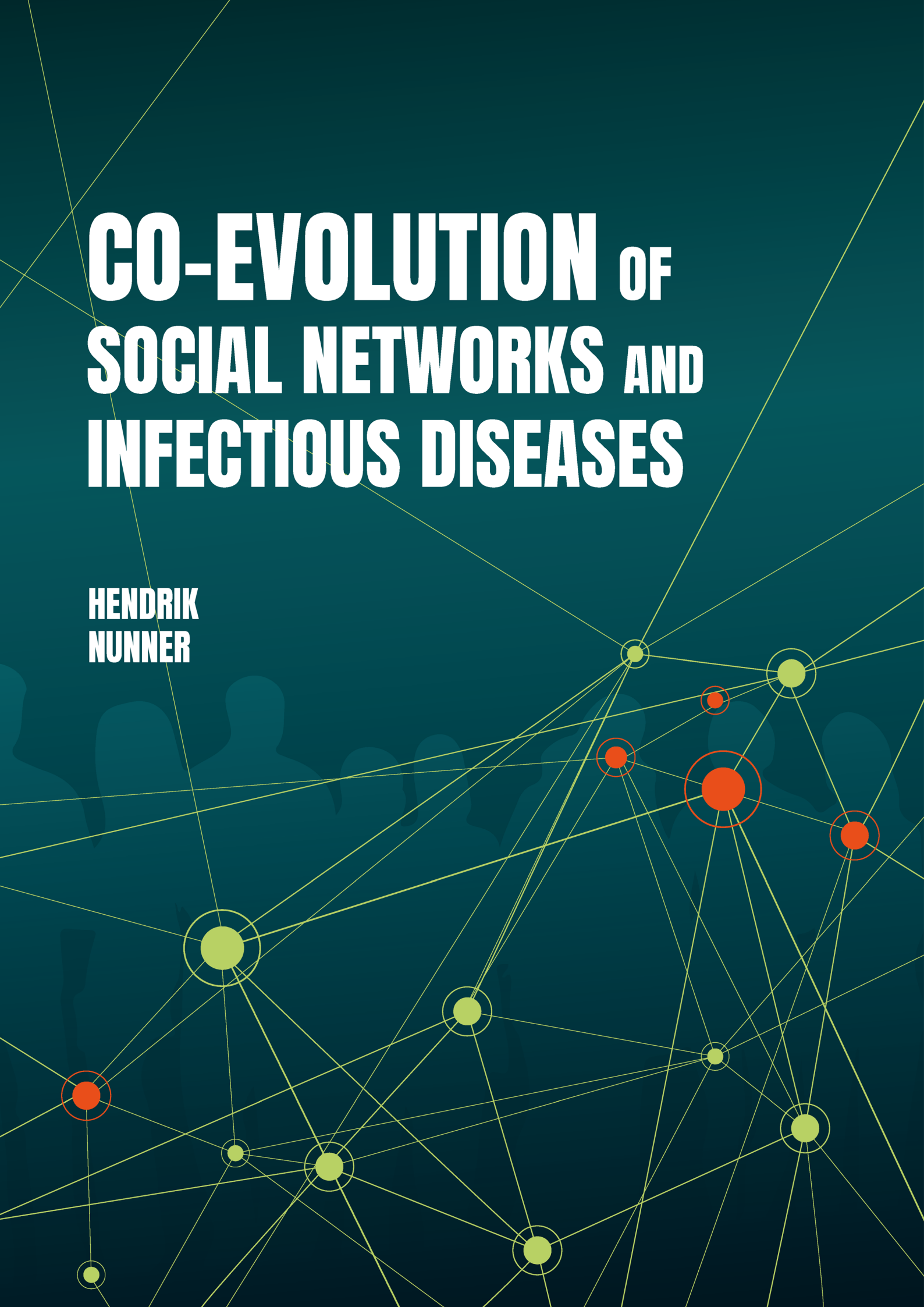


CO-EVOLUTION OF SOCIAL NETWORKS AND INFECTIOUS DISEASES

HENDRIK
NUNNER



CO-EVOLUTION OF SOCIAL NETWORKS AND INFECTIOUS DISEASES

**HENDRIK
NUNNER**

Co-evolution of social networks and infectious diseases
— Hendrik Nunner

Cover design: Benedikt Ploetz Design | benediktploetz-design.com

Printing: Ridderprint BV | ridderprint.nl

ISBN: 978-3-00-074502-7

© Copyright Hendrik Nunner, 2023

All rights reserved. No part of this dissertation may be copied, reproduced, or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage or retrieval system, without the prior written permission of the author. The copyright of the published articles has been transferred to the respective journals.

Co-evolution of social networks and infectious diseases

Co-evolutie van sociale netwerken en infectieziekten
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
vrijdag 14 april 2023 des ochtends te 10.15 uur

door

Hendrik Nunner

geboren op 29 januari 1983
te Henstedt-Ulzburg, Duitsland

Promotoren:

prof. dr. ir. V.W. Buskens
prof. dr. M.E.E. Kretzschmar

Beoordelingscommissie:

prof. dr. M.A.L.M. van Assen
prof. dr. ir. J.A.P. Heesterbeek
prof. dr. P.G.M. van der Heijden
prof. dr. M. Mäs
prof. dr. S.J. de Vlas

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van de Nederlandse organisatie voor gezondheidsonderzoek en zorginnovatie (ZonMw, subsidienummer 91216062) en van ODISSEI, de nationale onderzoeksinfrastructuur voor de Nederlandse sociale wetenschappen (<https://ror.org/03m8v6t10>).

Contents

Contents	v
List of figures	ix
List of tables	xi
1 Synthesis	1
1.1 Introduction	2
1.2 Network behavior in the context of infectious diseases	6
1.2.1 Network formation	6
1.2.2 Disease avoidance	7
1.2.3 An integrated model framework	7
1.3 Summaries of Chapters 2 through 5	8
1.3.1 Chapter 2: A model for the co-evolution of dynamic social networks and infectious disease dynamics	9
1.3.2 Chapter 3: Health behavior homophily can mitigate the spread of infectious diseases in small-world networks	10
1.3.3 Chapter 4: Disease avoidance may come at the cost of social cohesion	11
1.3.4 Chapter 5: Prioritizing high-contact occupations raises effectiveness of vaccination campaigns	12
1.4 Discussion	13
1.4.1 Knowledge gained and implications	13
1.4.2 Limitations and future work	16
1.5 Conclusion	18
2 A model for the co-evolution of dynamic social networks and infectious disease dynamics	21
2.1 Introduction	22
2.2 Theory	24
2.2.1 Infectious diseases	24
2.2.2 Social network formation	25
2.2.3 Health behavior and risk perceptions	26
2.3 The networking during infectious diseases model (NIDM)	27
2.4 A specific model case	27

2.5	Simulation	30
2.5.1	Simulation procedure	31
2.5.2	Parameter settings	32
2.5.2.1	Utilities	32
2.5.2.2	Network	33
2.5.3	Software	34
2.6	Data and analysis	34
2.7	Results and discussion	35
2.7.1	Model behavior	35
2.7.2	CIDM parameters	37
2.7.3	Network properties	39
2.7.4	Interaction effects	42
2.8	Conclusion and implications	44
3	Health behavior homophily can mitigate the spread of infectious diseases in small-world networks	49
3.1	Introduction	50
3.1.1	Network models of disease spread	50
3.1.2	Small-world networks	51
3.1.3	Health behavior homophily	52
3.2	The model	53
3.2.1	Purpose	53
3.2.2	Basic principles	53
3.2.3	Dynamics of disease transmission and network formation	54
3.2.4	Utility	55
3.3	Simulation	56
3.3.1	Simulation procedure and output data	56
3.3.2	Parameter settings	58
3.3.3	Data and source code availability	58
3.4	Analysis	59
3.5	Results and discussion	60
3.5.1	Epidemics in adaptive vs. static networks	60
3.5.2	Effects of health behavior homophily on epidemics in adaptive small-world networks	60
3.5.3	Effects of health behavior homophily on epidemics in static small-world networks	63
3.6	Conclusion and implications	63
4	Disease avoidance may come at the cost of social cohesion: Insights from a large-scale social networking experiment	69
4.1	Introduction	70
4.1.1	The experiment	71
4.1.1.1	Available actions	71

4.1.1.2	Point rewards	71
4.1.1.3	Disease transmission	75
4.1.2	Hypotheses and conditions	75
4.2	Results	77
4.2.1	The role of disease avoidance in decision-making	78
4.2.2	Group- and individual-level effects	80
4.2.3	Determinants of decision-making beyond avoidance behavior	83
4.3	Discussion	85
4.4	Methods	88
4.4.1	Experiment	88
4.4.1.1	Participants	88
4.4.1.2	Recruitment and compensation	88
4.4.1.3	Design	89
4.4.1.4	Procedure	90
4.4.2	Data and analysis	91
4.4.3	Availability of data and materials	92
5	Prioritizing high-contact occupations raises effectiveness of vaccination campaigns	95
5.1	Introduction	96
5.1.1	Related work	97
5.2	Simulation model	98
5.3	Results	102
5.3.1	Comparison of vaccination campaigns	102
5.3.2	Dynamics of epidemics	105
5.4	Discussion	106
5.5	Methods	109
5.5.1	Network formation model and simulation	109
5.5.2	Calibration of network structure with empirical data	110
5.5.3	Technical setup and runtimes	111
5.5.4	Code availability	111
5.5.5	Data availability	111
A	Supplement to Chapter 2	113
B	Supplements to Chapter 3	121
B.1	Additional model definitions	122
B.1.1	Utility	122
B.1.2	Process overview and scheduling	123
B.1.3	Entities, state variables and scales	124
B.1.4	Emergence	126
B.1.5	Stochasticity	126
B.2	Additional analyses	126

B.2.1	Network and disease properties	126
B.2.2	Effect of degree on final size	127
B.2.3	Additional effects on epidemics	128
B.2.4	Regression models	132
C	Supplements to Chapter 4	137
C.1	Communication with participants	138
C.1.1	Registration	138
C.1.2	Registration confirmation	141
C.1.3	Instructions	142
C.1.4	Pop-up Game A	146
C.1.5	Pop-up Game B	147
C.1.6	Survey	149
C.1.7	Cannot be assigned	150
C.2	Additional definitions	151
C.2.1	Simulations to determine parameter settings for the experiment . .	151
C.2.2	Additional simulations	153
C.3	Additional results	154
C.3.1	Demographic data of the participants	154
C.3.2	Effect of settings and conditions on epidemics	156
C.3.3	Avoidance of infected alters	156
C.3.4	Factors contributing to networking decisions	156
C.3.5	Progression of properties over time (extended)	159
C.3.6	Effect of parameter variations on network structure and epidemic size	160
D	Supplements to Chapter 5	163
D.1	Supplementary results	164
D.2	Supplementary methods	175
D.2.1	Network generation	175
D.2.2	Disease spread	177
D.2.3	Parameters and submodels	177
	Nederlandse samenvatting	181
	References	191
	Acknowledgments	209
	About the author	215
	ICS dissertation series	219

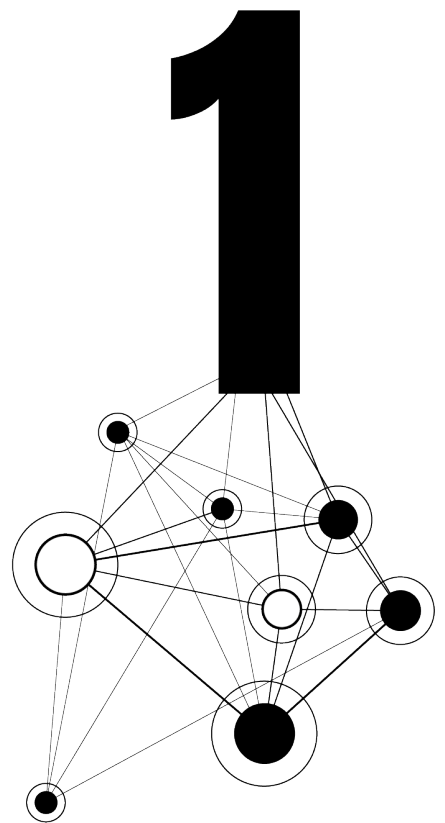
List of figures

1.1	Structure of this dissertation	10
2.1	Disease dynamics in networks with and without network changes	36
2.2	Disease and network dynamics by network size	41
2.3	Interaction of costs for infected ties and value of indirect connections	43
2.4	Interaction of disease severity and risk perception	44
3.1	Comparison of epidemics in adaptive and static networks	61
3.2	Bivariate effects of network properties on epidemics	61
3.3	Interaction effects in adaptive networks	63
4.1	Flowchart of the experiment	72
4.2	Graphical user interface of the Networking during Infectious Diseases Task	73
4.3	Marginal effects on the probability of a relation being desirable	79
4.4	Final epidemic size by clustering and social mixing	82
4.5	Progression of properties over time	84
5.1	Vaccination campaign scenarios	100
5.2	Flowchart of the simulation	101
5.3	Densities of final and peak size of epidemics	102
5.4	SIRV plots for vaccine availability per vaccination campaign condition . . .	105
A.1	Disease dynamics by network size in networks with network changes	115
A.2	Disease dynamics by network size in networks without network changes . .	116
B.1	Effects of model parameters on network properties	128
B.2	Effect of network properties on epidemics in adaptive networks	130
B.3	Effect of network properties on epidemics in static networks	131
C.1	Illustration of clustering settings	145
C.2	Illustration of clustering settings (Pop-up Game A)	146
C.3	Illustration of clustering settings (Pop-up Game B)	147
C.4	Effects of settings and conditions on epidemics	157
C.5	Avoidance of infected alters	157
C.6	Progression of properties over time (extended)	159
C.7	Parameter variations and their effects in simulations	160

List of tables

2.1	Overview of CIDM parameters	32
2.2	Two-level random-intercept logistic regression of final size of epidemics . .	38
2.3	Two-level random-intercept linear regression of duration of epidemics . . .	39
2.4	Properties of epidemics by network size	40
3.1	State variables, scales, and settings for model parameters	57
3.2	Regression analysis summary	62
4.1	Properties and corresponding point rewards per round	75
4.2	Proportions of decisions made and decisions increasing rewards	78
4.3	Factors contributing to the attractiveness of a relation	81
4.4	Risk factors for acquiring infection during one installment of the NIDT . .	82
5.1	Mean degrees per major occupational group (empirical networks)	99
5.2	Mean final size of baseline condition and difference by test condition	104
A.1	Descriptive statistics of network dynamics, epidemics, and network measures	117
A.2	Comparison of two-level random-intercept logistic regression models for final size of the epidemic by network size	118
A.3	Comparison of two-level random-intercept linear regression models for duration of epidemics by network size	119
B.1	State variables, ranges, and initial settings	125
B.2	Descriptive statistics	127
B.3	Logistic regression models for the effect of average degree on final size in adaptive networks	129
B.4	Logistic regression models for the effect of average degree on final size in static networks	129
B.5	Logistic regression models of final size with network dynamics	132
B.6	Linear regression models of epidemic duration with network dynamics . . .	133
B.7	Linear regression models of epidemic peak size with network dynamics . .	133
B.8	Linear regression models of number of network changes	134
B.9	Logistic regression models of final size without network dynamics	134
B.10	Linear regression models of epidemic duration without network dynamics .	135
B.11	Linear regression models of epidemic peak size without network dynamics .	135

C.1	Chance of infections	143
C.2	Points per number of relations	144
C.3	Chance of infections – Game A	146
C.4	Points per number of relations – Game A	146
C.5	Chance of infections – Game B	147
C.6	Points per number of relations – Game B	147
C.7	Parameter ranges used for network generation and determination of game parameters, and settings used for game parameters	152
C.8	Demographic data of the participants	154
C.9	Countries of residence of the participants	155
C.10	Three-level random-intercept logistic regression of networking decisions . .	158
D.1	Descriptive statistics of input networks for the simulations	165
D.2	Comparison of occupational groups in empirical and generated networks (1)	166
D.3	Comparison of occupational groups in empirical and generated networks (2)	167
D.4	Comparison of occupational groups in empirical and generated networks (3)	168
D.5	Comparison of occupational groups in empirical and generated networks (4)	169
D.6	Final sizes of the simulated epidemics by conditions and controls	170
D.7	Peak sizes of the simulated epidemics by conditions and controls	171
D.8	Mean peak size of baseline condition and difference by test condition . . .	172
D.9	Duration of the simulated epidemics by conditions and controls in time steps	173
D.10	Mean duration of baseline condition and difference by test condition	174
D.11	Parameters, ranges, and initial settings	178



Chapter 1

Synthesis

1.1 Introduction

Since the beginning of the COVID-19 pandemic in early 2020 until the time of finishing this dissertation in July 2022, the World Health Organization has registered almost 560 Million cases and well over 6 Million deaths from COVID-19 (World Health Organization, 2022b). Although a number of effective vaccines have been developed and large-scale vaccination campaigns have been implemented globally, only 58 of 194 WHO member states have met the 70% vaccine coverage goal (World Health Organization, 2022a). Understanding these disease dynamics in their social contexts remains crucial to mitigate the negative medical and social consequences of such epidemics. Mathematical models of disease spread are one of the most important tools to understand the spread of such dynamics and can thus support public health policymaking.

Classical deterministic models in epidemiology, for example, are powerful tools to predict the prevalence of the disease in large communities over time. This is important, *inter alia*, to estimate hospitalization rates and how many vaccinations are required to control the case numbers for an upcoming COVID-19 wave. These models typically use ordinary differential equations that assume the population of potential hosts to be perfectly mixed. We know, however, that people do not mix homogeneously but form social ties based on personal preferences and the opportunities provided by their environments (Feld, 1981; McPherson et al., 2001). Network models of disease spread, the type of model this dissertation addresses, relax the assumption of perfectly mixed populations and enable us to understand the effects of social structure and interaction patterns on the spread of infectious diseases. This is especially important for diseases that spread mainly via interpersonal contacts, such as COVID-19 (Li et al., 2020; Shen et al., 2020). Earlier simulation studies, for example, have repeatedly suggested that the presence of clusters within a social network can mitigate disease spread (Badham & Stocker, 2010; Keeling, 1999; Miller, 2009). This is because clusters describe densely connected areas within a network in which only a few relations form bridges to other areas. For an infectious disease, these bridges therefore represent a bottleneck for the spread from one cluster to another.

Still, just like any scientific model, network models make simplifying assumptions that hamper the interpretation of their theoretical results in the real world. A typical assumption is that human behavior (contact rates, health behavior) is constant over time resulting in static network structures (Bansal et al., 2010). Neglecting network changes may be a good approximation, especially when such changes occur on a much larger timescale than the spread of infections (Pastor-Satorras et al., 2015). Predictions about disease spread dynamics, however, could be substantially different if the co-evolutionary process between social networks and the spread of infection is taken into account (Bansal et al., 2010; Marceau et al., 2010). An approach that considers such processes is the design of individual-based models for infectious disease transmission (Bedson et al., 2021; Verelst et al., 2016; Willem et al., 2017). Simulations that are based on random network changes on

the host level, for example, suggest that considering network dynamics can lead to smaller and shorter epidemics (Gross et al., 2006; Leung et al., 2018; Risau-Gusman & Zanette, 2009).

We consider network changes not as random changes, but as conscious decisions related to people’s behavioral responses to and concerns about the disease. To illustrate how health behavioral responses can alter the structure of our social networks systematically, imagine it is early 2020 and the first cases of COVID-19 have been reported in your country. You are afraid to catch a disease that is not well understood yet and unsettled by the pondering question of how to protect yourself best. Even before official spatial distancing measures are announced, you decide to cancel a skiing trip to South Tyrol and to avoid social contacts that are not absolutely necessary. A few months later, after many people have died from COVID-19 infections, the first wave of the pandemic has subsided due to a nationwide lockdown. It is at that time that you and a friend plan a visit because neither of you experience any symptoms related to a COVID-19 infection and no one in your households has seen anyone else in a while.

This example describes a network mechanism that forms the basis for spatial distancing measures. That is, a change in network structure (here: reduction of social contacts) affects disease spread (here: lowering the number of new cases). The example shows further, that there are various reasons for structural change. First, people may die and all their social relations disappear permanently. Second, authorities may impose temporary measures for spatial distancing. Third, individuals may decide on their own accord to temporarily reduce their interpersonal contacts. In this dissertation, we are interested in network changes based on this latter reason, which represents a feedback loop between (i.) the spread of infectious diseases in social networks, (ii.) the adaptation of people’s networking behavior to disease spread, and (iii.) the influence of structural network changes on the spread of infections. We therefore ask *How does health behavior shape the co-evolution of social networks and infectious diseases?* (**Q1**).

To answer **Q1**, we introduce the *Networking during Infectious Diseases Model* (NIDM) in the following chapter. The NIDM is an individual-based mathematical model framework and agent-based simulation that integrates theory from sociology, health psychology, and epidemiology. Specifically, the NIDM describes networking decisions as a trade-off between the benefits, costs, and potential harms of infections that a social relation creates. That is, the NIDM rewards beneficial network positions (e.g., the number of social relations, being part of a cluster) and inflicts costs for acquiring an infection. The degree to which agents avoid infectious others depends on individual risk perceptions regarding personal susceptibility to, and the severity of an infectious disease. Using simulations in which artificial agents seek to myopically maximize expected rewards, we create predictions of how health behavioral reactions and network properties affect disease spread and vice versa (Chapters 2 and 3).

Another shortcoming of many network models of disease spread is that their results are not tested empirically and thus remain theoretical. To overcome this shortcoming regarding our own theoretical results presented in Chapters 2 and 3, we ask *How do the theoretical predictions about the co-evolution of social networks and infectious diseases stand up to empirical testing?* (**Q2**).

To answer **Q2**, we created the *Networking during Infectious Diseases Task (NIDT)*. The NIDT is the main task of a large-scale incentivized social networking experiment resembling the specific NIDM model case designed in Chapter 3. The added value of empirically testing our theoretical predictions is that agent-based simulations and incentivized behavioral experiments are complementary methods that in unison provide better insights than each method by itself (Buskens et al., 2022). More specifically, we can test the behavioral assumptions on which our model and agent-based simulations are based (i.e., risk perceptions regarding susceptibility and disease severity), and the resulting macro-level phenomena (i.e., extent of disease spread). Conducting an experiment is especially important for three reasons. First, our simulations have shown that even small behavioral changes can be decisive for whether a disease outbreak disappears quickly or spreads through the entire network. Second, human decision-making is not entirely based on cost-benefit considerations, but subject to cognitive limits and social preferences (Camerer, 2003). Using an experiment can therefore shed light on psychological mechanisms that are not included in the theoretical model (Lunn & Ní Choisdealbha, 2018). Third, we believe that hypothetical networking decisions reproduced in controlled experiments approximate realistic networking decisions better than simple rules of network formation models. Findings of this experiment can then be used to adjust parameter settings and extend model components.

