure to novelty was artificially high, while susceptibility to such exposures was artificially low.

In the urban scaling literature, existing theories implicitly postulate that the way that cities' total interconnectivity, productivity, and innovation change as the cities grow larger is that everyone (and to the same extent) becomes more connected, productive, and innovative—that is, that the whole city distribution is shifted upwards (Bettencourt, 2013). However, in Essay III, by acknowledging heterogeneity in individuals’ interconnectivity, productivity, and innovativeness within cities, we found that between 36% and 80% of the previously reported scaling effects can in fact be attributed to differences in the tails of cities’ distributions. That is, the outliers that Pareto identified a long time ago (Pareto, 1896) are disproportionately located in larger cities, and as a result not only drive the inequality within cities, but also the inequality between cities.

Thus, even though conclusions about the role of networks were (a) derived on the basis of fundamental and well-documented mechanisms in the workplace segregation literature, (b) derived on the basis of strong experimental evidence in the status-quality decoupling literature, and (c) corroborated by impressive aggregate-level predictions in the urban scaling literature, the findings of Essays I-III reveal that these conclusions nevertheless need to be revised to different degrees as a result of implicit or explicit assumptions about overly homogenous individuals and social environments. The fact that these areas of study otherwise have very little in common suggests that other micro-macro processes may also benefit from similar scrutiny in terms of the assumptions made about homogeneity.

Across Essays I–III, three micro-macro mechanisms were identified following the logic and practical steps of the proposed framework. Beyond addressing the specific research questions and research topics considered, these mechanisms all share the property of being semi-general, and may therefore potentially be of broader relevance, beyond the applications presented here.

Concerning the Trojan-horse mechanism (Essay I), the two fundamental building-blocks are (i) opportunity constraints in tie-formation within groups, and (ii) mobility between groups. As such, one could imagine this mechanism as potentially having relevance in any setting in which we encounter groups...
DISCUSSION AND CONCLUSIONS

(e.g., firms, schools, or neighborhoods) that are composed of individuals with different characteristics (e.g., ethnicity, gender, or social class) and where the individuals move between the groups in question. What is likely to determine the relative importance of the Trojan-horse mechanism in any given setting is a combination of (a) the magnitude of the peer effect, (b) the degree of consolidation (Blau & Schwartz, 2018) within groups, e.g., how similar/different individuals of two ethnicities are with respect to other dimensions such as education and age, and (c) the magnitude of homophily. One interesting way that this mechanism could be extended in future research would be by considering multidimensional homophily (Block & Grund, 2014). That is, rather than only focusing on a single dimension, such as gender, one could instead consider group composition in terms of combinations of attributes (e.g., gender $\times$ ethnicity $\times$ education $\times$ age). It may for example be the case that a move which produces segregation on one dimension might reduce segregation on a set of others.

With regard to wide social influence, Essay II developed this mechanism to explain how novel music can spread on Spotify. However, the core idea, that behavioral overlap has an important influence on both susceptibility and exposure to novel behaviors, but in opposite directions—and that it therefore involves a trade-off—is general enough to apply to a wider range of phenomena. One example of an area in which it might be interesting to explore the relevance of this mechanism is the political space, where there are many dividing lines between different political camps (see e.g., DellaPosta et al., 2015). Could it be, for example, that those who hold bridging positions have too little behavioral overlap with one of the sides to facilitate the diffusion of ideas across network clusters? Such questions bring to the fore other questions about how the effect of wide/narrow social influence may be dependent on network structure, which constitutes another interesting avenue to be explored. Tie-formation (creation and dissolution) is in turn affected by homophily, and hence, another set of questions to be explored involve how networks come to exhibit the conditions needed for the operation of wide social influence.

Finally, regarding the city-size dependent cumulative advantage mechanism, Essay III established the presence of this effect for individuals and their wages. However, the postulated reason underlying this effect—differential opportunity ceilings by city size—would seem to also extend to other outputs (e.g., innovation) and other units of analysis (e.g., firms). Supplementary simulations demonstrated that the inclusion of this mechanism in a recently proposed mathematical framework (Bettencourt, 2020) was able to resolve some of the framework’s empirical inconsistencies. Future research should further explore how this and related mechanisms might be incorporated into existing
models of urban scaling. Beyond urban scaling, the proposed mechanism also contributes to a broader research agenda that is interested in the way cumulative advantage effects are conditioned on system-level properties (Lynn & Espy, 2021). Finally, because one aspect of the city-size dependent cumulative advantage mechanism is that it produces differential inequality by city size, this mechanism may also be of potential interest beyond the field of urban scaling, for example to scholars of social inequality.

In conclusion, the analytical framework proposed in this thesis, together with the four essays, demonstrates how AS and CSS can fruitfully be integrated in order to construct empirically anchored micro-macro mappings that go beyond standard conceptions of generative sufficiency. In so doing, the thesis shows how a bridge can be formed between two traditionally hard-to-reconcile aspects of the AS agenda: empirical realism and bottom-up explanations. Reconciling the two has enabled the identification of real-world heterogeneities and their implications for the way micro properties translate into macro-outcomes. In line with Granovetter (1978), this thesis demonstrates that nuances in heterogeneity can alter the dynamics of real-world social processes in fundamental ways. Finally, and on a more general note, the results and conclusions of the thesis suggest that a promising path forward—that is enabled by the integration of AS and CSS—would be to search for what one might call "middle-range dynamics", i.e., generic descriptions of sequences of events in terms of influence-response functions and their macro consequences (e.g., Figure 6). Such generic descriptions do not seek to capture the full history of any given social process, but rather seek to identify the heart of the matter, the most important bits and pieces that are generic enough to be generalized to other cases.


BIBLIOGRAPHY


Dunbar, R. I. (2016). Do online social media cut through the constraints that limit the size of offline social networks? *Royal Society Open Science, 3*(1), 150292.


BIBLIOGRAPHY


Bibliography


Bibliography


Bibliography


BIBLIOGRAPHY


Bibliography


Ylikoski, P. (2016). Thinking with the coleman boat.

Essays
Essays

The essays associated with this thesis have been removed for copyright reasons. For more details about these see:

https://doi.org/10.3384/9789179293314
Beyond Generative Sufficiency

On Interactions, Heterogeneity & Middle-Range Dynamics

Martin Arvidsson