In the third and final part of this dissertation, we change the perspective from individual level behavior to network level properties and how they can be used for infectious disease control. Here, we ask *How can a better understanding of the interdependencies between social networks and infectious diseases enable us to design effective mitigation strategies?* (**Q3**). Specifically, the final part of this dissertation addresses the problem that only a limited number of people can receive vaccinations when supplies are scarce, a situation that occurred at the beginning of COVID-19 vaccination roll-out. Earlier simulation studies suggest that prioritizing high-contact individuals can be effective in containing disease spread (Dezső & Barabási, 2002; Pastor-Satorras & Vespignani, 2002). The network mechanism behind this phenomenon is that the immunization of high-contact individuals reduces the number of so-called *super-spreaders*, that is a few individuals who are responsible for large numbers of transmissions. A remaining problem, however, is how to identify high-contact individuals.

To answer **Q3**, we used NIDM simulations to compare the effectiveness of vaccination campaigns. That is, we first generated social networks based on empirical contact data with detailed occupational codes. Thereafter, we compared the effectiveness of vaccination campaigns with limited supplies that distribute vaccines either randomly or prioritize high-

contact occupations. Consequently, we test whether occupational groups can be a suitable and readily available proxy for vaccination campaigns rather than targeting single high-contact individuals.

The research done for this dissertation is one of three parts of the so-called *SocNetID* project (*Harnessing Social Networks for Infectious Disease Control*).¹ While the part presented here studied the feedback loop between disease spread and the adaptation of network behaviors to disease spread, another part of the project studied the feedback loop between disease spread and social influence on the adoption of health-promoting behaviors (Teslya et al., 2022a; Teslya et al., 2022b). This is important because a typical assumption of models of disease spread is that they consider health behavioral responses to be homogeneous and constant. From the literature, however, we know that health behavior can change over time. Again, imagine it is early 2020, and we are in the early stages of the COVID-19 pandemic. Although, you reduced your social contacts, you are still looking for measures how to protect yourself best from an infection. Earlier than most, you start wearing face masks in public because a friend who works as a physician in the nearby hospital reassured you that it is a simple yet effective tool to lower risks of infection. A few months later, you decide to stop wearing your mask in places where it is not mandatory. That is not just because the number of new cases has gone down, but rather because most people you know and see have gradually done so and look at you disparagingly when they see you still wearing a mask. This example describes two network mechanisms of social influence.² First, social contacts may convey or reinforce information (here: effectivity of mask-wearing) that affects behavior (here: starting to wear masks) (Centola & Macy, 2007). Second, social norms may change over time (here: mask-wearing in places where it is not mandatory) and non-conformity may be punished with subtle yet effective signals (here: giving a disparaging look). The third part of the SocNetID project used *Respondent-Driven Sampling* (RDS; Heckathorn, 1997) to measure networks and clustering in networks. In short, RDS combines snowball sampling (participants recruit other participants from acquaintances) and mathematical models to overcome the non-random way of data collection.

Since it provides the theoretical framework for this dissertation, the following section first introduces the basic theoretical concepts of NIDM. Thereafter, the results of the individual studies (Chapters 2 - 5) are summarized. The synthesis then concludes with an overarching discussion of what we learned, the implications of our findings, and opportunities for future work.

¹To provide temporal and societal context, it should be noted that the work on the SocNetID project began in September 2017, well before the COVID-19 pandemic. At that time, we embarked on a new path (especially in the Netherlands) to establish a structural collaboration between infectious disease epidemiology and social sciences.

²The example is for illustration purposes only, and is not intended to be complete regarding social influence regarding health-promoting behaviors.

1.2 Network behavior in the context of infectious diseases

1.2.1 Network formation

A fundamental question in the psychological and social sciences is what drives human behavior. Depending on the context, our motivations to act may be grounded in biological (e.g., hunger, reproduction), personal (e.g., career goals), or social needs (Morsella et al., 2009). A satisfying and fulfilling social life, for example, can contribute to social needs. That is, a family member, friend, colleague, neighbor, or acquaintance may provide affection, a sense of belonging, or emotional support in times of need. Dysfunctional social networks, in contrast, can have negative effects on someone's health. Early studies on social support showed that social isolation is associated with higher mortality (Berkman, 1984; Berkman & Leonard Syme, 1979; House et al., 1982) and poorer mental health (Goldberg et al., 1985; Lin et al., 1999). In addition, it is not necessarily the number of relations that matters, but also the quality of relations. Research on relapse prevention for alcohol abuse, for example, showed that artificially created support was ineffective, while support from the pre-existing social network increased the success of alcohol treatment (Barber & Crisp, 1995).

These examples suggest that it is in our best interest to actively surround ourselves with people who have positive effects on our lives. A common theme in the analytical social sciences is how the benefits of social relations can be used and optimized for personal advantage. According to Coleman (1994), closeness in social networks facilitates access to social capital, resulting in densely connected, cohesive structures. Proximity to the source of information, he argues, not only guarantees easier access, but also minimizes the risk of information being distorted by transmission through intermediaries. Furthermore, closed structures enable more effective sanctioning of violations, thus promoting trust and compliance with norms. Burt (1992), on the other hand, discusses the benefits of openness in interpersonal relationships. That is, if an actor's relationships are not shared among themselves, the actor can control who gets access to resources and information.

Consequently, both Coleman (1994) and Burt (1992) argue that adding social relations strategically can be beneficial. However, there are limits to the number of social relations a person pursues. Dunbar (1993), for example, argues that humans have a cognitive limit that allows them to maintain about 150 social relationships. Despite criticism about the accuracy of that number, there is support that there is an individual maximum number of relationships a single person can maintain (Lindenfors et al., 2021). From an economical/sociological perspective, Jackson (2008) argues that despite the benefits of social relations, they also come at a cost. That is, a social relation that is not maintained through the investment of time and effort will eventually dissolve.

1.2.2 Disease avoidance

Despite the positive effects of social relations on our lives, people in our social networks who are infected with a communicable disease can pose a health risk that triggers behavioral responses. Consider, for example, a highly motivated and popular colleague who shows up at work despite having symptoms during flu season. Since people tend to avoid others in times of increased risk of infection, the sick colleague has to eat lunch alone. Common examples of avoidance behaviors include avoiding public places (Jones & Salathé, 2009) and workplaces (Ahmed et al., 2018), voluntary quarantining (Tracy et al., 2009), or avoiding certain social contacts (Funk et al., 2010). Furthermore, avoidance behaviors are among the most commonly adapted behaviors to prevent infections for airborne diseases (Bish & Michie, 2010).

According to a large body of empirical studies (Bish & Michie, 2010; Ferguson, 2007; Goodwin et al., 2009; Jones & Salathé, 2009; Leppin & Aro, 2009), subjective risk perceptions on two dimensions are the main reasons for whether a person chooses to engage in avoidance behavior: (i.) perceived susceptibility to acquiring a disease and (ii.) perceived severity of a disease. In addition, the number and severity of symptoms determine how long and how much social contact is avoided (Bish & Michie, 2010). Consequently, individuals who consider themselves highly susceptible and expect severe symptoms in the event of an infection are more likely to avoid social contact than individuals who perceive a low risk of getting infected and expect to have mild symptoms.

1.2.3 An integrated model framework

A theoretical framework that allows to integrate these aforementioned considerations is *social production function* (SPF) theory (Ormel et al., 1999). That is, according to SPF theory, the optimization of well-being is the driving force behind human decision-making. Well-being, in turn, is organized hierarchically. On top, there are two universal, non-substitutable goals: *physical well-being* and *social well-being*. Both goals, however, need to be pursued independently, as lower levels of one type of well-being cannot be compensated with higher levels of the other type of well-being. Finally, in their pursuit of well-being, SPF theory assumes humans to act boundedly rational in order to maximize the net benefits for these goals. That is, people seek to minimize costs while maximizing benefits, given the constraints they face and the resources and information available to them.

Although SPF theory provides a theoretical framework to describe network behavior in the context of infectious diseases at an intuitive level, it lacks precision and a formal notation to make predictions about individual human behavior and emergent macro-level phenomena that are precise enough to be empirically tested. For this purpose, we created the *Networking during Infectious Diseases Model* (NIDM, presented in Chapter 2), which defines two interdependent processes: social network formation and disease transmission. In the NIDM, an infectious disease can be transmitted between two individuals that share a relation in a social network. The probability for a susceptible individual to get infected depends on the number of infected social relations and the infectivity of the disease. Once

infected, an individual recovers, and thus becomes immune to the disease, after a fixed number of discrete time steps. In Chapter 5, we introduce *vaccinated* as a fourth disease state. Vaccinated individuals are immune to the disease without having to recover from infection.

The social network formation process of the NIDM defines the behavioral mechanisms and assumptions of how people form social networks under consideration of infectious diseases. To be more precise, the NIDM is an individual-based mathematical model (for an overview, see Verelst et al., 2016) that defines network behavior in the context of infectious diseases as maximizing the net benefits (or utility) of one’s social relations. Net benefits are composed of three components. First, there are immediate benefits that a social relation creates. These benefits could take the form of affection, social capital, or sense of belonging. Second, there are immediate costs that a social relation creates. These costs could take the form of time and effort invested to maintain that relation. Third, there are (potential) costs a social relation creates as disease carrier. These (potential) costs of infections depend on the disease states of the agents in a dyadic relation. To illustrate this, imagine two agents: an ego that makes a networking decision (create, maintain, or break a relation), and an alter that is the subject of that decision. If the alter is not infected, the networking decision is solely based on whether the immediate benefits outweigh the immediate costs. If, however, the alter is infected, the potential costs of an infection for a susceptible ego depend on how infectious and severe the ego perceives the disease to be. It follows, that a susceptible ego who perceives high risks of an infection is more likely to avoid the infected alter than a susceptible ego who perceives low risk of an infection.

The NIDM is implemented as an agent-based simulation. In this simulation, an arbitrary number of agents make networking decisions in order to myopically maximize individual utility. The creation of a relationship, however, requires the consent of both agents involved, while the dissolution of a relation can be made unilaterally. The agent-based simulation therefore allows predictions about the outcomes generated by the feedback loop between (i.) how disease spread influences networking decisions, (ii.) how these decisions may (temporarily) alter social relations as the main pathway for disease transmission, and (iii.) how the structural changes in the social network affect disease spread.

1.3 Summaries of Chapters 2 through 5

The NIDM constitutes the microtheory that enabled us to study the co-evolution of social networks and infectious diseases. To obtain a comprehensive picture, we combined micro- and macro-level perspectives as well as exploratory and explanatory methods (see Figure 1.1). In Chapters 2 and 3, we introduce the theoretical foundations of the NIDM and propose two specific model cases. Simulations allowed us to make predictions about how variations in risk perception (on the individual and on the group level) and structural properties of the social networks affect disease spread in adaptive social networks. To test the behavioral assumptions of the NIDM and the hypotheses derived from the predictions, Chapter 4 describes the results of a large-scale online network experiment. At its core,

the experiment is a replication of the specific NIDM model introduced in Chapter 3. In the experiment, artificial agents were replaced by human participants. Finally, Chapter 5 describes a simulation study that uses the NIDM to fit empirical contact data to contact networks, and studies whether prioritizing high-contact professions can improve the distribution of vaccinations when supplies are scarce. A summary of the main findings and how earlier chapters are related to later chapters can be found in Figure 1.1.

It remains to be mentioned that the chapters were written as individual manuscripts and can be read as such. That is, although the chapters build on each other, each chapter is self-contained and discusses relevant information from an earlier chapter if needed. In consequence, some redundancy could not be avoided due to this set-up. Furthermore, chapters follow the structure required by the journals to which they were submitted. To accommodate the naming conventions of the journal’s target audience, terminology may vary slightly between chapters (e.g., dynamic/adaptive networks).

1.3.1 Chapter 2: A model for the co-evolution of dynamic social networks and infectious disease dynamics

The main contribution of Chapter 2 is the introduction of the NIDM and the presentation of a simple yet relevant use case. In a first step, we identified relevant theories from sociology, health psychology, and epidemiology and integrated them into the NIDM. As previously discussed, the NIDM defines a utility function for network behavior in the context of infectious diseases consisting of three components: the benefits, costs, and potential harm social relations create.

In a second step, we introduced the *Connections during Infectious Disease Model* (CIDM). The CIDM is based on the *Truncated Connections Model* (Jackson & Wolinsky, 1996) and describes social scenarios in which the benefits of social relations extend to the relations of someone’s immediate relations (*indirect relations*, such as friends of friends). Costs, however, are only incurred by the immediate relations. The decision to choose a well-understood network formation model was guided by the hope of facilitating interpretation of our results.

In a third step, we used agent-based simulations to study the effect of disease avoidance on epidemics in the CIDM. To acquire a broad picture of the interdependent dynamics of social networks and infectious diseases, we followed an exploratory approach by systematically varying at least one parameter from each NIDM component. The results based on 36,000 simulations suggest that higher levels of perceived health risks elicit stronger disease avoidance responses and thus lower numbers of social relations, which, in turn, result in fewer agents acquiring infections. Higher benefits of indirect relations, however, facilitate the creation of social relations, which in turn, result in more potential transmission routes, higher peaks, and quicker disease spread. Additionally, we found that small changes in disease avoidance can make the difference whether a disease outbreak becomes an epidemic. Finally, we found that the reward structure of the CIDM causes the average number of social relations to increase with network size. As a result, networks with 50+ agents are so densely connected that outbreaks almost always infect the entire network. Consequently,

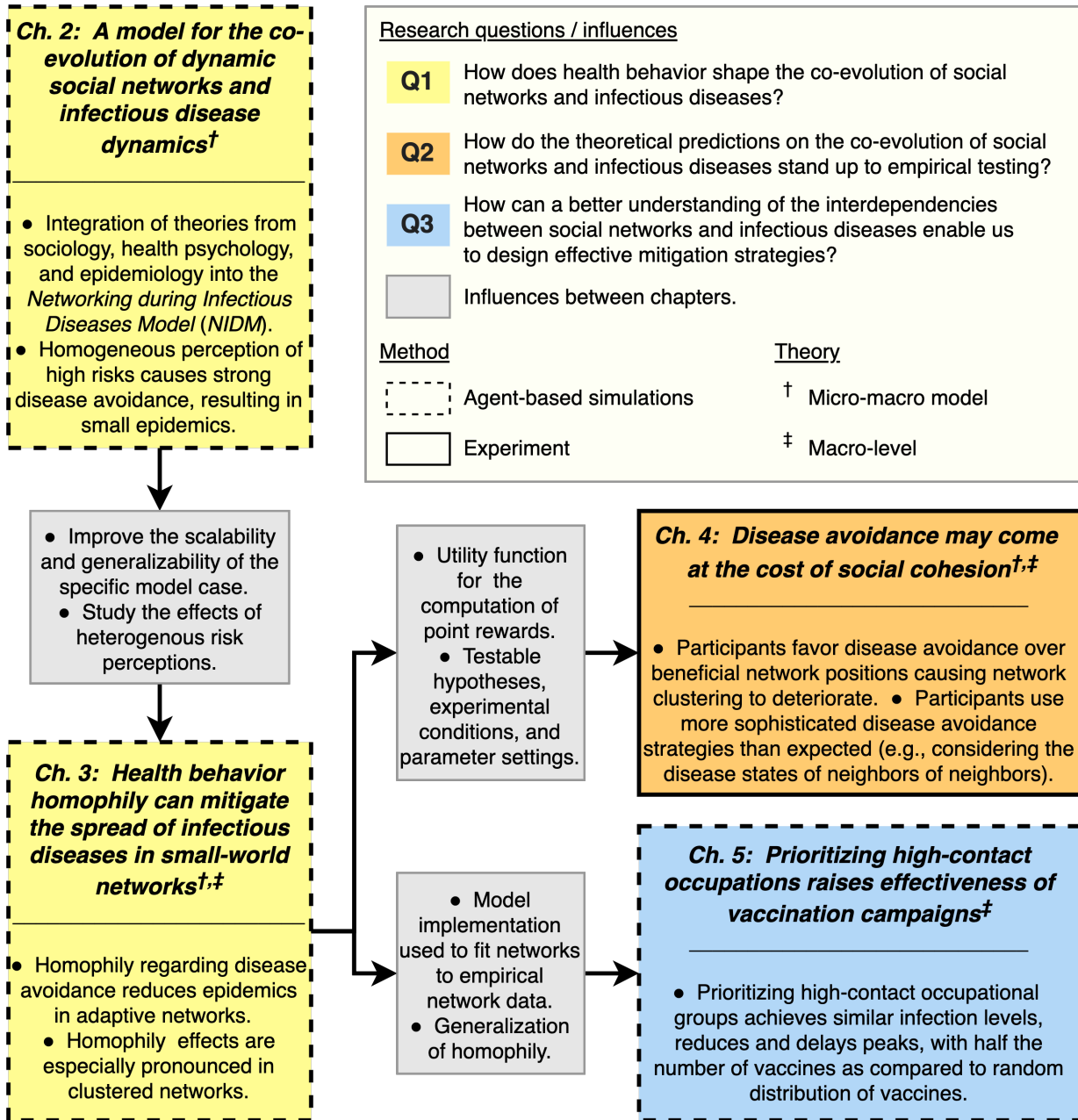


Figure 1.1: Structure of this dissertation.

the insights from this chapter are interesting for social contexts of limited size, such as classrooms, small companies, extended family, or groups of friends. More importantly, the NIDM has proven to be a powerful framework to produce plausible insights into the co-evolution of social networks and infectious disease dynamics.

1.3.2 Chapter 3: Health behavior homophily can mitigate the spread of infectious diseases in small-world networks

Addressing the limitations of the CIDM, Chapter 3 introduces a new specific NIDM model case, the so-called *Small-Worlds during Infectious Disease Model (SWIDM)*. To improve scalability, we defined a utility component that allows to control the optimal number of

relations independent of network size. To improve generalizability, we chose to model the formation of small-world networks, a type of network that shares properties commonly found in real-world contact networks: high clustering (densely connected areas within a network) and short average path length (a few bridging relations between clusters that provide shortcuts from one part to another part of the network). To be more precise, the utility function of the SWIDM allows us to control the rewards for the proportion of closed triads on the individual-level and thus network clustering and average path length on the group-level. Additionally, the SWIDM allows controlling the likelihood of relations between agents similar in risk perception (*health behavior homophily*).

Our results, based on more than 80,000 agent-based simulations, suggest that both network clustering and health behavior homophily can mitigate the spread of infectious diseases. On the one hand, the existence of bridges between clusters constitute a bottleneck for the disease to spread from one cluster to another cluster. This bottleneck, in turn, can quickly be dissolved by only a few agents and thus requires only minimal structural change to isolate an entire cluster from the disease. On the other hand, high degrees of health behavior homophily cause groups of agents perceiving high risks to protect each other from infection. That is, a single susceptible agent can lower their individual probability to acquire an infection by breaking a relation to an infectious neighbor. If, however, the agent's neighbors are also perceiving high risks of infections, the homogeneous preference to quickly dissolve relations to infectious neighbors lowers the probability to acquire an infection for the entire neighborhood. Finally, we observe that the combination of the two effects (networks with clusters of health homophilous agents) produce the smallest outbreaks. That is because entire clusters of agents perceiving high risks of infections are not only more likely to isolate themselves, but can also isolate themselves quickly enough to prevent the disease from spreading into the cluster (only a few bridges with at least one agent that is likely to dissolve infectious relations).

1.3.3 Chapter 4: Disease avoidance may come at the cost of social cohesion

In Chapter 4, we present the experimental study to test the behavioral assumptions of the NIDM and hypotheses derived from SWIDM simulations. Specifically, we tested whether infectious alters are avoided by susceptible egos (**H4.1**), whether higher perceived risk of infection causes stronger avoidance of infectious alters (**H4.2**), whether higher perceived risk of infection lowers the individual probability to get infected (**H4.3**), whether epidemics are smaller in networks with a higher degree of clustering (**H4.4**), and whether epidemics are smaller in networks with a higher degree of health behavior homophily (**H4.5**). For this purpose, we developed the *Networking during Infectious Diseases Task* (NIDT), a round-based network game that rewards points based on the utility function of the NIDM. That is, participants received points for the number of social relations and the proportion of closed triads, and lost points for acquiring infection. Points were converted into monetary rewards and paid to the participants at the end of the experiment.

Our results support the behavioral assumptions of the NIDM framework. That is, participants actively avoid infected alters (**H4.1**) and more risk averse participants show stronger avoidance reactions (**H4.2**). Avoidance behavior, however, was so strong that the disease hardly spread within the networks. As a result, we could neither confirm nor reject the remaining hypotheses (**H4.3-H4.5**). Additional simulations suggested that stronger disease avoidance alone cannot explain the experimental results. Participants rather used more sophisticated strategies to avoid infection and infectious others than expected. That is, infectious participants were not only avoided by susceptible participants but by all participants irrespective of disease state. Furthermore, not only infectious participants were avoided but also susceptible participants with infectious neighbors. Finally, we find that connectedness of neighbors played only a minor role for networking decisions so that clustering deteriorated in both clustering conditions (increasing in the low clustering condition, decreasing in the high clustering condition). To summarize, we find that participants prioritize disease avoidance over maintaining network structure at the cost of potentially higher point rewards.

We identify three reasons for the discrepancy between the model predictions and the experimental results. First, infections provide a salient signal (red color indicating infected participants) that triggers immediate behavioral responses (avoidance). In comparison, clustering requires more effort to keep track off (monitoring the number of closed triads someone is part off). Second, according to *Prospect Theory* (Kahneman & Tversky, 2013; Tversky & Kahneman, 1992) people tend to reject available benefits to avoid low probabilities of loss (Denes-Raj & Epstein, 1994). This tendency can have induced the strong disease avoidance reaction at the cost of network benefits. Third, the media, cultural groups, and interpersonal networks may have *socially amplified the risk* (Kasperson et al., 1988) of the COVID-19 pandemic, which was ongoing at the time we conducted the experiment, and thus triggered responses that went beyond the immediate mitigation of harm.

1.3.4 Chapter 5: Prioritizing high-contact occupations raises effectiveness of vaccination campaigns

In Chapter 5, we switch the perspective from how individual network behavior affects the spread of infectious diseases to how we can make use of network properties for epidemic intervention. Specifically, we take up on a twenty-year-old idea that prioritizing high-contact individuals can improve the effectiveness of vaccination campaigns (Dezsó & Barabási, 2002; Pastor-Satorras & Vespignani, 2002). However, an open question was how to identify high-contact individuals, as contact information is not only hard to retrieve but also private. Rather than identifying high-contact individuals, we propose the affiliation to an occupational group a proxy for connectedness in a social network.

To test whether occupational group can function as a suitable proxy, we first fitted SWIDM networks to contact data that has been collected during the early stages of the COVID-19 pandemic in China, South Korea, Japan, Italy, UK, and the US (Belot et al., 2020). We then simulated epidemics varying in distribution strategy (baseline / no vaccination, random distribution, prioritization of high-contact occupations), vaccine availability

(5% - 50% of the population), and probability of immunization (60% - 90%). Our results suggest that, in comparison to the baseline scenario, the number of infections can be lowered to a similar degree when either distributing vaccines randomly or when prioritizing high contact-occupational groups with only half the number of vaccines. In addition, the prioritization of occupational groups can reduce and delay peaks, so that fewer simultaneous infections occur. In conclusion, our results suggest occupational group as a readily available and operational proxy to increase the effectiveness of vaccination campaigns.

1.4 Discussion

The goal of this dissertation was to obtain a better understanding of the co-evolution of social networks and infectious diseases. That is, most network models of disease spread either neglect the dynamic nature of our social lives or model network changes as a random process. Therefore, the conclusions based on these models do not allow substantial insights into the co-evolutionary dynamics between social networks and infectious diseases. To address the problem of lacking or inaccurate network dynamics, we identified relevant theories from sociology, health psychology, and epidemiology, and integrated them into the NIDM. The NIDM is a highly adaptable model framework that describes network behavior as the tradeoff between the immediate benefits and costs, and the potential harm a social creation creates. To our knowledge, the NIDM is the first network model of disease spread that integrates health behavior theory explicitly.

1.4.1 Knowledge gained and implications

Using agent-based simulations we first studied how disease avoidance on the individual-level affects disease spread on the group-level. Some of the results we found were not surprising. We found, for example, that the more relations exist in a social network, the higher the probability for the disease to spread across the entire network. We also found that higher degrees of clustering can mitigate disease spread in small-world networks. Although someone may argue that these effects have already been reported by various studies before, we consider this to be a strong argument in favor of the NIDM and the model implementations presented in this thesis. Models based on the NIDM that support well-established findings thus raise confidence in findings that are new, surprising, or not necessarily in line with previous studies.

An opposing result that we found between static and adaptive networks is that higher degrees of clustering can shorten the duration of epidemics in adaptive networks, while higher degrees of clustering can prolong the duration of epidemics in static network. That is, non-adaptive agents retain bridging relations between two clusters during an outbreak, and thus retain the bottleneck for a disease to spread from one cluster to another. Adaptive agents, however, can dissolve cluster bridging relations to infectious neighbors, and thus isolate entire groups from the infection relatively easy. We believe that this insight provides an interesting addition to the seminal study of Watts and Strogatz (1998), who argue that bridges between clusters are shortcuts that allow infectious diseases to quickly reach any

part of a network. Although we agree with the general notion of Watts and Strogatz (1998), we also believe that these shortcuts are weak points for disease spread and can therefore be used to improve spatial distancing measures. That is, on the one hand reducing contact with distant relations (spatial distance, emotional distance) results in less disruption to the network. On the other hand, reducing contact with distant relations decreases the likelihood that the disease spreads rapidly throughout the entire network.

Another finding suggests that health behavior homophily can increase the mitigating effect of clustering. That is because the more agents perceiving high risks of infection inhabit a cluster, the more likely a high risk perceiving agent is part of a bridging relation, and the more likely the cluster is isolated from the infection. If, however, clusters are randomly mixed, all clusters are equally likely to get infiltrated by the disease. We believe that this effect could be exploited for epidemic intervention, such as targeted information campaigns for groups that are more likely to perceive low health risks. The literature shows various examples of how risk perception coincides with personal characteristics that tend to create homophily on the group-level, such as gender, age, educational status, etc. (Bish & Michie, 2010). One way to estimate the effectiveness of targeted information campaigns could be to use NIDM simulations in a manner similar to that described in Chapter 5. That is, first we identify a proxy relevant for a specific disease (e.g., men, younger than 25, HIV). Second, we generate contact networks with characteristics based on empirical data, such as degree distribution, degree of clustering, degree of homophily regarding the characteristics of the proxy (e.g., gender, age), and risk perception. Third, we define various strengths of how information campaigns affect the risk perception of the proxy group. Fourth, we use NIDM simulations to compare to what extent changes in risk perception affect disease spread.

Although our experiment did not provide sufficient data to either support or refute our theoretical finding that health homophilous clusters mitigate disease spread, we gained valuable insights into disease avoidance behavior of our participants. On the one hand, we found support for the behavioral assumptions of the NIDM. That is, infectious alters were avoided by susceptible egos and higher risk aversion caused stronger avoidance of infectious alters. On the other hand, we found that disease avoidance was not only stronger compared to what our simulations suggested, but also that participants used more sophisticated strategies for avoiding infection. That is, susceptible participants avoided not only infected neighbors but also susceptible neighbors with infected neighbors. Disease avoidance, however, came with the loss of beneficial network relations, and thus with the loss of experimental payoff.

Despite the difficulty to generalize the experimental results to the real world, we believe that this prioritization of disease avoidance over maintaining social cohesion can be considered a warning signal. That is, infection can be considered a salient signal (e.g., visible symptoms, self-reports) with clear consequences (e.g., experiencing symptoms, absenteeism) that trigger immediate behavioral responses (i.e., avoiding infectious alters). Loss of social cohesion, however, is comparatively hard to detect as it is a consequence of disease avoidance that only occurs over time. Furthermore, according to SPF theory a

person cannot compensate for a low level of social well-being with a high level of physical well-being, as both of these two universal, non-substitutable goals have to be pursued independently. Studies during the COVID-19 pandemic have already shown that social isolation can cause a whole variety of personal problems, such as depression, anxiety, lack of sleep, substance abuse, etc. (Banerjee & Rai, 2020; Heape, 2021; Kim & Jung, 2021; Pietrabissa & Simpson, 2020; Sepúlveda-Loyola et al., 2020). The prioritization of disease avoidance may thus reinforce isolation through imposed distancing measures and thus reinforce a slow and hard to recognize decay of social cohesion and overall well-being in the long term.

A way to mitigate the negative effects of disease avoidance on social cohesion could be the organization of social events under secure circumstances. During the COVID-19 pandemic, there have been various approaches to maintaining social contact. Among colleagues there have been interactive online events, such as virtual scavenger hunts, trivia competitions, and cook-alongs. Online gaming has shown to influence offline social support positively (Trepte et al., 2012), while outdoor activities performed at physical distance, such as yard work or picnics may be appropriate tools to foster social contacts for the elderly (Heape, 2021). Although, there are ideas to overcome negative effects of social isolation, there is much need to evaluate the effectiveness of existing ideas, and design new measures based on scientific research. Furthermore, it is important to carefully consider the wording and communication around non-pharmaceutical interventions. From cognitive psychology we know that language and the use of words shape our perceptions of the world (Boroditsky, 2011). The commonly used term *social distancing* thus may create unwanted side effects, such as reinforcing feelings of loneliness. Abel and McQueen (2020) therefore proposes to use the term *spatial distancing* rather than *social distancing*, while promoting *social closeness* at the same time. Another term, typically used in the epidemiological literature is *physical distancing*.

Furthermore, with the NIDM we have provided a toolbox that allows to combine formal models and computer simulations with empirical data in two ways. On the one hand, the NIDM allows producing testable hypotheses and conditions for experiments (see Chapters 3 and 4). The combination of simulations and experiments allows more substantial conclusions. That is, experiments can, for example, test the assumptions the model is built upon, and reveal the influence of cognitive limits or social preferences on assumed behavior. Findings from these experiments can then be used to refine the model and to fit parameter settings. On the other hand, the NIDM allows fitting network properties (e.g., degree distribution, clustering, homophily / assortative mixing) to empirical data. This allows more reliable predictions on how network properties and different interventions can shape the course of epidemics (see Chapter 5).

1.4.2 Limitations and future work

This dissertation has shown the capabilities of the NIDM to shed light on the co-evolution of social networks and infectious diseases. Additionally, the findings and limitations of our work can guide future research. Specifically, we identify two major pathways for future research: modelling and informing policymakers.

Regarding modelling, the NIDM can only mark the beginning of integrating health behavioral theory into network models of disease spread. That is, we have presented a highly adaptable stepping stone model, and we have shown how it can be used to provide meaningful insights through either simulations or experiments. With the SWIDM, for example, we have provided a model that provides generalizable insights because it can capture common properties of social networks (degree distribution, clustering, homophily). However, the experiment in Chapter 4 has shown that not only parameter settings of the SWIDM require a better empirical foundation, but also extensions may be required to capture the sophisticated strategies used by the participants to avoid infection. Plausible extensions based on our experiment could be the consideration of disease states of neighbors of neighbors for networking decisions, increasing the attractiveness of recovered alters, and decreasing the attractiveness of infected alters irrespective of the disease state of the ego.

Furthermore, in this dissertation, we used models that considered risk perception as a constant innate preference. However, surveys during the early times of the COVID-19 pandemic have shown that risk perceptions do not only differ between individuals and countries (Dryhurst et al., 2020) but also changed over time (Wise et al., 2020). One element that affects risk perception and health behavior is the consumption of media. Information on the global prevalence of a disease, as reported by the media, could therefore change the risk perception of the agents homogeneously. Local prevalence information, as reported among personal contacts, could change the parameter in highly infected areas of the network only. Such an approach could pave the way for understanding the influence of global information campaigns and social influence on disease avoidance.

In addition, behavioral adaptation in the NIDM is modeled from the perspective of susceptible individuals. That is, depending on heterogeneous risk perceptions, individuals may either engage in or avoid social contacts that carry infections. Although this is in line with a large body of empirical studies (e.g., Bish & Michie, 2010; Leppin & Aro, 2009) and a commonly used mechanism in individual-based health behavior models, such as the *Health Belief Model (HBM)*, the NIDM could be extended with a component that makes it less attractive for infected individuals to engage in social contact. Depending on the disease, an infected individual may experience symptoms that may force them to stay at home. Additionally, being infected comes with an ethical/normative responsibility not to endanger other people.

Another interesting path to follow is to study the effect of health opinions and social influence on the spread of infection. A recent model study by Teslya et al. (2022a), for example, investigated the conditions under which competing health opinions can co-exist and how they affect the spread of infections. A key finding is that the epidemic burden is smaller if the presence of an infection causes individuals to switch from health neutral

to health promoting behaviors. Additional model studies on COVID-19 suggest that early adoption of health promoting behaviors can delay epidemic peaks (Teslya et al., 2020), while diminishing compliance with spatial distancing during vaccination rollout promote the development of additional epidemic peaks (Teslya et al., 2022b). However, not only health promoting opinions affect the spread of infection. Especially fake news has been a highly visible problem since the beginning of the COVID-19 pandemic. People that are susceptible to irrational beliefs, for example, have been shown to be less likely to adhere to self-protective guidelines, such as handwashing or spatial distancing (Teovanović et al., 2021). Furthermore, people on Facebook tend to prefer friends who have a similar interest in either scientific news or conspiracy theories (Bessi et al., 2016). This form of homophily may have significant effects on the course of the pandemic, as it may form clusters of individuals that do not believe in the existence of COVID-19, therefore do not take self-protective measures, and are thus at higher risk of infection. To study to what degree conspiracy theory homophily keeps a pandemic alive, a specific NIDM model case could be designed that considers homophily according to whether individuals believe in fake news or not, accompanied by a lower risk perception towards health risks for individuals believing in fake news.

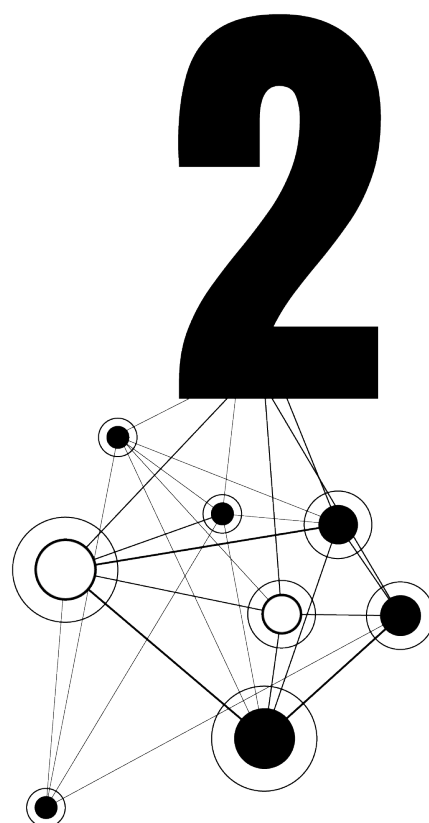
Another challenge is how to design models and experiments that find a good balance between abstraction and reality. How plausible is it, for example, that somebody knows the disease state of a neighbor or even that of a neighbor's neighbor? People may show symptoms and thus reveal their infectiousness involuntarily. People may also tell others that they are not feeling well or that their spouses are not joining for dinner because they are feeling ill. However, somebody who does not show symptoms may nevertheless be infectious. Follow-up simulations and experiments could therefore explicitly investigate to what extent disease avoidance is influenced by different degrees of uncertainty about the disease states of network partners. Surveys could furthermore inform how openly people communicate disease states when feeling ill and how strongly they avoid potentially infectious others. In addition, some relations may never be questioned (e.g., children) even in case of highly severe and infectious diseases. To integrate the quality of relations, a specific model case could add a weight parameter to the benefit component. A higher weight could then increase the benefit of a relationship, making it less likely to be cut despite infection.

Regarding policymaking, the dynamics of the NIDM can reveal important mechanisms that ought to be considered for the design of interventions. The SWIDM, for example, has shown that disease avoidance can cause the loss of social cohesion. However, the SWIDM is a very general model that does not necessarily allow immediate inferences on specific diseases and specific social context. Specific diseases rather need to be modeled based on empirical data given a plausible social context. Chapter 5 provides an example of how NIDM simulations and empirical data can be expressive for the comparison of vaccination campaigns within occupational networks. In addition, Chapter 5 shows that lacking social network dynamics are not a problem for network models of disease spread per se. It rather depends on the research questions at hand whether network dynamics

are expected to change infection dynamics significantly. Static network models remain a plausible abstraction when considering, for example, time scales or types of networks that are less prone to substantial relationship changes.

1.5 Conclusion

In conclusion, this dissertation provides a novel approach to describe the co-evolution of social networks and infectious diseases. With the NIDM, we propose a highly adaptive model framework that describes individual networking decisions as a trade-off between a relation's benefits, costs, and potential harm of infection. Specific model cases can reveal insights that can be used to design or improve epidemic interventions and guide future research. These contributions have gained much relevance since the COVID-19 pandemic demonstrated how vulnerable our increasingly globalized world is to infectious disease outbreaks.



Chapter 2

A model for the co-evolution of dynamic social networks and infectious disease dynamics

Abstract *Recent research shows an increasing interest in the interplay of social networks and infectious diseases. Many studies either neglect explicit changes in health behavior or consider networks to be static, despite empirical evidence that people seek to distance themselves from diseases in social networks. We propose an adaptable steppingstone model that integrates theories of social network formation from sociology, risk perception from health psychology, and infectious diseases from epidemiology. We argue that networking behavior in the context of infectious diseases can be described as a trade-off between the benefits, efforts, and potential harm a connection creates. Agent-based simulations of a specific model case show that: (i) high (perceived) health risks create strong social distancing, thus resulting in low epidemic sizes, (ii) small changes in health behavior can be decisive for whether the outbreak of a disease turns into an epidemic or not, (iii) high benefits for social connections create more ties per agent, providing large numbers of potential transmission routes and opportunities for the disease to travel faster, and (iv) higher costs of maintaining ties with infected others reduce final size of epidemics only when benefits of indirect ties are relatively low. These findings suggest a complex interplay between social network, health behavior, and infectious disease dynamics. Furthermore, they contribute to solving the issue that neglect of explicit health behavior in models of disease spread may create mismatches between observed transmissibility and epidemic sizes of model predictions.*

This chapter has been published as Nunner, H., Buskens, V., & Kretzschmar, M. (2021). A model for the co-evolution of dynamic social networks and infectious disease dynamics. *Computational Social Networks*, 8(1), 19. Nunner and Buskens developed the mathematical model. Kretzschmar contributed to its conceptualization. Nunner programmed the simulation, performed analyses, created visualizations, and wrote the manuscript. All authors reviewed the manuscript.

2.1 Introduction

In an increasingly globalized world, we face risks of outbreaks of infectious diseases, such as AIDS, Ebola, COVID-19, or new strains of influenza. Human behavior is known to have major influence on infectious disease dynamics (Heesterbeek et al., 2015). Examples are how we interact with another when we are sick, whether we break social ties over the course of infections, what we believe to know about health risks, and how we derive actions from this information.

Past studies show an increasing interest in this interplay of health behavior and infectious diseases in both sociological and epidemiological scholarship (Balcan et al., 2009; Ferguson et al., 2006; Gross et al., 2006; Reluga, 2010; Valdez et al., 2012a). Mao and Yang (2012), for example, emphasize that neglect of health behavior may create mismatches between observed transmissibility of diseases and epidemic sizes of model predictions. Based on a 2-layer network (contact, information) they predict lower epidemic sizes when more individuals adopt preventive behaviors disseminated through communication networks. Further integration of health behavior, the authors argue, is necessary and requires an interdisciplinary approach combining social sciences, health psychology and epidemiology.

From studies in health psychology we know, for example, that perceived risks of infections may cause people to avoid social gatherings during times of increased risks of infection (Funk et al., 2010; Jones & Salathé, 2009). Despite the widely recognized importance of incorporating dynamical aspects in epidemiological contexts (Bansal et al., 2010; Ferguson, 2007), many studies on infectious diseases consider social networks to be static, implying that health behavior does not affect social behavior. Static relations, however, do not describe the dynamic nature of our social networks (Christakis & Fowler, 2007; Palla et al., 2007), while disease dynamics is known to be sensitive to topology changes in adaptive networks (Pastor-Satorras et al., 2015).

Leung et al. (2018) recently addressed the problem of lacking social network dynamics. Using a formal model of disease spread, they showed that social distancing implemented by stochastic rewiring of network ties, rather than dropping edges, and thus keeping average degree constant, has an immediate effect on the epidemic size. High rewiring rates of susceptible agents when connected to infected peers results in a lower epidemic size than in static networks. Moderate rewiring rates, in contrast, result in larger epidemic sizes. The authors propose, however, that individual decision-making is a non-stochastic, but highly complex health behavioral process waiting to be formally integrated into epidemiological social network models. We posit that a more realistic description of epidemics is possible by integrating social network dynamics based on health behavior into epidemiological models. This paper identifies the necessary theories and provides a general model to fill this gap.

Considering health behavior, Funk et al. (2009) argue that people need to become aware of disease-related information (e.g., personal observations, public information campaigns) to change their behavior. Further, two systematic reviews on airborne diseases (Bish & Michie, 2010; Leppin & Aro, 2009) conclude that subjective risk perception with regard to (i) the perceived probability of getting infected and (ii) the perceived severity of a disease

is the main reason for health-related behaviors in the context of infectious diseases. Thus, a person perceiving risk to be high may prefer to avoid social contacts during a wave of influenza, while a person perceiving risk to be low may not alter social behavior at all (see also, Ferguson, 2007; Goodwin et al., 2009; Jones & Salathé, 2009). The model by Poletti et al. (2012) also suggests effects on epidemic impact depending on perceived risks and the corresponding behavioral reaction (lowering probability of infection), even for small reductions in contact numbers. Finally, Bish and Michie (2010) argue that the number and severity of symptoms are significant factors for the duration and degree of separation from social contacts.

The elaborations above demonstrate a complex co-evolution of infectious disease dynamics and social network dynamics mediated by risk perceptions of actors. It remains, however, unclear how this co-evolution comes about. We therefore ask: *How does health behavior shape the co-evolution of epidemics and dynamic social networks?* Despite the enormous increase in the number of studies of disease spread in social networks since the beginning of the COVID-19 pandemic (e.g., Abdulkareem et al., 2020; Arenas et al., 2020; Brotherhood et al., 2020; Della Rossa et al., 2020; Firth et al., 2020; Herrmann & Schwartz, 2020; Karaivanov, 2020; Liu et al., 2020; Maher et al., 2020; Paré et al., 2020; Peirlinck et al., 2020; Scata et al., 2020; Silva et al., 2022; Weeden & Cornwell, 2020), we are not aware of any study explicitly considering a choice-based model including health behavior and interdependent dynamics of social networks and infectious disease spread. Modeling such network related health behavior allows to compare the effects of different control measures against epidemic outbreaks, a crucial factor for policy-making (Davis et al., 2015; Heesterbeek et al., 2015). Furthermore, our work adds to the rich literature on co-evolution between social networks and different types of behavior (Burk et al., 2007; Corten & Buskens, 2010).

To answer the question, we provide a general model for the co-evolution of dynamic social networks and infectious diseases mediated by risk perceptions of actors. The model follows an interdisciplinary approach and combines theories of social network formation from sociology and economics, infectious diseases from epidemiology, and individual health behavior from health psychology. Furthermore, rather than using purely stochastic rewiring processes or artificially lowering transmissibility (see Ferguson et al., 2006; Leung et al., 2018; Mao & Yang, 2012), we address two crucial shortcomings of contemporary models to couple the dynamics of social networks and disease behavior based on a more explicit theoretical mechanism: (i) the integration of dynamic network structure in consideration of (ii) behavioral changes due to infection risks (Wang et al., 2015).

We first identify the minimal requirements for such a model with solid footing in the literature (Section 2.2), including theory from epidemiological scholarship (Hens et al., 2012) and theory on how people form social ties (Doreian & Stokman, 1997; Jackson, 2008; Weesie & Flap, 1990). In addition, we use psychological theory to describe health behavior and how people translate the exposure to health risks into action. We integrate these theories to propose a general model describing the co-evolutionary processes of dynamic social networks and infectious diseases (Section 2.3). A specific model case (Section 2.4)

implemented as agent-based simulation (Section 2.5), serves to illustrate how the general model can be tied to specific types of social networks (here: 10 to 50 individuals, such as groups of friends, school classes, extended family, small companies, or teams within larger companies), and how different social, psychological, and disease related conditions affect the course of epidemics within these networks (Section 2.7).¹ We conclude with the implications of our study and opportunities for further research (Section 2.8).

2.2 Theory

2.2.1 Infectious diseases

Infectious diseases differ in many aspects, such as transmission routes (e.g., airborne, sexual contact, animal vectors, food contamination), the symptoms they cause (e.g., fever, coughing, fatigue, diarrhea, muscle aches), the body parts they affect (lungs, skin, inner organs), virulence, the cause of infection (bacteria, viruses, fungi, parasites), the course and duration of the disease, and many more. However, despite the numerous differences between specific infectious diseases, there are also commonalities that apply to all. The infection rate, or probability of an infection per contact, is central for the spread of any infectious disease and thus an important factor whether the invasion of an infection may turn into an epidemic (Funk et al., 2009). Furthermore, the stages individuals go through during an infection are similar. This progress typically starts with an individual being susceptible to a disease. After being infected with a pathogen, the susceptible individual becomes infectious and possibly symptomatic. Finally, the individual either recovers or dies from the infection. Recovered individuals may either be immune or once again become susceptible to the disease.

This general course of infectious diseases can be translated into the well-known compartmental *SIR model* (Hens et al., 2012), which divides a population into three different compartments according to disease states (Susceptible, Infectious, and Recovered). Depending on transmission and recovery rates, ordinary differential equations allow predicting different properties of infectious disease dynamics, such as the total number of infected, peaks and duration of epidemics, and effects of vaccines. Models are designed depending on, inter alia, which states of the infection are considered or whether recovered individuals may die, become immune, or become susceptible again.

Despite the undisputed importance of compartmental models in epidemiology, Wang et al. (2015) describe the lack of theoretical foundation in behavioral responses and social mixing as a pitfall. Especially the assumption that individuals have the same probability to get in contact with others is delusive, as people get more likely in contact with others who are close to them (spatially, emotionally, characteristically) (McPherson et al., 2001). More recent studies addressed this issue by dividing populations into spatially distinct groups, so-called *meta-populations* (Balcan et al., 2009; Colizza et al., 2007), or using

¹Clearly, the NIDM also allows modeling larger networks. That, however, requires a different social network formation model than the one used in this paper.

network approaches to study the effect of network topology, interpersonal relationships, informational exchange, or protective behavior on disease spread (Bansal et al., 2010; Durham & Casman, 2011; Kitchovitch & Li, 2010; Leung et al., 2018; Mao & Yang, 2012; Pastor-Satorras et al., 2015; Perez & Dragicevic, 2009). While static networks provide good approximations when networks change at a much slower pace than diseases spread, an increasing number of studies is concerned with network dynamics and disease dynamics occurring on similar timescales (Pastor-Satorras et al., 2015). Model studies on adaptive networks allow unique insights on the non-trivial feedback loop between infectious disease spread and spontaneous changes in human behavior. Game-theoretic approaches study, for example, the conditions that make social distancing a beneficial response to infection risks (Chang et al., 2020; Reluga, 2010). Studies to model temporary interruption of contacts between infected actors (Tunc et al., 2013; Valdez et al., 2012a; Valdez et al., 2012b) show that pausing relationships during epidemics increase the epidemic threshold. The underlying network structure in these studies, however, remains static. In a study to model epidemic dynamics in adaptive networks, Gross et al. (2006) showed that disease spread is inhibited when susceptible actors choose to distance themselves from infected others. Furthermore, they showed that depending on rewiring rate and initial state of the system, healthy and endemic states can coexist.

While all these approaches impressively show that network dynamics may alter the course of epidemics, they neglect modeling health behavior explicitly.² We therefore seek to identify the relevant theories to describe how people form their networks and how this behavior differs when being at risk of getting infected with a disease.

2.2.2 Social network formation

Depending on the context, different positions within a network may create structural advantages, which in turn create social or economic benefits (Burt, 2001; Coleman, 1988; Cook et al., 1983; Granovetter, 1973). As a result, individuals consciously change their networks to their personal advantage (Flap & Völker, 2004; Watts, 1999).

Benefits in our general model are considered personal well-being, a combination of social and physical well-being, as suggested by *Social Production Function* theory (*SPF* theory; Ormel et al., 1999). Social well-being can be satisfied through affection or behavioral confirmation by direct personal contacts. This creates an incentive to establish connections to other persons. Maintaining ties, however, comes at the cost of time and effort (Jackson, 2008). It is therefore not necessarily rational for an actor to just randomly connect to everyone else, as some ties might create higher value than others while others have higher costs for maintenance.

According to *SPF* theory, a major factor of physical well-being is the absence of physical harm. Consequently, the presence of an infectious disease within a social network creates a potential harm to each individual and thus creates an incentive not to form or even to break

²Note that health behavior is not necessarily ignored in models of disease spread. If, for example, estimation of epidemic parameters, such as R_0 is based on empirical data, behavior is reflected in the data and thus implicitly included.

existing ties with (infectious) others. As a result, networking behavior becomes a trade-off (utility U_i) of an individual (i) between the social well-being his/her contacts create (benefit B_i), the costs to maintain the ties (C_i), and the physical harm the connection may cause due to an infectious disease (D_i):

$$U_i = B_i - C_i - D_i. \quad (2.1)$$

The first two components of the equation describe the net social utility of contacts disregarding the potential harm of infections. Jackson (2008) provides a framework to model this. These *Strategic Network Formation Models* (Jackson, 2008, ch. 6) formalize how and why individuals form connections based on the costs and benefits of ties. Analyses of the resulting network and utility structures enables to explain the psychological and societal constraints under which certain network properties come about.

It has to be noted that appropriate models of utility structures depend on the type of disease and its mode of transmission (HIV: sexual contacts, measles: classrooms). Further, potential costs of infections affect expected utility. Actors, however, may apply behavioral changes regardless of the exact shape of the utility function to avoid infections as we argue below.

2.2.3 Health behavior and risk perceptions

Health behaviors affect many everyday decisions including nutritional issues, personal and sexual encounters, or substance use (Catania et al., 1990; Ferguson, 2007; Institute of Medicine and National Research Council, 2011). Regarding infectious diseases, Bish and Michie (2010) classify health behaviors into three categories: (i) preventive behaviors (e.g., hand washing, mask wearing, vaccinations), (ii) avoidant behaviors (e.g., work absence during a wave of influenza), and (iii) management of disease behaviors (e.g., consulting medical experts). All these behaviors lower or even eliminate the risks of getting infected. Here, we focus on avoidant behaviors, as we expect it to have the greatest effect on social network structures. For social networks this translates to *avoiding potentially infectious social connections*. Such social distancing can be achieved through either choosing not to form a tie to, or breaking an existing tie from, an infectious other. Empirical studies support that social distancing successfully reduces infection risks through link removal (Funk et al., 2010), voluntary quarantine (Tracy et al., 2009), or avoidance of public places (Jones & Salathé, 2009). Ahmed et al. (2018) show further that social distancing at the workplace reduces epidemic size. But how do people choose to avoid others when being at risk of infections?

In their systematic review of 30 articles on SARS and avian influenza, Leppin and Aro (2009) describe that independent of how different studies conceptualize risk (i.e., subjective expected utility vs. psychometric), the main commonality among all studies is that risk perception is the driving factor for health-related decisions. Further, risk perception is affected mainly by the perceived susceptibility to and the perceived severity of a disease (Bish & Michie, 2010; Leppin & Aro, 2009). A systemic review by Bults et al. (2015) em-

phasizes the subjective nature of perceived susceptibility: people typically make inaccurate predictions about health risks and thus actual and perceived risks of getting infected may differ greatly. While the findings for severity are not as consistent, there is also abundant evidence that the perceived severity of a disease (e.g., expected harm for health, expected fatality) has a major effect on risk assessment as well (Bults et al., 2015; Leppin & Aro, 2009).

A behavioral model that describes how individuals make health-related decisions based on the aforementioned considerations is the ego-centered *Health Belief Model* (in short: *HBM*; Green & Murphy, 2014). Similar to the HBM, our model integrates the empirically documented concepts *perceived susceptibility* and *perceived severity* of a disease to capture health behavior driven by risk perception.

2.3 The networking during infectious diseases model (NIDM)

Based on the foregoing theoretical considerations a formal representation of networking behavior and infectious diseases needs to satisfy two requirements: First, social distancing is the result of a deliberation process that weighs the net benefits for keeping a connection to an infectious peer and the harm of a disease that potentially results from the same connection. Second, objective measurements for the harm of a disease (susceptibility, severity) need to be modifiable to satisfy the subjective nature of risk perceptions.

The integration of these requirements is expressed by the *Networking during Infectious Diseases Model* (*NIDM*). The NIDM assumes an actor (i) to optimize the utility function in Equation 2.2 (U), an elaborate version of Equation 2.1, combining the benefit of social connections (B_i), the costs to maintain ties (C_i), and the potential physical harm of infectious contacts (D_i):

$$U_i(\mathbf{G}, \mathbf{d}, \mathbf{R}) = B_i(\mathbf{G}, \mathbf{d}) - C_i(\mathbf{G}, \mathbf{d}) - D_i(\mathbf{G}, \mathbf{d}, \mathbf{R}). \quad (2.2)$$

Utility depends on the network structure (\mathbf{G}), the disease state of all actors (\mathbf{d}), and their risk perceptions (\mathbf{R}). We assume individuals to act boundedly rational using the information available through their environment. That is, individuals act strategically in order to maximize their myopic personal benefits.

In the following we present a specific case of the NIDM to illustrate how it can be used to study the co-evolution of social networks and infectious diseases for a specific type of social network.

2.4 A specific model case

In a first step to create a specific model from the NIDM, we define a baseline utility function for network formation. This utility function describes the social context and thus determines the properties of the social network to be studied. We choose the truncated

version of the *Connections Model* (CM; Jackson, 2008; Jackson & Wolinsky, 1996). By choosing a well-known model for an initial investigation, we hope to facilitate understanding of the results. The CM defines utility (U) of an actor (i) as the combination of benefits α of connections at distance 1 (direct connections), benefits β of connections at distance 2 (indirect connections), and the costs to maintain direct connections c :

$$U_i(\mathbf{G}) = \alpha \cdot n_i + \beta \cdot m_i - c \cdot n_i, \quad (2.3)$$

where n_i is the number of direct and m_i is the number of indirect connections. We use the truncated version implying that only benefits of connections at distances 1 and 2 are considered rather than also providing benefits for longer distances. The model represents that people do not only benefit from direct contacts by receiving help, support, information, etc., but that they can also obtain benefits indirectly from others connected to direct contacts.

Note that the CM is a model typically describing social contexts of limited group size, such as groups of friends, school classes, extended family, small companies, or teams within larger companies. Furthermore, while the CM is an ad hoc choice, it contains characteristics relevant for infectious diseases. Consider a sick person, who receives care by direct friends (doing groceries, taking over chores). In addition, friends of friends may be beneficial, for example by enabling the friend to help (taking over his/her chores) or in the form of practical matters (borrowing an inhaler, providing information on care).³ Additionally, the model provides an interesting example for the co-evolution of social networks and infectious diseases as networking behavior depends on benefits from direct and indirect connections, while diseases are transmitted only between actors in direct connection.

We consider connections to be unweighted, undirected, and non-reflexive; presented by the adjacency matrix $\mathbf{G} = g_{ij}$, with $g_{ij} \in \{0, 1\}$, $g_{ii} = 0$, and $g_{ij} = g_{ji} = 1$ if a tie between actors i and j exists. The *degree* is the number of ties at distance 1 of an actor:

$$n_i = \sum_j g_{ij}. \quad (2.4)$$

The *distance 2 degree* is defined by the sum of all indirect connections of an actor (t_{ij}):

$$m_i = \sum_j t_{ij}, \text{ with} \quad (2.5)$$

$$t_{ij} = \begin{cases} 1, & \text{if } (\mathbf{G}^2)_{ij} > 1 \text{ and } g_{ij} = 0 \text{ and } i \neq j \\ 0, & \text{otherwise.} \end{cases} \quad (2.6)$$

³Note that friendship is merely an illustrative example. The NIDM describes general social network and infectious disease dynamics that can be tied to any context. That is, connections may form and dissolve over time based on health behavior. Thus, if someone gets sick, interactions may discontinue independent of their nature (e.g. going for lunch, having a business meeting, having sexual contact), but have the possibility to recover at a later point in time.

We extend the CM by considering a generic infectious disease and health behavior driven by risk perception. The utility in the resulting *Connections during Infectious Diseases Model* (CIDM) is therefore only affected when a disease is introduced into the network. We define the vector of disease states for all actors:

$$\mathbf{d} \in \{S, I, R\}^N, \quad (2.7)$$

with $d_i = S$ if i is susceptible, $d_i = I$ if i is infected, and $d_i = R$ if i is recovered.

Consequently, the ties of an actor (g_{ij}) can be categorized by disease state. For distance 1, we define:

$$n_{i_X} = \sum_{j, d_j=X} g_{ij}, \quad \text{where } X = S, I, \text{ or } R, \quad (2.8)$$

as the number of actors at distance 1 with disease state X , while for distance 2 that is:

$$m_{i_X} = \sum_{j, d_j=X} t_{ij}, \quad \text{where } X = S, I, \text{ or } R. \quad (2.9)$$

In addition to the network structure (\mathbf{G}), social benefits (B_i) of actor i depend now on the disease state of connected peers (\mathbf{d}):

$$B_i(\mathbf{G}, \mathbf{d}) = \alpha \cdot (n_{i_S} + \kappa \cdot n_{i_I} + n_{i_R}) + \beta \cdot (m_{i_S} + \lambda \cdot m_{i_I} + m_{i_R}), \quad (2.10)$$

where $0 \leq \kappa \leq 1$ is a discount factor for the value of infected direct (n_{i_I}) and $0 \leq \lambda \leq 1$ is a discount factor for the value of infected indirect connections (m_{i_I}). Thus, Equation 2.10 captures that infected connections may not be able to provide support as usual. A similar approach is used for the costs of actor i to maintain social relations:

$$C_i(\mathbf{G}, \mathbf{d}) = c \cdot (n_{i_S} + \mu \cdot n_{i_I} + n_{i_R}), \quad (2.11)$$

where $\mu \geq 1$ is a cost increase for infected direct connections (n_{i_I}). Therefore, Equation 2.11 suggests that maintaining infected connections may result in higher efforts, for example due to nursing care.

(Potential) harm of infections for actor i (D_i) is the product of probability to get infected (p_i) and severity of the infection (s_i). D_i depends on network structure (\mathbf{G}), the disease state of all actors (\mathbf{d}), and risk perceptions (\mathbf{R}):

$$D_i(\mathbf{G}, \mathbf{d}, \mathbf{R}) = p_i(\mathbf{G}, \mathbf{d}, \mathbf{R}) \cdot s_i(\mathbf{d}, \mathbf{R}). \quad (2.12)$$

In order to model the subjective nature of risk perceptions, we distinguish between two modifiers:

$$\mathbf{R} : \mathbf{r}_\pi \in \mathbb{R}, \mathbf{r}_\sigma \in \mathbb{R}, \quad (2.13)$$

where r_π modifies the actual probability to get infected (see Equation 2.14) and r_σ modifies actual severity of the infection (see Equation 2.16).

The probability to get infected is divided into three different cases, depending on an actor's own disease state:

$$p_i(\mathbf{G}, \mathbf{d}, \mathbf{R}) = \begin{cases} \pi_i(\mathbf{G}, \mathbf{d})^{2-r_\pi}, & \text{if } d_i = S \\ 1, & \text{if } d_i = I \\ 0, & \text{if } d_i = R. \end{cases} \quad (2.14)$$

In case an actor is in the recovered state ($d_i = R$) the probability to get infected is 0 (immune). In case the actor is infected ($d_i = I$), the infection is present ($p_i = 1$). If an actor is susceptible ($d_i = S$), the probability to get infected depends on two factors. First, the actual probability to get infected:

$$\pi_i(\mathbf{G}, \mathbf{d}) = 1 - (1 - \gamma)^{n_{iI}}, \quad (2.15)$$

where γ is the objective probability to get infected per single contact. Second, the risk perception factor for the probability to get infected ($0 \leq r_\pi \leq 2$). The power function accounts for the uncertainty of actors with regard to their own susceptibility. Note that we use $2 - r_\pi$ as the exponent so that the interpretation of r_π is such that if r_π increases, actors subjectively estimate the risk of infection to be higher. Thus, $r_\pi = 1$ implies an accurate estimate, while $r_\pi < 1$ and $r_\pi > 1$ represent an underestimation and an overestimation of personal susceptibility, respectively.

Finally, disease severity (s_i) describes how strongly an actor is affected by the symptoms of the disease. Perceived severity of the disease is again dependent on an actor's disease state (\mathbf{d}) and risk perceptions (\mathbf{R}):

$$s_i(\mathbf{d}, \mathbf{R}) = \begin{cases} \sigma^{r_\sigma}, & \text{if } d_i = S \\ \sigma, & \text{if } d_i = I \\ 0, & \text{if } d_i = R. \end{cases} \quad (2.16)$$

An actor in the recovered state ($d_i = R$) is immune and thus cannot be affected by the disease ($s_i = 0$). Infected actors ($d_i = I$) experience the objective severity of the disease ($s_i = \sigma$, with $\sigma > 1$). For susceptible actors ($d_i = S$), however, the risk perception factor ($0 \leq r_\sigma \leq 2$) transforms actual severity into subjectively perceived risks of a disease.

2.5 Simulation

In the following, we illustrate the framework for agent-based CIDM simulations. First, we describe the simulation procedure. Parameter settings used to study model behavior are explained thereafter.

2.5.1 Simulation procedure

A single simulation consists of two stages, each composed of a number of time steps (iterations): (i) network initialization (150 time steps) and (ii) epidemic (200 time steps). The number of time steps is fixed to standardize subsequent analyses. It is also sufficiently large to always attain *pairwise stable* networks (Jackson & Wolinsky, 1996) at the end of both stages, and disease-free networks at the end of the epidemic stage. Pairwise stability means that no agent either benefits from breaking an existing tie unilaterally and no pair of actors from creating a non-existing tie. Pairwise stability is tested at the end of each time step by checking for all possible pairs of agents whether adding a non-existing tie or removing an existing tie increases utility.

In the first time steps of the epidemic stage a single randomly selected agent is infected with a communicable disease. By introducing the disease into a pairwise stable network we ensure that changes in network structure are solely based on health behavioral reactions of the agents.

A single time step consists of two distinct, consecutive, and interdependent processes: (i) disease dynamics followed by (ii) social network dynamics. Disease dynamics define the transmission of infections between agents:

- Repeat until all agents have been processed:
 - Randomly select an unprocessed agent i .
 - If i is susceptible, compute whether i gets infected from infected direct connections (equation 2.15).
 - If i is infected, compute whether agent recovers: passed time steps since infection $\geq \tau$.

Social network dynamics define the formation and termination of ties between agents:

- Repeat until all agents have been processed:
 - Randomly select an unprocessed agent i .
 - i randomly retrieves a proportion ϕ of all other agents j in the network: $\phi \cdot (N - 1)$.
 - Repeat until all retrieved agents have been processed:
 - Randomly select (another) retrieved co-agent j :
 - If i is connected to j :
 - Terminate tie ij , if the utility for i without ij is larger than the current utility with ij .
 - If i is not connected to j :
 - Form tie ij , if utility for both i and j is larger with ij than current utility without ij .

Table 2.1: Overview of CIDM parameters.

Parameter	Admissible	Used
<i>Utilities</i>		
<i>I. Social benefits (B_i)</i>		
Benefit of direct ties	$\alpha \in \mathbb{R}$	$\alpha = 10$
Discount of infected direct ties	$0 \leq \kappa \leq 1$	$\kappa = 1.0$
Benefit of indirect ties	$\beta \in \mathbb{R}$	$\beta \in \{2, 8\}$
Discount of infected indirect ties	$0 \leq \lambda \leq 1$	$\lambda = 1.0$
<i>II. Social maintenance costs (C_i)</i>		
Costs to maintain direct ties	$c \in \mathbb{R}$	$c = 9$
Cost increase for infected direct ties	$\mu \geq 1$	$\mu \in \{1.0, 1.5\}$
<i>III. Potential harm of infections (D_i)</i>		
Diseases severity	$\sigma > 1$	$\sigma \in \{2, 10, 50\}$
Probability of getting infected per contact	$0 \leq \gamma \leq 1$	$\gamma = 0.1$
Risk perception (disease severity)	$0 \leq r_\sigma \leq 2$	$r_\sigma = r_\pi = r \in \{0.5, 1.0, 1.5\}$
Risk perception (probability of infection)	$0 \leq r_\pi \leq 2$	
Recovery time	$\tau > 0$	$\tau = 10$
<i>Network</i>		
Network size	$N > 1$	$N \in \{10, 15, 20, 25, 50\}$
Initial network structure	$\iota \in \{\text{empty}, \text{full}\}$	$\iota \in \{\text{empty}, \text{full}\}$
Proportion of ties to evaluate per time step	$0 < \phi \leq 1$	$\phi = 0.4$

2.5.2 Parameter settings

Table 2.1 shows an overview of admissible and used CIDM parameters. These parameters fall into two categories: (i) ego-centered utility parameters and (ii) network parameters. We decided on a minimal but expressive selection of parameter variations that allows to investigate the effects of each model component on overall behavior. That is, we varied only one parameter per term in the utility function (i.e. benefit of indirect ties for social benefits, cost increase for infected direct ties for social maintenance costs, disease severity as property of diseases and risk perception as property of agents for potential harm of infections), and selected settings that allow strong variations in dynamics. Further, we used a limited number of fixed values rather than randomized values for each simulation run. This allows to disentangle model behavior with regard to parameter settings and stochastic processes of the simulations, and to exercise more control over the parameter combinations we believe to be of interest.

2.5.2.1 Utilities

In accordance with the literature on the CM, we consider relatively low net benefits of direct connections and relatively high net benefits of connections at distance 2 ($\alpha - c < \beta$). That is because net benefits of direct connections that are higher than benefits of connections at distance 2 ($\alpha - c > \beta$) inevitably result in fully connected networks. To distinguish settings in which indirect ties are valued relatively highly (e.g., professional networks providing

access to resources), and settings in which indirect ties are valued relatively lowly (e.g., friendships likely to result in triadic closure), we compare two parameter settings: $\beta = 2$ and $\beta = 8$. To limit the change in benefits to a single varied parameter, we neglect changing benefits due to infections ($\kappa = 1.0$, $\lambda = 1.0$).

To investigate overall lowered social value of direct relations ($B_i - C_i$) we vary costs for infected connections. One setting without increased costs ($\mu = 1.0$) and one with increased costs for infected ties ($\mu = 1.5$). This translates to situations in which infected individuals typically recover alone while staying independent and situations that require support from others, because infected individuals cannot master their everyday life alone (e.g., requiring somebody to get groceries). Costs to maintain connections is set to a constant value ($c = 9$) to accomplish two things. First, we do not obtain fully connected networks, because net benefits for direct connections ($10 - 9 = 1$) are smaller than benefits of connections at distance 2 for both settings. Second, net benefits of direct and indirect connections combined vary greatly depending on benefit of connections at distance 2 (β).

Diseases in the current implementation of the CIDM are considered generic. Thus, they may range from mild ($\sigma = 2$) through moderate ($\sigma = 10$) to severe ($\sigma = 50$). The probability to get infected per contact and time step is constant at $\gamma = 0.1$. We assume that risk perceptions with regard to disease severity and susceptibility coincide ($r_\sigma = r_\pi = r$). Further, we assume agents of the same population to perceive risks equally. This form of homogeneity applies more to social context and equal distribution of information rather than individual differences. Consequently, we compare three different types of populations with regard to risk perception. First, agents perceiving risk to be lower than actual risk (*low risk*; $r_\pi = r_\sigma = r = 0.5$). Second, agents perceiving risk realistically (*realistic risk*; $r_\pi = r_\sigma = r = 1.0$). Third, agents perceiving risk to be higher than actual risk (*high risk*; $r_\pi = r_\sigma = r = 1.5$). Infected agents in the CIDM require a fixed number of time steps to recover ($\tau = 10$). Simulation test runs showed that this combination of settings allows to prevent all agents to get infected or to disconnect from infectious others immediately after introduction of the disease into the network, and consequently to create strong variations of network and disease dynamics.

2.5.2.2 Network

We simulated populations of 10, 15, 20, 25, and 50 agents. Although larger populations are of empirical interest for disease dynamics, the CM typically describes social contexts of limited group size. Furthermore, pilot simulations have shown that results for larger network sizes do not differ qualitatively from networks with 50 agents, due to increasingly large degrees of agents in larger networks. Additionally, we differ between two starting conditions with regard to network density: empty ($\iota = \text{empty}$, with $g_{ij} = 0$ for all ij) and fully connected networks ($\iota = \text{full}$, with $g_{ij} = 1$ for all $i \neq j$ and $g_{ii} = 0$). This ensures that pairwise stability emerges from networks with different densities. Finally, we set the number of existing or potentially new ties an agent may evaluate per time step to 40% of the population ($\phi = 0.4$). This is done to account for the idea that people do not always have full control over their social relations (a train conductor at work, parents

picking up their children from daycare) and the possibility of disease transmissions from asymptomatic and thus unrecognized infectious persons. Again, this holds true for every agent, thus referring to social context (e.g., work environment) and leveling out individual differences within the population.

We systematically combined all previously defined parameter values, resulting in a total of 360 different parameter combinations. Further, we ran 100 simulations for each parameter combination, resulting in a total number of 36,000 simulation runs.

2.5.3 Software

The simulation was programmed using the Java 8 programming language and the Graph-Stream 1.3 library (Pigné et al., 2008) for graph handling. The complete code, including an executable program and an easy-to-use graphical user interface, is freely accessible under the GPLv3 license (Nunner, 2020).

2.6 Data and analysis

We log detailed network and disease information for each agent and time step of each simulation run. To understand how network dynamics shape the course of epidemics, we use regression analyses of simulation data (cf., Buskens & Snijders, 2016; Buskens & Yamaguchi, 1999) with regard to: (i) the proportion of recovered agents at the end of each simulation (*final size of the epidemic*, or short: *final size*) and (ii) the number of time steps until all once infected agents have recovered (*duration of the epidemic* or short: *duration*). To disentangle how much of our results are driven by parameter settings and how much is due to stochasticity of the simulation, we use two-level random-intercept regressions (level 2: 360 parameter combinations; level 1: 100 simulation runs). Random elements in the simulation are: (i) agent who is initially infected, (ii) order of agents that change their disease states on network connections, (iii) which network connections are considered and in which order. Furthermore, we create four models, each based on the previous model: (i) *Empty*, (ii) *CIDM effects*, (iii) *Network effects*, and (iv) *Interaction effects*. The Empty model shows how much unexplained variance is due to stochasticity between and within the same parameter settings. It further allows to compare the reduction of unexplained variance when controlled for the varied CIDM model parameters (CIDM effects model). The Network effects model further controls for network measures (network density, degree of patient-0), while the Interaction effects model adds statistically significant interaction effects of all previously described main effects. Data points for regression analyses are single simulation runs.

According to Long (1997), logit (and probit) models are most suitable when the dependent variable is a proportion of a binary response. Final size corresponds to this requirement as it describes the proportion of agents that have been infected with the disease and those that have not. The statistical model estimates how the logit of final size can

be approximated linearly by a combination of the parameters. Further, we use two-level random-intercept linear regression to analyze the model with regard to duration of the epidemic.

We consider linear relationships for all varied parameters, because they are either simulated with only two different values (β , μ , ι), or separate model parameters for more than two values (σ , r) and manipulation of parameters (N^x) did neither improve model fit (ℓ) nor substantively reduced unexplained variance (ρ). Interactions are constructed using grand mean centering, while maximum likelihood is used to estimate parameters. To allow easier interpretation of model coefficients, we rescale disease severity ($\frac{\sigma}{50}$) and network size ($\frac{N}{50}$).

For visual inspection of the progression of epidemics we generate SIR model plots, a method commonly used to display the temporal development of compartments during epidemics. We plot only time steps relevant for the epidemics. That is, rather than showing all simulated time steps, plots begin 10 time steps prior to the introduction of a disease (epidemic stage) and display 80 time steps in total. Next to showing the dynamics over all simulations, we also report the dynamics for networks in which agents change relations as a consequence of the introduction of a disease next to the dynamics for simulation runs in which this does not happen. Although we never actively turned off the option to change ties, there are quite some parameter constellations in which actors do not have a reason to change ties. A comparison between networks with and without tie changes therefore allows a clearer picture of the effects of network dynamics on epidemics.

To relate the network and disease dynamics to some overall network characteristics, we compute several characteristics of networks at different moments during the simulation. Density is computed as the proportion of all possible connections present in the network (Wasserman & Faust, 1994, p. 101 ff.). Clustering is calculated as the average proportion of closed triads over the number of all possible triads per agent (Watts & Strogatz, 1998). Finally, closeness is computed by reversing the normalized average distance between any two nodes in the network; network size (N) is imputed as distance for pairs of nodes that are not connected through any path in the network (Buechel & Buskens, 2013, p.163). For analyses we use R version 3.6.0 (R Core Team, 2019) with additional packages *lme4* (Bates et al., 2015) for logit regressions, *texreg* (Leifeld, 2013) for export of results, and *ggplot2* (Wickham, 2016) for data visualization.

2.7 Results and discussion

2.7.1 Model behavior

Figure 2.1 shows a clear interaction between social network and infectious disease dynamics, especially when comparing simulation runs with network changes (row 2) to simulation runs without network changes (row 3). Column 1, for example, shows the progression of epidemics over time. While in both cases (rows 2 and 3), after the introduction of the disease at time step 10, the proportion of infected agents increases and peaks at approx-

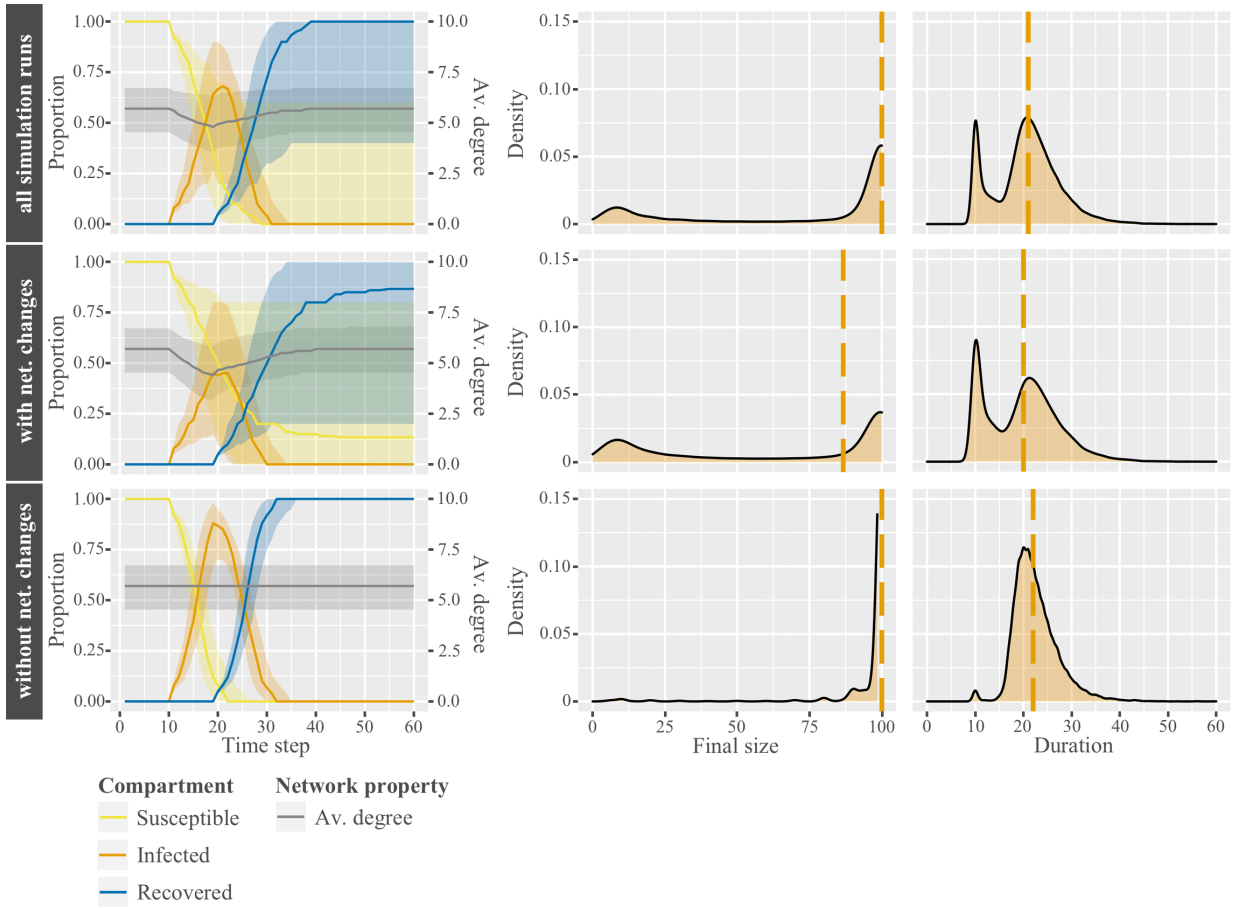


Figure 2.1: Disease dynamics in networks with and without network changes. Column 1 shows plots for the median proportional changes of agents according to disease states (yellow, orange, and blue lines), changes in median average network degree (gray lines), and interquartile ranges (shaded areas around median lines) over time. Column 2 shows the distribution of final size of epidemics. Column 3 shows the distribution of duration of epidemics. Row 1 contains the data of all simulation runs combined. Row 2 shows the data for simulation runs with network changes. Row 3 shows the data for simulations without network changes. The dashed orange line indicates the median of the respective distributions.

imately time step 21 (orange line), the peak for simulation runs with network changes is significantly lower. At the same time, average degree drops from about 6 to below 5 ties per agent in row 2, while average degree remains stable in row 3. Considering that networks were pairwise stable before the initial infection, the average number of 3.41 network changes per agent (sum of dissolution and creation of ties) is caused and mainly driven by the presence of infectious agents. Additionally, we observe that even more network changes were initiated by agents trying to connect to others (on average 18.00 tie requests per agent). These tie requests, however, were accepted only in 9% of the cases. This is a vast difference to the network initialization stage prior to the epidemics, where 40% tie

requests were accepted from an average of 7.97 tie requests per agent.⁴ Over time, infected agents recover, resulting in decreases in the infected compartment, increases in the recovered compartment, and a restoration of degree to the pre-epidemic state in row 2.

Furthermore, column 2 of Figure 2.1 reveals a bimodal distribution of final size in simulation runs with network changes (row 2). On the left a comparably small peak shows the simulation runs with only a single agent infected. On the right a large peak shows the simulation runs with all agents infected. It follows, that in most simulations the epidemic either died out immediately or spread across the entire network. This bimodal effect is the driving force behind the large width of interquartile ranges in row 2, especially for the final proportions of susceptible and recovered agents. In simulation runs without network changes (row 3), however, we observe virtually no variance in final size, as the vast majority has all agents infected at some point.

Considering the average duration of epidemics, we can observe a corresponding bimodal effect for simulation runs with network changes (column 3, row 2). The left peak shows the recovery time for a single infection. The right peak shows the average duration of epidemics for final sizes of close to 100%. Note that the left peak is more distinct for duration than final size, because one single infection requires exactly 10 time steps for recovery, while final size is a proportion depending on network size.

These data suggest that, on an individual level, agents succeed under certain conditions to avoid potential harm of diseases by distancing themselves from infected contacts during an epidemic. That is, agents cut ties in order to move into low risk network positions, unless social benefits outweigh the potential harm of an infection. On the population level, cutting ties ultimately leads to smaller and shorter epidemics. In the following we investigate which CIDM and network parameters drive the variances in these two epidemic measures.

2.7.2 CIDM parameters

Tables 2.2 and 2.3 show the effects of CIDM parameters on final size and duration of epidemics, respectively. The *Empty* regression model shows that in our following elaborations we will be able to attribute 69% of the overall variance in final size and 31% of variance in duration (see intraclass correlation coefficients ρ) to variation of parameters, while the remainders are caused by the stochastic elements of the simulation.

The *CIDM effects* regression model reveals that higher benefits of indirect connections (β) result in increases of final size and duration of epidemics. That is because higher social benefits may outweigh the potential harm of infections, and thus agents are more likely to stay connected. This causes more agents to get infected. Furthermore, the results of our model suggest that higher numbers of infected require more time for all agents to recover. Note that the longer duration is not an inevitable consequence of larger final sizes, as more intense epidemics may create many simultaneously infected agents without significant impact on duration.

⁴For more detailed descriptive statistics, we refer to Table B.2 in the Supplement to Chapter 2

Table 2.2: Two-level random-intercept logistic regression of final size of epidemics.

	Empty	CIDM effects	Network effects	Interaction effects
<i>Fixed/observed effects</i>				
Intercept	2.27 (0.15)	2.29 (0.08)	2.29 (0.08)	2.59 (0.07)
<i>Main effects</i>				
<i>I. CIDM parameters</i>				
Benefit distance 2 (β)		0.25 (0.02)	0.25 (0.02)	0.23 (0.02)
Cost increase for infected ties (μ)		-4.97 (0.30)	-5.00 (0.30)	-5.30 (0.21)
Disease severity ($\frac{\sigma}{50}$)		-1.87 (0.17)	-1.87 (0.17)	-1.75 (0.11)
Risk perception (r)		-2.74 (0.18)	-2.75 (0.18)	-3.16 (0.13)
Network size ($\frac{N}{50}$)		5.56 (0.29)	4.66 (0.37)	6.23 (0.38)
<i>II. Network properties</i>				
Density (den_{start})			-7.68 (1.00)	-7.33 (0.97)
Degree of patient-0 (deg_{start}^0)			3.86 (0.32)	6.46 (0.51)
<i>Interaction effects</i>				
$\beta \times \mu$				0.52 (0.06)
$\mu \times \frac{\sigma}{50}$				3.16 (0.44)
$\mu \times r$				6.40 (0.49)
$\frac{\sigma}{50} \times r$				-2.04 (0.27)
$r \times den_{start}$				9.44 (1.49)
$\frac{N}{50} \times deg_{start}^0$				13.80 (2.10)
<i>Random/unobserved effects (Intercept)</i>				
s^2	7.33	1.44	1.41	0.47
Log likelihood (ℓ)	-12,079.15	-11,814.47	-11,734.67	-11,581.32
Intraclass correlation (ρ)	0.69	0.31	0.30	0.13
Observations	36,000	36,000	36,000	36,000
Groups: parameter combinations	360	360	360	360

Notes: Bold coefficients are significant at $p < 0.001$. Others are not significant ($p > 0.05$). SEs are shown in parentheses.

Increasingly severe diseases, both objectively (σ) and subjectively perceived (r), and higher costs for infected direct ties (μ) create smaller final sizes and shorter duration of epidemics. That is because more severe diseases and higher costs for infected ties reduce the net benefit of a relationship to such a degree that potential costs of infections or existing costs exceed the social gains of the relation to an infected agent. Consequently, more agents disconnect from their partner, causing less agents to be infected, which in turn requires less time steps for the disease to disappear.

Finally, larger networks result in higher proportions of infected agents and longer lasting epidemics. The effect on duration is straightforward, as more consecutively infected agents require more time to recover. The effect on final size, however, is not intuitively comprehensible on its own. Thus, we consider the interplay with additional network properties in the following section.

Table 2.3: Two-level random-intercept linear regression of duration of epidemics.

	Empty		CIDM effects		Network effects		Interaction effects	
<i>Fixed/observed effects</i>								
Intercept	20.66	(0.20)	20.66	(0.17)	20.66	(0.16)	21.23	(0.19)
<i>Main effects</i>								
<i>I. CIDM parameters</i>								
Benefit distance 2 (β)			0.36	(0.06)	0.36	(0.05)	0.36	(0.04)
Cost increase for infected ties (μ)			−3.40	(0.66)	−3.42	(0.63)	−3.42	(0.44)
Disease severity ($\frac{\sigma}{50}$)			−1.82	(0.39)	−1.81	(0.38)	−1.81	(0.26)
Risk perception (r)			−2.80	(0.41)	−2.81	(0.39)	−2.81	(0.27)
Network size ($\frac{N}{50}$)			2.05	(0.59)	−3.20	(0.70)	−0.11	(1.15)
<i>II. Network properties</i>								
Density (den_{start})					−26.87	(1.96)	−19.93	(3.20)
Degree of patient-0 (deg_{start}^0)					3.30	(0.62)	3.31	(0.62)
<i>Interaction effects</i>								
$\beta \times \mu$							1.05	(0.15)
$\beta \times \frac{N}{50}$							−1.06	(0.13)
$\mu \times den_{start}$							−38.17	(5.00)
$\frac{\sigma}{50} \times r$							−4.16	(0.64)
$\frac{\sigma}{50} \times \frac{N}{50}$							6.48	(0.93)
$r \times \frac{N}{50}$							8.20	(0.96)
$\frac{N}{50} \times den_{start}$							32.97	(9.35)
<i>Random/unobserved effects (Intercept)</i>								
s^2		13.66		9.58		8.63		3.96
Log likelihood (ℓ)	−113,544.27		−113,481.97		−113,388.50		−113,274.74	
Intraclass correlation (ρ)	0.31		0.24		0.22		0.11	
Observations	36,000		36,000		36,000		36,000	
Groups: parameter combinations	360		360		360		360	

Notes: Bold coefficients are significant at $p < 0.001$. Others are not significant ($p > 0.05$). SEs are shown in parentheses.

2.7.3 Network properties

By adding network density and degree of the first infected agent at the time when the disease is introduced into the network, the *Network effects* regression model reveals larger final sizes of epidemics when: (i) network size increases, (ii) the network is less dense, or (iii) patient-0 has more ties. The positive effect of higher degrees for patient-0 can be explained with statistical probabilities. That is, the more connections an infected agent has, the more likely the disease is being transmitted to other agents. The other two effects require a closer look.

Table 2.4 shows a negative relationship between network size and density. Hence, larger networks have a lower proportion of actual connections compared to potential connections. Network size, however, has a positive effect on average degree. In other words, when networks grow larger the overall number of connections increases while fewer potential connections are formed. Applying the previous argument of statistical probabilities in

Table 2.4: Properties of epidemics by network size.

Network size	Density	Av. degree	Final size	Duration
10	0.404 (0.027)	3.632 (0.244)	56.987 (38.023)	18.406 (7.788)
15	0.338 (0.016)	4.736 (0.230)	66.636 (38.890)	20.434 (7.552)
20	0.299 (0.012)	5.688 (0.230)	73.880 (37.272)	21.481 (6.803)
25	0.273 (0.009)	6.545 (0.227)	80.395 (33.792)	21.999 (5.999)
50	0.204 (0.004)	9.994 (0.206)	95.855 (18.140)	20.971 (3.922)

context of average degree, we observe that a larger total number of connections increases the chance for a disease to spread. Thus, the effect of network size is related to the effect of average degree, which has a decisive impact on the final size of the epidemic.⁵

Figure 2.2 reveals that epidemics are larger for larger network sizes, despite similar behavioral reactions of the agents. Column 1, for example, shows that the reduction of average degree during the epidemics is similar for all network sizes (about 1 degree on average), while the overall average is at considerably different levels (3.632 for $N = 10$, 9.994 for $N = 50$). As a result, more connections remain in larger networks providing more potential transmission routes. Consequently, in large networks agents simply cannot dissolve enough ties to distance themselves from the disease sufficiently resulting in almost all agents getting infected in every simulation run. A comparison of disease dynamics between simulation runs with and without network changes during epidemics (Figures A.1 and A.2 in the Supplement to Chapter 2) shows that the differences of final size and duration of epidemics between different network sizes is largely driven by agents cutting ties and not merely an effect of higher average degrees in larger networks.

Furthermore, we observe a negative effect of network size on the duration of epidemics when combined with density and degree of patient-0 (*Network effects*, Table 2.3). That is, epidemics last longer when: (i) network size decreases, (ii) the network is less dense, or (iii) patient-0 has more ties. At first sight, these results may not seem plausible. In fact, networks with 10 agents have the lowest average duration of epidemics (18.406 time steps, see Table 2.4), while networks with 25 agents have the longest lasting epidemics on average (21.999 time steps). The largest simulated networks with 50 agents, however, show a decrease in average duration of epidemics (20.971 time steps). This non-linear effect for the duration of epidemics is explained by two opposing factors: First, the larger a network gets (assuming constant average degree) the more time is needed to infect the entire network. Second, the larger the network the larger the average degree (see Table 2.4). As discussed earlier, larger average degrees result in higher chances to spread a disease. Consequently, larger average degrees increase the likelihood for disease transmissions per

⁵Note that the effect of density should be interpreted in relation to the other network characteristics. One must realize that within the CM-model, density hardly varies given network size and as a result if we analyze data for one network size only and we do not control for the degree of patient-0, there is no effect of density. However, if we control for degree of patient-0 for a given network size, the negative effect can be interpreted as a compensation for the overestimation of the effect of patient-0 for the whole network. If we analyze data for different network sizes simultaneously, the density effect interferes with the effect of network size discussed above.

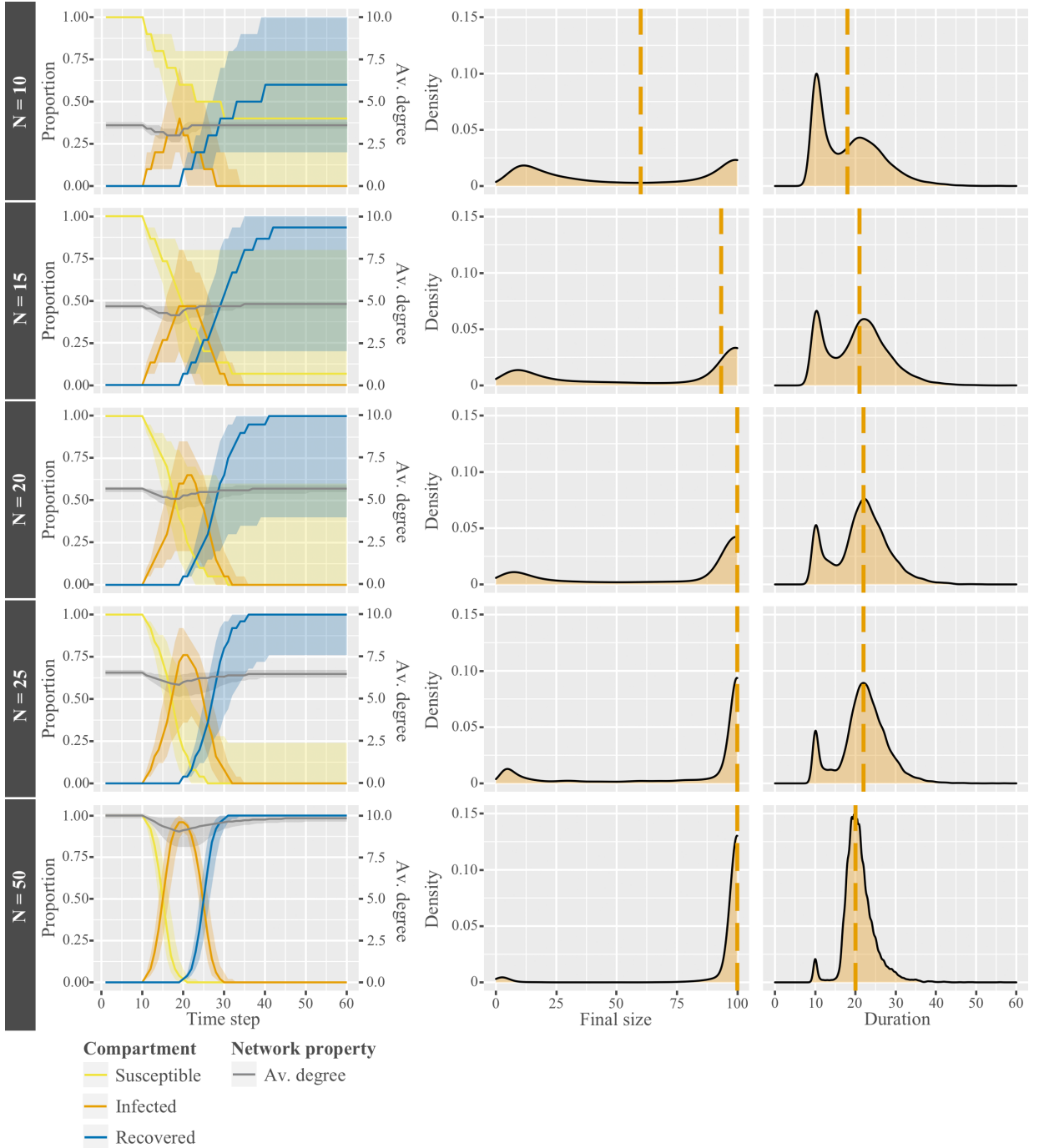


Figure 2.2: **Disease and network dynamics by network size.** Column 1 shows plots for the median proportional changes of agents according to disease states (yellow, orange, and blue lines), changes in median average network degree (gray lines), and interquartile ranges (shaded areas around median lines) over time. Column 2 shows the distribution of final size of epidemics. Column 3 shows the distribution of duration of epidemics. Rows show data split by network size. The dashed orange line indicates the median of the respective distributions.

time step and therefore facilitate faster disease spread than lower average degrees. Networks up to 25 agents approximate the tipping point of these opposing effects: Smaller networks are large enough to slow down disease spread stronger than average degree speeds it up.

Note that the relation between network size, network density, and average degree is an artifact of the CM. Although implications of degree scaling with network size become more unrealistic the larger the networks get, it illustrates that dynamics can significantly change with network size. Furthermore, it shows the importance of choosing appropriate network formation models for specific model cases.

2.7.4 Interaction effects

The *Interaction effects* regression models help to understand the interdependence between disease and network dynamics. They were selected exploratively based on their effect on improvement of model fit (ℓ) and explained variance (ρ). As these effects are not all intuitive, we take a closer look at two representative interactions that affect both final size and duration of the epidemics: (i) benefit of connections at distance 2 (β) \times cost increase for infected ties (μ), and (ii) disease severity ($\frac{\sigma}{50}$) \times risk perception (r).

As discussed before, increased benefits at distance 2 (β) show positive effects, while cost increases for infected ties (μ) show negative effects on final size and duration of epidemics. This is shown again in Figure 2.3. The bottom left plot shows that for parameter combinations with standard care costs and high social benefits for connections at distance 2, epidemics are typically large ($Mdn = 100\%$) with long duration ($Mdn = 22$ time steps). Parameter combinations that differ from the bottom left in one parameter only (increased care costs for infected ties, bottom right; low social benefit, top left), show similar final sizes and average duration. Parameter combinations with both increased care costs for infected ties and low social benefits (top right), however, show significantly smaller final sizes ($Mdn = 40\%$) and shorter duration of epidemics ($Mdn = 17$ time steps).

The average degrees of the networks (gray lines) show that highly valued connections at distance 2 (bottom row) create low dissolution of ties independent of whether the sickness requires more care for an infected connection. In contrast, lower valued connections remain only intact when care for the sick requires no extra costs (top left). Consequently, parameter combinations with low social benefits and increased costs for infected ties (top right) cause many agents to disconnect and thus successfully distance themselves from the disease once it enters the network.

A similar picture is drawn for disease severity (σ) and risk perception (r), where both parameters show negative main effects on final size and duration of epidemics individually. Again, these effects are stronger in combination than individually (Figure 2.4). Parameter combinations with high risk populations and varied disease severity (top row) show equal average final sizes ($Mdn = 100\%$) and similar duration of epidemics ($Mdn = [21, 22]$ time steps). The same applies for mild diseases and various risk perceptions (left column; final size: $Mdn = 100\%$; duration: $Mdn = [21, 22]$ time steps). Parameter combinations with low risk populations and highly severe diseases (bottom right), however, have the lowest final size ($Mdn = 30\%$) and duration of epidemics ($Mdn = 15$ time steps). Again, the decisive factor for these results is average network degree: Agents disconnect more if

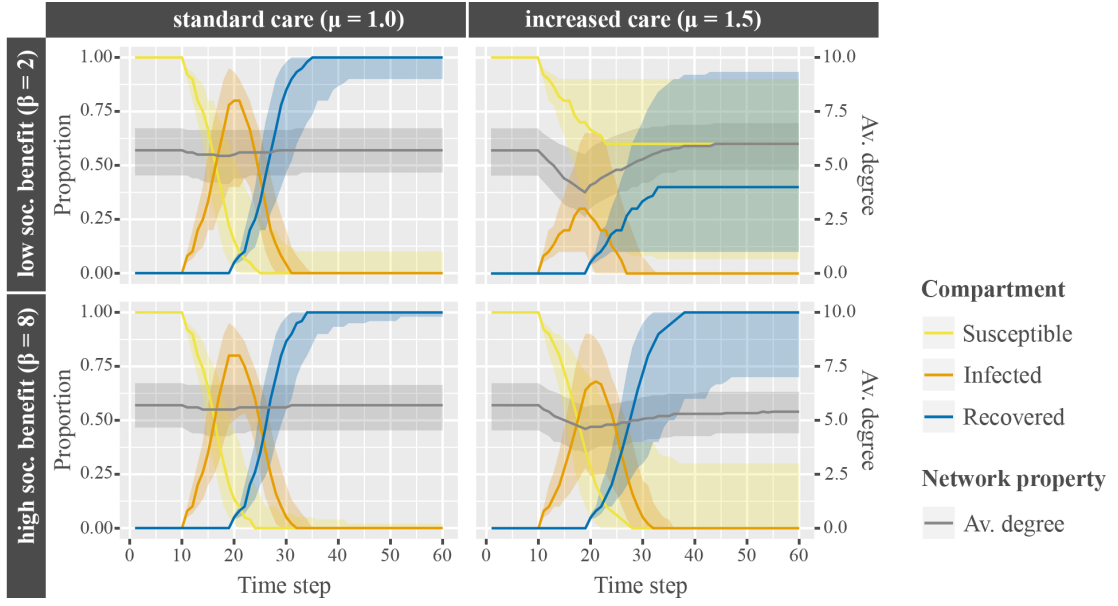


Figure 2.3: **Interaction of costs for infected ties and value of indirect connections.** Interaction of increased costs for infected ties (μ , columns) and value of indirect connections (β , rows) on the median proportional changes of agents according to disease states (yellow, orange, and blue lines), changes in median average network degree (gray lines), and interquartile ranges (shaded areas around median lines) over time.

they perceive risk to be high *and* diseases are severe (bottom right) compared to parameter combinations with agents perceiving risk to be low (top row) *or* mild diseases (left column) alone.

Both depicted interactions demonstrate another highly important effect: an underlying two-way interaction between the social network and the disease. First, the higher the average degree the more serious the epidemic. Second, the more serious the (subjective) consequences of a disease (here: increased costs for infected ties or perceived severity) the stronger the behavioral reaction and thus the lower the average degree.

Further, Figure 2.4 allows looking at the effect of more subtle changes of tie dissolution. Consider the effect of average degree on final size of the epidemic in parameter combinations with highly severe diseases and differing risk perceptions. First, populations with agents perceiving the risks realistically (center right), and second, populations with agents perceiving high risks of infections (bottom center). We see that small changes of average degree (from $Mdn = 4.53$ to $Mdn = 4.13$) create large differences in final size (from $Mdn = 95\%$ to $Mdn = 30\%$), peak height (from $Mdn = 55\%$ to $Mdn = 25\%$), and duration of epidemics (from $Mdn = 31$ to $Mdn = 25$). This enables two important insights: First, epidemics are highly sensitive to social distancing in dynamic social networks. Second, even small changes on the individual level may have major consequences on the large scale.

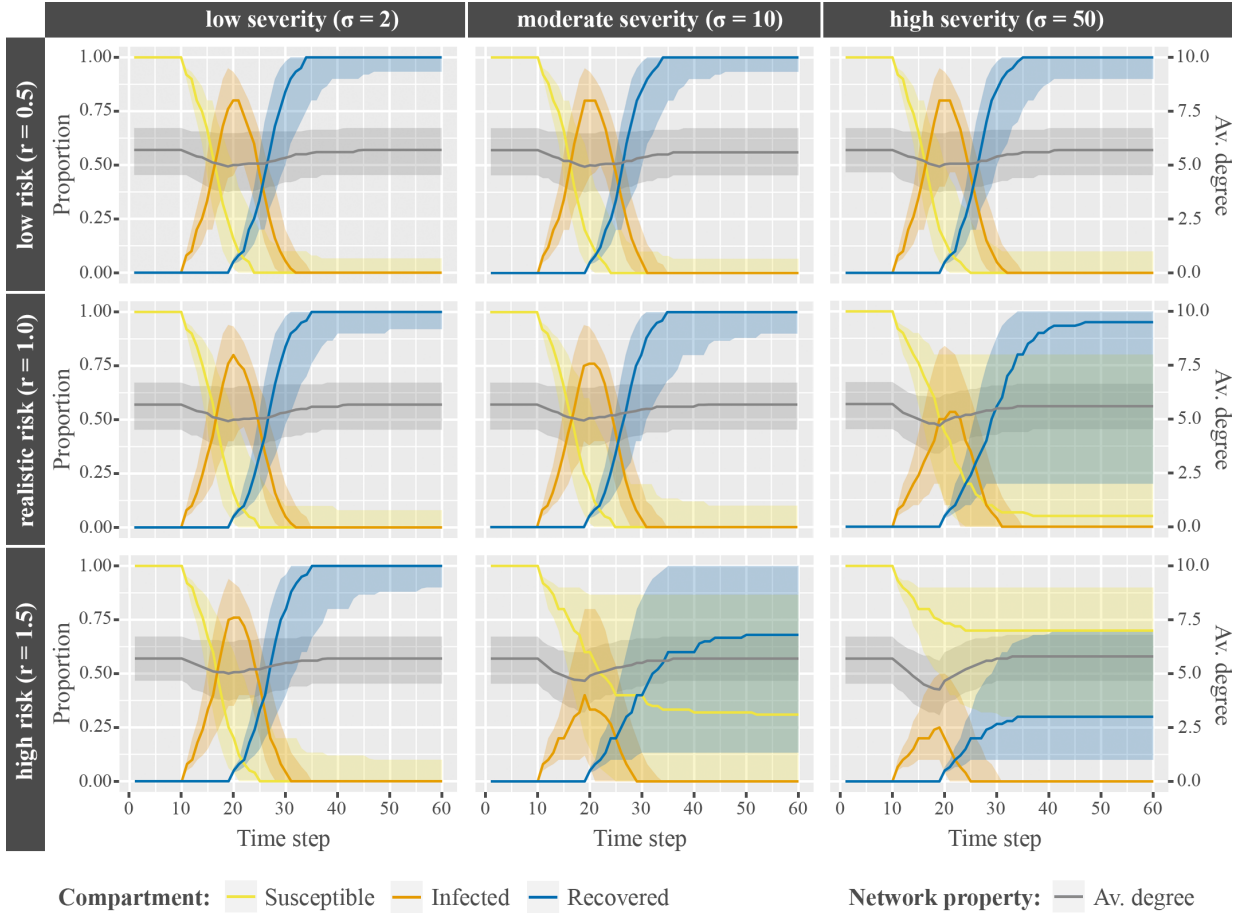


Figure 2.4: **Interaction of disease severity and risk perception.** Interaction of disease severity (σ , columns) and risk perception (r , rows) on the median proportional changes of agents according to disease states (yellow, orange, and blue lines), changes in median average network degree (gray lines), and interquartile ranges (shaded areas around median lines) over time.

2.8 Conclusion and implications

The role of social networks in the spread of infectious diseases has gained more attention in recent years. Networks are commonly considered static or network dynamics are modeled as random processes devoid of health behavioral foundations (see, Leung et al., 2018; Mao & Yang, 2012). We know, however, that people deliberately seek to distance themselves from diseases in social networks (Ahmed et al., 2018; Ferguson et al., 2006; Funk et al., 2010), or are instructed to distance themselves from others at times of increased risk of infection (Courtemanche et al., 2020; Fong et al., 2020; Glass et al., 2006).

To address this, we propose the *Networking during Infectious Diseases Model (NIDM)*, a highly adaptable steppingstone model for the co-evolution of social networks and infectious diseases. This model combines theories of social network formation from sociology, infectious diseases from epidemiology, and individual health behavior from social psychology. We argue that networking behavior in the context of infectious diseases is a trade-off between the social well-being a contact creates, the costs required to maintain a social

tie, and the potential physical harm infectious contacts create. Furthermore, we created a specific model case of the NIDM based on the truncated version of the *Connections Model* (CM; Jackson, 2008; Jackson & Wolinsky, 1996), a model typically describing social contexts of limited group size, including a multi-agent simulation in which agents deliberately distance themselves from infectious contacts in social networks whenever the potential harm of getting infected through a tie outweighs its social benefit.

Based on our simulations, we gain a number of theoretical insights: We find average degree as a major determinant for epidemic size and duration of epidemics. That is, larger numbers of connections provide more opportunities for disease spread, thus causing more agents to get infected and shorter epidemics. Furthermore, we see a highly interdependent process between the properties of agents, diseases, networks, and epidemics: the higher the (perceived) risks of a disease, the lower the net benefit of a tie, the stronger the social distancing, and consequently the lower the epidemic size. While these results align with our expectations, we gained more surprising insights thanks to our novel network approach. First, networks with benefits for social connections that induce many ties per agent, provide large numbers of potential transmission routes, and consequently allow the disease to travel faster. Second, higher costs of maintaining ties with infected others reduce final size of epidemics only when benefits of indirect ties are relatively low. Third, we find that small changes in social behavior have large effects on the epidemic, which may be decisive for whether the outbreak of a disease turns into an epidemic or not.

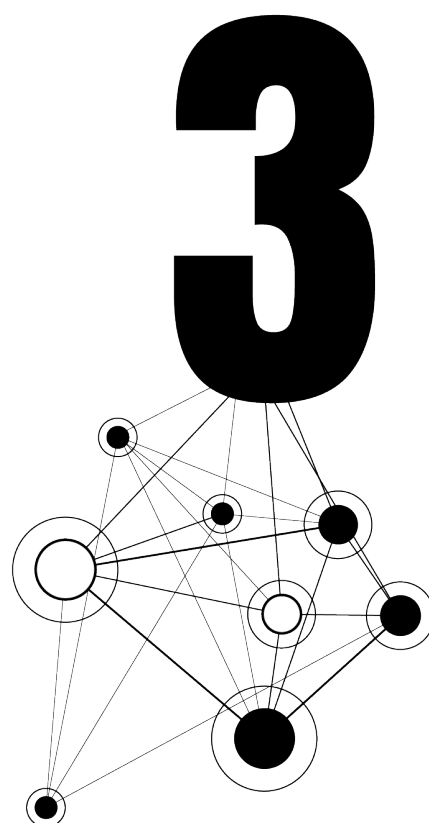
Our theoretical outcomes are in line with empirical observations and other model studies, inter alia, from Ahmed et al. (2018) and Danon et al. (2011). The former is a review of empirical studies that showed reduction in the number of infections and lower epidemic peaks due to workplace social distancing. The latter is an epidemic network model considering the social structure of populations suggesting that high average degrees result in increases of final size of epidemics. Similar to Poletti et al. (2012), our results suggest that even small reductions in contact numbers can reduce epidemic impact. While in line with the results of Leung et al. (2018) (strong stochastic rewiring creates small epidemic sizes, moderate stochastic rewiring creates large epidemic sizes), our model can provide a more explicit theoretical mechanism for their findings. More precisely, our model provides a tool to improve models of infectious diseases by adding network formation processes based on theories from sociology and health psychology, rather than using purely stochastic rewiring processes or artificially lowering transmissibility (Ferguson et al., 2006; Leung et al., 2018; Mao & Yang, 2012). Additionally, our model has proven to be useful to reveal previously unknown and less intuitive mechanisms (e.g., higher costs of maintaining ties with infected others reduce final size of epidemics only when benefits of indirect ties are relatively low). Furthermore, we contribute to solving the issue that neglect of health behaviors may create mismatches between observed transmissibility of diseases and epidemic sizes of model predictions (Mao & Yang, 2012). Although some of our theoretical insight could help to understand earlier empirical findings, the empirical validity of our new predictions needs tailored empirical testing in subsequent studies.

It is most important to stress that our general model, the NIDM is a highly adaptive steppingstone to capture the basic characteristics of the co-evolution of dynamic social networks and infectious disease dynamics, which allows many opportunities for extensions and specifications in each part of the model. First, social utilities, and thus resulting network structures, can be modeled with regard to context (social, disease, cultural). HIV, for example, is typically modeled using sexual contact networks (benefits for direct contacts only) that allow great control over single contacts (large ϕ parameter). Further, social maintenance costs (C_i) can be extended with moral costs for infected individuals to infect others. Second, diseases can be defined according to their objective risks (lethality, stigma, financial), disease states, and treatments. As there is no cure yet for HIV, infected individuals do not recover but remain infected until death. Adoption of preventive behaviors (e.g., intake of PrEP) lowers or may even eliminate the probability of transmission. Third, subjective perception of health risks (r_π , r_σ) can differ according to the availability of information or risk perceptions, an effect utilized by awareness-raising campaigns.

It is furthermore important to stress that any specific model case of the NIDM will come with limitations depending on the modeling choices made. While our current model case, for example, the *Connections during Infectious Diseases Model (CIDM)*, provides an interesting interplay of social and infectious disease dynamics, the CM is designed and typically used to describe small networks. Our simulations suggest that the CM is not feasible for network sizes of 50 and above, as average degree grows with network size ultimately causing all agents to get infected. One approach to realize larger plausible social networks could be the extension of the CM with an explicit saturation term for the number of social relations. Another approach is to design a different baseline utility to describe the utility of social relations for a specific context. Especially the choice of social contexts need to be considered carefully. In friendship networks, for example, we would expect to observe triadic closure, and thus increasing network density and clustering, when indirect contacts are beneficial. Thus, professional networks, as studied by Ahmed et al. (2018), seem more plausible as an application for the CM as baseline utility, as indirect contacts may provide benefit through access to rare information unavailable through direct contacts, without the same likelihood for closing the triad. Further, while it seems plausible that agents assess health risk on local information, the assumption that agents have complete knowledge about the contacts of potential new network partners is not quite realistic. Still, we think that in a realistic process of network formation, people investigate into relations before establishing certain ties. Our assumption can also be interpreted as a simplification of this process, assuming that people indeed are able to obtain this information if they want to establish a relation. Alternatively, one could investigate alternative assumptions, such as random creation of new ties followed by dissolution if the relation is not satisfactory. Furthermore, due to low value and variance of network clustering in the CIDM, we cannot make inferences on how networks with dynamic clusters affect the course of epidemics. Empirically, we know that social networks are typically highly clustered (a property that hardly emerged in our networks) while having small average path lengths (*small worlds*) and that these properties can facilitate disease spread (Watts, 1999; Watts & Strogatz, 1998).

Furthermore, people differ in how they perceive risks and translate this into action. It follows, that an outbreak in a cluster of individuals perceiving risks to be high may prevent an epidemic, while a few low risk individuals in a high risk population may accelerate disease spread.

In conclusion, our theoretical insights suggest that modeling health behavior explicitly in network models of disease spread can help to gain a broader theoretical foundation and deeper insights into the co-evolutionary processes of dynamic social networks and infectious diseases. Furthermore, the steppingstone model we propose — the *Networking during Infectious Diseases Model (NIDM)* — has proven to be a valuable tool for such an undertaking.



Chapter 3

Health behavior homophily can mitigate the spread of infectious diseases in small-world networks

Abstract *Research has repeatedly shown that the spread of infectious diseases is influenced by properties of our social networks. Small-world like structures with densely connected clusters bridged by only a few connections, for example, are not only known to diminish disease spread, but also to increase the chance for a disease to spread to any part of the network. Clusters composed of individuals who show similar reactions to avoid infections (health behavior homophily), however, might change the effect of such clusters on disease spread. To study the combined effect of health behavior homophily and small-world network properties on disease spread, we extend a previously developed ego-centered network formation model and agent-based simulation. Based on more than 80,000 simulated epidemics on generated networks varying in clustering and homophily, as well as diseases varying in severity and infectivity, we predict that the existence of health behavior homophilous clusters reduce the number of infections, lower peak size, and flatten the curve of active cases. That is because agents perceiving higher risks of infections can protect their cluster from infections comparatively quickly by severing only a few bridging ties. A comparison with epidemics in static network structures shows that the incapability to act upon risk perceptions and the low connectivity between clusters in static networks lead to diametrically opposed effects with comparatively large epidemics and prolonged epidemics. These findings suggest that micro-level behavioral adaptation to health risks mitigate macro-level disease spread to an extent that is not captured by static network models of disease spread. Furthermore, this mechanism can be used to design information campaigns targeting proxies for groups with lower risk perception.*

This chapter has been published as Nunner, H., Buskens, V., Teslya, A., & Kretzschmar, M. (2022a). Health behavior homophily can mitigate the spread of infectious diseases in small-world networks. *Social Science & Medicine*, 312, 115350. Nunner and Buskens developed the mathematical model. Kretzschmar and Teslya contributed to its conceptualization. Nunner programmed the simulation, performed analyses, created visualizations, took care of data curation, wrote the original draft of the manuscript, and integrated the feedback of all co-authors. All authors reviewed and discussed the manuscript repeatedly.

3.1 Introduction

The COVID-19 pandemic has revealed the vulnerability of our globalized world to the spread of infectious diseases. As a reaction to this, many countries implemented physical distancing measures to lower the number of infections for a disease mainly transmitted through human-to-human contact in social networks (Li et al., 2020; Shen et al., 2020; World Health Organization, 2021). In other words, reducing social contacts should reduce the transmission of the virus.

However, authority-initiated interventions are not the only factors affecting social network structure to lower the risk of infections. Avoidant health behaviors, such as keeping away from large crowds, avoiding symptomatic coworkers, or avoiding sexual contact because a partner has a sexually transmitted disease, are common self-imposed measures known to mitigate disease spread (Bish & Michie, 2010; Funk et al., 2009). These health behaviors depend on heterogeneous risk perceptions (Bish & Michie, 2010; Leppin & Aro, 2009) that, depending on the composition of a group, may facilitate or hinder the spread of infectious diseases on the macro-level, such as COVID-19 (d’Andrea et al., 2022) or HIV (Koku & Felsher, 2020). Furthermore, risk perceptions depend on personal properties, such as age, gender, ethnicity, educational level, or marital status (Bish & Michie, 2010); properties known to facilitate *homophily* (McPherson et al., 2001), a tendency of people to be connected to others with similar characteristics.

Additionally, certain network properties are known to affect the spread of diseases in networks. On the one hand, small average path length reduces the probability of a disease to die out before reaching distant points of the network (Watts & Strogatz, 1998). On the other hand, the presence of clusters, densely connected areas within a network that have only few ties bridging to nodes outside this area, can reduce the size of an epidemic (Badham & Stocker, 2010; Keeling, 1999; Miller, 2009). These two features are of particular interest for the spread of infectious diseases, as they constitute the typical structure of our *small-world*-like social networks (Milgram, 1967; Watts & Strogatz, 1998).

Although we know quite well how small-world properties and homophily affect disease spread in networks separately, it is less clear how the two shape epidemics in combination. Furthermore, network models of disease spread typically neglect adaptive changes in the network structure, despite their potentially large effect on predictions (Bansal et al., 2010). We therefore ask: *What are the effects of health behavior homophily on epidemics in adaptive small-world networks?* and *How do epidemics differ between adaptive and static networks?*

3.1.1 Network models of disease spread

For almost 40 years, scientists use network theory to study the spread of infectious diseases (e.g., Klov Dahl, 1985; May & Anderson, 1987). A typical approach is to study how the structural properties of relationships (edges) between individuals (nodes) affect the spread of infectious diseases.

Many network models of disease spread, however, assume that relationships do not change over time (Bansal et al., 2010). In their review on epidemic processes in complex networks, Pastor-Satorras et al. (2015) argue that static networks are good approximations, when the network evolves at a much slower pace than the dynamic process working within it. Examples for processes that change network topology slowly are demographic changes, such as births and deaths, migration, or social or economic changes. Short-term changes, however, can be the result of behavioral changes to adapt to health risks thus requiring the consideration of the interdependent dynamics between social networks and infectious diseases (Bansal et al., 2010). An established approach to account for such adaptations as well as heterogeneity on the host level are individual-based models for infectious disease transmission (Bedson et al., 2021; Verelst et al., 2016; Willem et al., 2017). Simulation studies on adaptive networks suggest, for example, that epidemic size is smaller in adaptive networks compared to their static counterparts, no matter whether ties to infectious nodes are simply cut (Gross et al., 2006) or whether infectious nodes rewire randomly (Leung et al., 2018; Risau-Gusman & Zanette, 2009). Another typical application of individual-based models is to test the effectivity of network-based interventions (Verelst et al., 2016), such as vaccination (Brebán, 2011; Nunner et al., 2022b) or information campaigns (Mao & Yang, 2012; Nyabadza et al., 2010).

3.1.2 Small-world networks

“My it’s a small world.” is according to Milgram (1967, p. 61) a typical statement we make upon realizing that a stranger knows the same person we do. In a seminal study, he discovered that it takes only about five intermediaries to send a letter from distant places in the US (i.e., Wichita, Kansas and Omaha, Nebraska) to a stockbroker in Boston, Massachusetts. This surprisingly low number is caused by the typical properties of our social networks: high clustering and short average path length (Watts & Strogatz, 1998). That is, small-world networks typically have densely connected areas that are connected by just a few bridges. Translated into our social lives that means we form clusters with people who are also connected among each other (e.g., friends, family), while we typically know few people who are not part of that inner social circle (e.g., a former fellow student). The tendency for high clustering or *triadic closure* (Simmel, 1950) is often studied in the social sciences (Granovetter, 1973), also because social cohesive structures have advantages in terms of cooperation and trust for the people in these structures (Coleman, 1994).

Despite the advantages of social cohesion, we typically do not find fully connected networks. In addition to obvious reasons (spatial distance, too many people, personal sympathies), Jackson (2008) presents a process of consideration preceding the formation of contacts. He argues that while social ties provide benefits (e.g., emotional support, sense of belonging), there are also costs to maintain the relationships (e.g., time, effort). Thus, the answer to whether a social tie is formed or kept is a trade-off between the benefits from having the tie and the costs to maintain it. On the one hand, this trade-off creates tightly-knit clusters within our social networks that are limited in size. On the other hand, clusters are not disconnected from each other, but bridged by a few ties. Burt (1992) argues

that bridging ties provide their own advantages such as access to and control of resources and information. Consequently, network properties, such as the number of relations, the degree of clustering, or average path length, depend on costs and benefits shaped by social context.

Previous research on disease spread in social networks has shown repeatedly that increased clustering leads to smaller epidemics (Badham & Stocker, 2010; Keeling, 1999; Miller, 2009). Small-world networks that combine high clustering with short average path lengths, however, facilitate the spread of infectious diseases (Watts & Strogatz, 1998). That is because transmission events are subject to probabilities, and the more transmissions are required to reach distant parts of the network, the more likely the transmission chain breaks. Ties bridging two clusters lower the average path length by providing shortcuts to access distant nodes with only a few steps. It follows that while clusters hamper the spread of infections, increasing numbers of bridges between clusters facilitate the disease to quickly spread through the entire network.

3.1.3 Health behavior homophily

In the context of infectious diseases, Kasl and Cobb (1966) define health behavior as “[...] any activity undertaken by a person believing himself to be healthy, for the purpose of preventing disease or detecting it in an asymptomatic stage” (p. 531). A systematic review of 26 studies on a variety of airborne diseases (SARS, avian influenza, H5N1, swine flu, H1N1) reveals avoidant behavior as one of the most commonly adapted behaviors to prevent disease (Bish & Michie, 2010). Typical avoidant behaviors are removing social ties (Funk et al., 2010), voluntary quarantining (Tracy et al., 2009), or avoiding public places (Jones & Salathé, 2009).

A systematic review of 28 studies on SARS and avian influenza (Leppin & Aro, 2009) showed that independent of how risk is conceptualized, the perception of risks is the most important determinant of health decisions. The two main drivers of risk perception are (i) how probable a person believes it is to get infected, and (ii) how severe that person thinks the disease is (Bish & Michie, 2010; Leppin & Aro, 2009). Risk perceptions, however, are highly subjective in nature (Bults et al., 2015), meaning that the more people perceive themselves to be susceptible to a disease and the more severe the consequences of the disease are perceived, the more likely they are to change behavior to save themselves from a potential infection. It follows, that a person who perceives a high risk of catching a disease with presumably severe effects on personal health is more likely to avoid potentially infectious contacts than a person who does not.

Although risk perception is an individual characteristic and the resulting behavioral changes occur on the individual level, homophily regarding risk perception exerts effects beyond that. Homophily indicates the extent to which it is more likely for a social tie to exist between similar people than between dissimilar people. Causes for homophily can be either the demographic structure of a group (*baseline* homophily) or explicit preferences for similarity (*inbreeding* homophily; McPherson et al., 2001). Independent of what causes homophily to emerge, risk perceptions and health behavior often correlate with other prop-

erties facilitating homophily, such as age, gender, ethnicity, educational level, or marital status (Bish & Michie, 2010). McPherson et al. (2001) argue that as a result of homophily transmission processes tend to be socially localized. Kitchovitch and Li (2010), for example, showed that higher levels of risk avoidance in scale-free networks limits the disease to the area around highly connected nodes, while nodes with lower contact numbers remain unaffected.

3.2 The model

Although numerous studies were concerned with the effects of small-world properties (Badham & Stocker, 2010; Keeling, 1999; Miller, 2009; Watts & Strogatz, 1998) or homophily (d’Andrea et al., 2022; Kitchovitch & Li, 2010; Koku & Felsher, 2020) on epidemics, there is, to our knowledge, no study that looks into their interplay. Furthermore, simulation studies typically consider static networks (Badham & Stocker, 2010; d’Andrea et al., 2022; Keeling, 1999; Kitchovitch & Li, 2010; Miller, 2009; Watts & Strogatz, 1998). To study the combined effects of small-world properties and homophily on epidemics in adaptive social networks, we created the so-called *Small-Worlds Infectious Disease Model (SWIDM)*, a specific model case of the *Networking during Infectious Diseases Model (NIDM)* (Nunner et al., 2021).

To facilitate comparability, reproducibility, and understanding of our model we are guided by the updated version of the *ODD protocol* (“Overview”, “Design concepts”, “Details”) of individual- and agent-based models (Grimm et al., 2010). Here, we offer mostly top-level descriptions of the model components relevant for this paper. In Section B.1 of the Supplements to Chapter 3, we offer more detailed and formal model definitions.

3.2.1 Purpose

The SWIDM was developed to investigate how (i) social mixing of heterogeneous risk perception (random, homophilous), (ii) variations in small-world network properties (clustering, path length), and (iii) the possibility of network changes (static, adaptive) affect the transmission of infectious diseases in social networks.

3.2.2 Basic principles

The SWIDM is a network formation and disease spread model for a closed population of autonomous agents (nodes) that can create infection-relevant relationships (edges). Networking decisions are based on the trade-off between the social benefits (e.g., sense of belonging, affection), costs (e.g., time, effort, money), and (potential) harm of infections (e.g., symptoms, absence from work, hospitalization) that a tie has for an agent. Agents seek to myopically maximize personal utility by (i) forming ties to other agents (requires the consent of the opposite agent), (ii) maintaining existing ties to other agents, and (iii) dis-

solving ties to other agents (unilateral decision). Decisions on these actions are based on an agent's position in the network, individual risk perception related to the disease, and the distribution of disease states among agents in the network (see Section 3.2.4).

In their choice to create or maintain ties, agents have a preference for triadic closure versus openness in their networks. This preference balances two things: first, a preference for closed triads (Coleman, 1994), and second, the strategic advantage to bridge structural holes (Burt, 1992). Finally, the underlying agent selection process captures baseline homophily (McPherson et al., 2001) by implementing larger probabilities for changing ties with agents who have similar risk perceptions. This process does not integrate explicit preferences regarding risk perception of alters (inbreeding homophily) and is in line with findings that correlations between personal properties, such as risk perceptions and health behavior, are driven by higher likelihoods to meet similar others (e.g., regarding age, gender, ethnicity, educational level, marital status) (Bish & Michie, 2010; McPherson et al., 2001).

3.2.3 Dynamics of disease transmission and network formation

Disease dynamics and adaptive network dynamics are simulated in discrete time steps (for detailed pseudocode refer to Section B.1.2 of the Supplements to Chapter 3). At the start of each time step, the simulation updates the epidemiological state of the population. That is, for each agent in random order, we first determine its disease state. If an agent is susceptible, the likelihood of the agents to become infected is:

$$\pi_i = 1 - (1 - \gamma)^{t_{i_I}}, \quad (3.1)$$

with γ being the probability to get infected per single contact and t_{i_I} the number of infected neighbors. If an agent is infected and has been infected for a fixed number of time steps (τ), the agent recovers. If an agent is recovered, the agent cannot get re-infected until the simulation run ends. Note that the random order ensures that all nodes get an equal chance to be processed first, as later processed agents may have an increased number of infectious neighbors than if they had been processed earlier.

Following the update of the population's epidemiological state, social network dynamics are computed. That is, for each agent i in random order a set J consisting of a fixed proportion (ϕ) of agents from the entire population (N) is selected. To realize the concept of *focused* interaction in an agent's social vicinity (Feld, 1981), J is composed of neighbors at distance 1 and distance 2, as well as agents randomly selected from the entire network. Furthermore, to account for homophily, each single agent j is selected either to be the most similar agent regarding risk perception (r) or an agent selected without regard to risk perception. More formally, we draw a random number u from the uniform distribution $U(0, 1)$ and compare this to the homogeneity parameter ω such that:

$$j = k, \text{ with } k \begin{cases} \text{such that } |r_i - r_k| = \min_{\forall k' \in N^*} |r_i - r_{k'}|, & \text{if } u \leq \omega \\ \text{is a randomly selected element of } N^*, & \text{if } u > \omega \end{cases} \quad (3.2)$$

For each agent j in J , i dissolves an existing tie, if i 's utility (see Section 3.2.4) without tie ij exceeds the utility with tie ij ; or i proposes a new tie to j , if i 's utility with tie ij exceeds the utility without ij . Note that only if j accepts the proposal (j 's utility with tie ij also exceeds the utility without ij), a new tie ij is formed.

3.2.4 Utility

Utility for an agent i is defined in the SWIDM as the trade-off between the social benefits (B_i), social maintenance costs (C_i), and (potential) harm (D_i) of infections based on the network connections held by i :

$$U_i = B_i - C_i - D_i. \quad (3.3)$$

Social benefits are defined as the weighted sum of the benefits for ties (left summand) and the benefits for the proportion of closed triads (right summand):

$$B_i = b_1 \cdot t_i + b_2 \cdot \left(1 - 2 \cdot \frac{|x_i - \alpha|}{\max(\alpha, 1 - \alpha)} \right), \quad (3.4)$$

with x_i denoting the actual proportion of closed triads i belongs to, α the preferred proportion of closed triads, and t_i the number of ties agent i possesses. Social maintenance costs are assumed to be quadratic in the number of ties t_i to model increasing marginal costs of additional ties:

$$C_i = c_1 \cdot t_i + c_2 \cdot t_i^2. \quad (3.5)$$

The combination of benefits and costs for social ties ($[b_1 \cdot t_i] - [c_1 \cdot t_i + c_2 \cdot t_i^2]$) allows us to control the number of ties the agents seek to establish apart from clustering considerations. This largely controls the degrees of the agents and thus the average degree of the network. Further, we control the degree of clustering in the network by defining the proportion of closed triads agents seek to be a part of (α).

(Potential) harm of infections is a combination of *perceived* probability to get infected (p_i) and *perceived* severity of the disease (s_i):

$$D_i = p_i \cdot s_i. \quad (3.6)$$

That is, agents transform actual probability to get infected π_i (Equation 3.1) into a subjective version of the same, depending on their disease state:

$$p_i = \begin{cases} \pi_i^{2-r}, & \text{if } i \text{ is susceptible,} \\ 1, & \text{if } i \text{ is infected,} \\ 0, & \text{if } i \text{ is recovered.} \end{cases} \quad (3.7)$$

Furthermore, agents transform actual severity of the disease (σ) into a subjective version of the same, depending on their disease state:

$$s_i = \begin{cases} \sigma^r, & \text{if } i \text{ is susceptible,} \\ \sigma, & \text{if } i \text{ is infected,} \\ 0, & \text{if } i \text{ is recovered.} \end{cases} \quad (3.8)$$

Consequently, agents with risk perception values (r) below 1 underestimate, while agents with risk perception values above 1 overestimate the probability of infections and disease severity. We can therefore control the severity (σ) and infectivity (γ , operationalized as transmission probability per contact and time step) of the disease, as well as the agents' perception of these parameters (r).

3.3 Simulation

3.3.1 Simulation procedure and output data

A simulation run is initialized by setting the model parameters (see Table 3.1). Thereupon, we simulate the social network dynamics in a disease free population until a *pairwise stable* network (Jackson & Wolinsky, 1996) emerges. That is a notion of equilibrium where no agent benefits from unilaterally breaking an existing tie and no pair of agents benefits from jointly creating a non-existing tie. This ensures that changes in network structure at the onset of an epidemic depend solely on the presence of an infectious disease. If this initialization phase did not lead to a pairwise stable and connected network within 50 time steps, we restarted the initialization until we had over 80,000 networks suitable for the next phase (80,305 to be precise).

The observation phase begins when a randomly selected agent gets infected (*index case*). The observation phase ends when no infected agents are left (every agent is either susceptible or recovered). Each observation phase is executed twice for each network and parameter setting. That is, one simulation without network dynamics and one simulation with network dynamics. Thus, both conditions start with the same network structure and index case. The only difference is that agents in the static condition cannot change their network ties, while agents in the adaptive condition can.

During the observation phase we keep track of all varied parameters and a range of outcome variables (see Table 3.1). On the agent level, we record the number of social ties (t_i), the proportion of closed triads (x_i), and the disease state (d_i). On the network level, we record the number of susceptible ($|\mathbf{S}|$), infected ($|\mathbf{I}|$), and recovered agents ($|\mathbf{R}|$), the number of broken (t_G^-) and created ties (t_G^+), network clustering (\mathcal{C}_G), average path length (\mathcal{L}_G), and homophily (\mathcal{H}_G). Furthermore, we keep track of average degree (\mathcal{D}_G), as it is known to be a significant factor for diffusion processes even for small variations.

Table 3.1: State variables, scales, and settings for model parameters.

State variable	Scale	Setting
<i>Agent</i>		
<i>I. Parameters</i>		
Benefit per social tie*	$b_1 \in \mathbb{R}_0^+$	$b_1 = 1.0$
Benefit for triadic closure	$b_2 \in \mathbb{R}_0^+$	$b_2 = 0.5$
Preferred proportion of closed triads [†]	$0 \leq \alpha \leq 1$	$\alpha \sim U[0, 1]$
Simple cost per tie*	$c_1 \in \mathbb{R}_0^+$	$c_1 = 0.2$
Marginal cost per tie*	$c_2 \in \mathbb{R}_0^+$	$c_2 = 0.05$
Risk perception	$0 \leq r \leq 2$	$r \sim U[r_{min}, r_{max}]$
<i>II. Outcomes</i>		
Number of social ties	$t_i \in \mathbb{N}_0$	
Proportion closed triads	$0 \leq x \leq 1$	
Disease state	$d \in \{S, I, R\}$	
<i>Network</i>		
<i>I. Parameters</i>		
Number of agents	$N \in \mathbb{N}_0$	$N = 80$
Minimum risk perception	$0 \leq r_{min} < 2$	$r_{min} \sim U[0, 1]$
Maximum risk perception	$0 < r_{max} \leq 2$	$r_{max} \sim U[1, 2]$
Likelihood of ties similar in risk perception [‡]	$0 \leq \omega \leq 1$	$\omega \sim U[0, 1]$
<i>II. Outcomes</i>		
Number of susceptible agents	$ \mathbf{S} \in \mathbb{N}_0$	
Number of infected agents	$ \mathbf{I} \in \mathbb{N}_0$	
Number of recovered agents	$ \mathbf{R} \in \mathbb{N}_0$	
Number of broken ties	$t_G^- \in \mathbb{N}_0$	
Number of created ties	$t_G^+ \in \mathbb{N}_0$	
Clustering	$0 \leq \mathcal{C}_G \leq 1$	
Average path length	$\mathcal{L}_G \in \mathbb{R}_0^+$	
Homophily	$-1 \leq \mathcal{H}_G \leq 1$	
Average degree	$\mathcal{D}_G \in \mathbb{R}_0^+$	
<i>Infectious disease, parameters</i>		
Disease severity	$\sigma > 1$	$\sigma \sim U(1, 100]$
Infectivity [§]	$0 \leq \gamma \leq 1$	$\gamma \sim U[0.01, 0.20]$
Recovery time in time steps	$\tau > 0$	$\tau = 5$

Notes: *: The combination of $b_1 = 1.0$, $c_1 = 0.2$, and $c_2 = 0.05$ sets the preferred number of ties to 8. [†]: The effect of α on network clustering and average path length is depicted in Figure B.1 in the Supplements to Chapter 3. [‡]: The effect of ω on homophily is depicted in Figure B.1 in the Supplements to Chapter 3. [§]: Infectivity is operationalized as transmission probability per contact and time step.

Network clustering is defined as the average of all agents' local clustering coefficients, while average path length consists of the average of all shortest paths between all pairs of agents in the network (Watts & Strogatz, 1998). Following the concept of quantifying degree-based assortative mixing by Newman (2002), we quantify homophily with the Pearson correlation coefficient of risk perception between all pairs of tied agents.

3.3.2 Parameter settings

We fixed the population size to $N = 80$, the preferred number of ties per agent to 8 (realized through the combination of $b_1 = 1.0$, $c_1 = 0.2$, and $c_2 = 0.05$; see Section 3.2.4 for details), the number of agents to be evaluated per time step to 16 ($\phi = 0.2$), and the recovery time to $\tau = 5$ time steps. These parameters constitute a relevant and interesting framework for our further studies. That is, keeping average degree constant, and thus creating a homogeneous degree distribution, enables us to minimize the otherwise strong effect of degree on epidemic dynamics (e.g., Danon et al., 2011; Nunner et al., 2021), and thus to study the network properties of interest (clustering, average path length) in isolation. Additionally, empirical studies support an average of 8 contacts per day being a relevant magnitude for networks of respiratory disease spread (Danon et al., 2013; Leung et al., 2017), while 5 days recovery time is within the average range for respiratory diseases, like influenza (Longini Jr et al., 2005), and people typically have between 5 and 30 contacts relevant for respiratory disease transmissions per day (Mossong et al., 2008). Finally, we opted for a constant homogeneous setting to minimize sources of variation, while test simulations have shown that these settings produce interesting dynamics.

Submodels are realized through randomized initial settings of the remaining model parameters. This allows us to generate epidemics on a whole variety of different scenarios and to study how sensitive the outcomes are to changes in parameter settings. Variations in clustering and average path length are realized through uniform random samples of preferred proportion of closed triads ($\alpha \sim U[0, 1]$). Note that the joint variation of clustering and path length is key in the collective dynamics of small-world networks (Watts & Strogatz, 1998). Uniform random samples of likelihood of ties between agents that are similar in risk perception ($\omega \sim U[0, 1]$) are used to realize variations in homophily. Further, we set randomized bounds for risk perception on the network level ($r_{min} \sim U[0, 1]$, $r_{max} \sim U[1, 2]$) and assign randomized values within the previously set bounds to each agent ($r_i \sim U[r_{min}, r_{max}]$) to realize variations of risk perception on the network and individual level. Finally, we vary disease severity ($\sigma \sim U(1, 100]$) and infectivity ($\gamma \sim U[0.01, 0.20]$).

Note that, although we choose plausible settings supported by empirical examples, we do not study a specific disease in a specific context but consider parameter variations that contribute to answering our specific research questions. Please refer to Nunner et al. (2021) for detailed elaborations on the model’s theoretical foundations, its functional form, and additional sensitivity analyses.

3.3.3 Data and source code availability

The data, the Java 8 source code to generate the data (including an executable program and an easy to use graphical user interface), and the R scripts to analyze the data during the current study are available under the GPLv3 license in the GitHub repository, <https://github.com/hnunner/nidm-simulation>.

3.4 Analysis

To investigate the effect of small-world properties and health behavior homophily on epidemics, we divide the variables into two categories. First, independent variables describing network and disease properties: (i) network clustering, (ii) average path length, (iii) homophily, (iv) average degree, (v) disease severity, (vi) infectivity, and (vii) average risk perception. Note that the listed network properties are outcome variables resulting from the network formation process, but independent variables in the analysis of epidemics. Second, dependent variables characterizing epidemics and network dynamics: (i) the percentage of infected agents throughout the entire epidemic (*final size*), (ii) the number of time steps from first infection until all agents are either susceptible or recovered (*duration*), (iii) the maximum number of simultaneously infected agents (*epidemic peak size*), and (iv) the number of network changes during the epidemic (number of broken and created ties).

We use box-and-whisker plots to compare dependent variables between adaptive and static networks. Box-and-whisker show the median, interquartile range ($IQR : [Q_1, Q_3]$), minimum (at most $Q_1 - 1.5 \cdot IQR$), maximum (at most $Q_3 + 1.5 \cdot IQR$), and outliers. We use Kendall rank correlation coefficients to describe the relationship between non-normally distributed variables (i.e., final size - duration, final size - peak size).

Bivariate effects are analyzed visually using scatter plots with independent variables on the x-axis and dependent variables on the y-axis. Due to the large number of simulations, we sort the independent variables and bin them into 100 points on the x-axis, thus, showing the mean of $\frac{80,305}{100} \approx 803$ simulations per point. Standard errors and 95% confidence intervals are omitted, as they hardly exceed the size of the dots in every plot.

Finally, we perform regression analyses to study model behavior in a multivariate manner. We realize that significance tests do not have the conventional interpretation because we do not have a usual sample. Tests should be interpreted more descriptively and are used to find the best fitting models given the sample of model parameters drawn (cf. Buskens & Snijders, 2016; Buskens & Yamaguchi, 1999). We use models to linearly approximate the logit of final size by combinations of the independent variables, since final size describes the percentage of a binary response (infected / not infected) (Long, 1997). We use linear models for duration, peak size, and number of network changes. We first create models for each dependent variable and the main effects for all standardized independent variables. Second, we create interaction effects starting from all possible interactions. We then remove all effects that are not significant at $p < 0.001$ or do not contribute to reduce unexplained variance at $R^2 \geq 0.001$. For highly collinear parameters (e.g., clustering and path length) we select parameters to minimize collinearity, while maximizing R^2 . We use an inductive approach rather than testing theoretically informed interaction effects to get a more complete picture of the model behavior and not to miss potentially counterintuitive interactions.

3.5 Results and discussion

In the following, we discuss the results of our simulations starting with a comparison of epidemics in adaptive and static networks. Thereafter, we discuss how health behavior homophily affects epidemics in adaptive and small-world networks and conclude with how these dynamics differ in static networks. Additional analyses can be found in Section B.2 of the Supplements to Chapter 3, where we show, among others, that the effect of degree on final size is negligible (*Effect of degree on final size*) and can therefore be omitted for further analyses.

3.5.1 Epidemics in adaptive vs. static networks

While in the majority of simulated epidemics almost all agents get infected in the static networks ($Mdn = 97.50\%$), only about one fifth of the agents get infected in the adaptive networks ($Mdn = 21.25\%$). Because duration and peak size are correlated with final size (final size - duration: $r_\tau = 0.61$, $p < 0.05$; final size - peak size: $r_\tau = 0.89$, $p < 0.05$), epidemics in static networks are on average longer (static networks: $Mdn = 17$ time steps; adaptive networks: $Mdn = 15$ time steps) and have higher maximum numbers of simultaneously infected agents (static networks: $Mdn = 53.75\%$; adaptive networks: $Mdn = 11.25\%$). That is, few infected agents need on average less time to recover than many infected agents, given they do not get infected at the same time step. This relation, however, reverses for large final sizes in static networks (Fig. 3.1b, orange box). Here, epidemics take significantly less time than for medium final sizes (Fig. 3.1b, green box). In adaptive networks, on the other hand, the duration is largely independent of final size when final size is larger than 10%. Consequently, many infected agents increase the rate of transmission events and thus the speed of disease spread if agents cannot actively avoid infections. Adaptive agents in networks with large final sizes, on the other hand, manage to delay infections, thus slowing down the course of epidemic. As a result, static networks show not only higher epidemic peaks, but must also have more simultaneously infected nodes on average.

3.5.2 Effects of health behavior homophily on epidemics in adaptive small-world networks

Figure 3.2 shows bivariate effects of network properties on epidemics in adaptive networks. The data reveal negative effects on final size for homophily, network clustering, and average path length. Furthermore, final size is mostly stable at high levels in networks with low levels of homophily (< 0.45) and clustering (< 0.45). In contrast, final size drops off quickly for low levels of average path length and remains stable for higher levels of average path length (> 3.0). Interestingly, multivariate regression analyses show that the effect of average path length on final size vanishes when controlled for all other independent variables (Table 3.2).

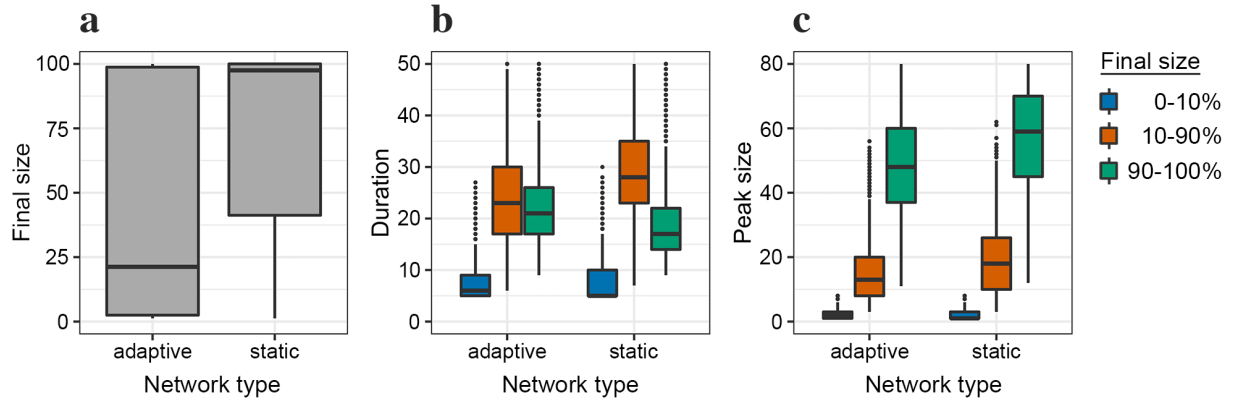


Figure 3.1: **Comparison of epidemics in adaptive and static networks.** Each panel shows box-and-whisker plots divided by network type (left: adaptive, right: static). Panels show the average outcome for final size (a), duration (b), and peak size (c). Plots for duration and peak size are further divided into three groups depending on final size: low (0-10%; blue), medium (10-90%; orange), and high (90-100%; green).

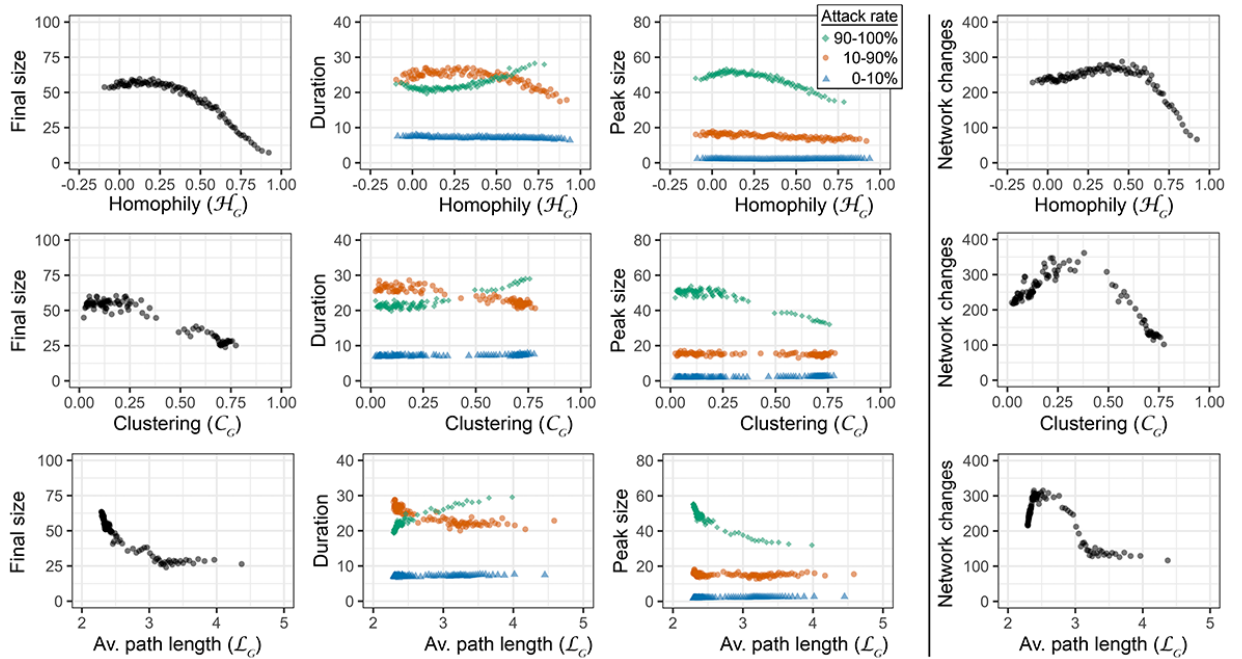


Figure 3.2: **Bivariate effects of network properties on epidemics.** Plots show the effects of homophily (row 1), network clustering (row 2), and average path length (row 3) on final size (column 1), duration (column 2), and peak size (column 3) of epidemics, as well as number of network changes (column 4) in adaptive networks. There are 100 points per group and plot, with each point showing the average of ≈ 803 ($\frac{80,305}{100}$) observations. Measures of variance are omitted, as they merely exceed the size of the dots. Data for duration and peak size have been grouped by final size: 0 – 10% (blue), 10 – 90% (orange), and 90 – 100% (green). Additional effects (average degree, average risk perception, disease severity, infectivity) and effects of all parameters on epidemics in static networks can be found in Figures B.2 and B.3 in the Supplements to Chapter 3.

Table 3.2: Regression analysis summary.

	Adaptive networks				Static Networks		
	Final size	Duration	Peak size	Network changes	Final size	Duration	Peak size
Main effects							
Number of network changes ($t_G^{+/-}$)	\oplus	\oplus	+++				
Homophily (\mathcal{H}_G)	---	-	---	-			-
Clustering (\mathcal{C}_G)	--	-	---	++	-	++++	-
Average path length (\mathcal{L}_G)		++		----	-	++	--
Average risk perception ($r_{\sigma,\pi}$)	-	-	-	-			
Disease severity (σ)	-	-	-				
Infectivity (γ)	++++	--	\oplus	\oplus	\oplus	-	\oplus
Interaction effects							
$t_G^{+/-} \times \mathcal{H}_G$	--		-				
$t_G^{+/-} \times \mathcal{C}_G$	--	--	--				
$t_G^{+/-} \times \mathcal{L}_G$		++++					
$t_G^{+/-} \times \gamma$	++	--	++				
$\mathcal{H}_G \times \mathcal{C}_G$	+	-	+				
$\mathcal{H}_G \times \mathcal{L}_G$				-			
$\mathcal{H}_G \times \gamma$	--		--	+			
$\mathcal{L}_G \times \gamma$			--	-	-	\oplus	-
Adjusted R ²	0.92	0.80	0.85	0.39	0.76	0.07	0.87
Number of observations	80,305	80,305	80,305	80,305	80,305	80,305	80,305

Notes: All variables are standardized. All effects shown are significant at $p < 0.001$. Effect direction is shown as sign (+ for positive effects, - for negative effects), while effect sizes are shown in relation to the largest effect (\oplus/\ominus): $+/- = 0\text{-}20\%$, $++/-- = 20\text{-}40\%$, ..., $+++++/- = 80\text{-}100\%$. Individual regression models can be found in Section B.2.4 of the Supplements to Chapter 3.

The number of network changes helps to understand the complex dynamics in adaptive networks. That is, all independent variables have similar effects on final size and the number of network changes (Fig. 3.2, column 4). It follows that the further the disease spreads in the network, the more agents reposition themselves to avoid or delay infections. If we, however, control for all independent variables (Table 3.2), the number of network changes increases, although final size decreases as network clustering increases. That is because infections inside densely connected clusters require agents to dissolve comparatively many ties to distance themselves from the disease. Due to the low number of ties bridging any two clusters, however, it is less likely that the disease spreads through the entire network.

Furthermore, the interaction effect between homophily and clustering (Fig. 3.3a) reveals that the combination of high clustering and high health behavior homophily produces the lowest final sizes. As the negative effect of risk perception (in combination with disease severity and infectivity) on final size (Table 3.2) suggests, the less risk taking the agents are, the quicker they dissolve their ties to infectious neighbors. Consequently, a cluster of low risk taking agents is especially hard to infiltrate due to the quick dissolution of ties to comparatively few infectious neighbors outside the cluster. Consequently, the average epidemic in clustered networks is also shorter and has a lower peak size (Table 3.2). Considering final sizes of 90% and above (Fig. 3.2a, column 2, green points), however, the effect is inverted: If all agents get infected at some point, the disease requires more time to bridge between loosely connected clusters. Figure 3.3 reveals further that even for highly infectious diseases, higher levels of homophily produce smaller final sizes (b) and larger average path lengths produce lower epidemic peaks (c).

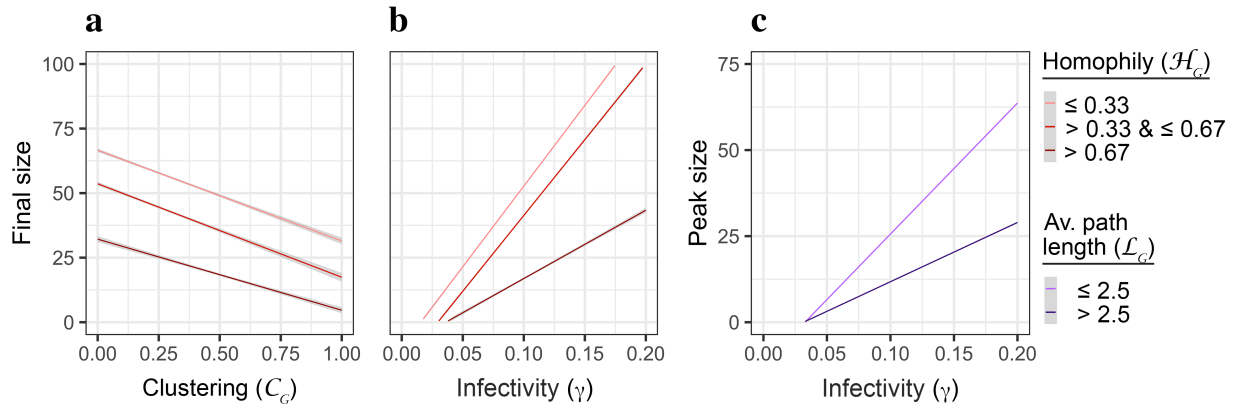


Figure 3.3: **Interaction effects in adaptive networks.** Final size of epidemics decreases the more clustered the networks (a) and the lower the infectivity (b). At the same time, larger degrees of health behavior homophily (darker red lines) create lower numbers of infections. Epidemic peak size increases the higher the infectivity (c) and the shorter the average path length (light purple line).

Finally, we observe that epidemics with final sizes up to 90% were insensitive to variation in every independent variable (Fig. 3.2, column 3). Peak sizes for final sizes of more than 90%, however, follow the shape of the final size curves. That is, only if the disease spreads through almost the entire network, the number of simultaneously infected agents correlates with the final size of the epidemic.

3.5.3 Effects of health behavior homophily on epidemics in static small-world networks

Dynamics in static networks show several significant differences to their adaptive counterparts. First, there is no observable effect of homophily and risk perception on final size in static networks (Table 3.2). This is, of course, due to the operationalization of homophily in our study, which renders agents incapable to act upon their perceptions in static networks. Second, we observe a positive effect of clustering on epidemic duration in static networks (Table 3.2). That is, while the capability to cut off bridging ties quickly in adaptive networks causes epidemics to die out quickly (unless the disease spreads to the entire networks), the incapability to distance oneself from the disease leads to slow spread and on average larger final sizes in static networks. Third, we observe negative effects of average path length on final size and peak size in static networks only (Table 3.2).

3.6 Conclusion and implications

Although the literature contains numerous publications on how diffusion processes are shaped by small-world properties and homophily individually (Badham & Stocker, 2010; d’Andrea et al., 2022; Keeling, 1999; Kitchovitch & Li, 2010; Koku & Felsher, 2020; Miller, 2009; Watts & Strogatz, 1998), it is still unknown how the diffusion of infectious diseases is

shaped by their combination. We therefore asked: *What are the effects of health behavior homophily on epidemics in adaptive small-world networks?* and *How do epidemics differ between adaptive and static networks?*

To answer these questions, we created the so-called *Small-Worlds Infectious Disease Model (SWIDM)*, a specific model case of the *Networking during Infectious Diseases Model (NIDM)*; Nunner et al., 2021). The SWIDM is a network formation model that considers 1. subjective risk perception regarding infectious diseases, 2. the propensity of forming closed triads, and 3. the likelihood of forming ties between individuals similar in risk perception. That is, autonomous agents form, maintain, or break social ties based on the trade-off between the benefits and costs of social relations, and potential harm of infections. Agent-based simulations of epidemics enabled us to consider a wide range of infectious diseases (mild vs. severe, low vs. high infectivity), health behaviors (low risk taking vs. high risk taking), homophily (low vs. high chance of ties between agents similar in risk perception), and small-world networks (fixed average degree, low vs. high clustering, short vs. long average path length).

In line with previous studies (Leung et al., 2018), our results suggest that epidemics in static small-world networks are on average larger, thus longer, and have higher peaks compared to epidemics in adaptive networks. Furthermore, we see that the dynamics of disease transmission are heavily shaped by network properties. Take the effect of clustering on an epidemic, for example. While we confirm a negative effect of clustering on final size in both adaptive and static networks (Badham & Stocker, 2010; Keeling, 1999; Miller, 2009), we observe opposing effects of clustering on duration in adaptive (–) and static (+) networks. In both cases, the disease needs to cross one of the few available bridges to reach another cluster. Consequently, only a few ties need to be cut to isolate a cluster, causing the disease to die out quickly in the adaptive networks. In the static networks, however, intrusion of a cluster is more likely, while spreading across one of the few bridges slows down disease spread.

Furthermore, our results suggest that health behavior homophily in adaptive networks may cause the epidemic to end earlier than compared to randomly mixed networks. That is, not only the properties of individual actors but also their composition in heterogeneous networks affect disease spread. Consider, for example, a risk averse ego connected to an infectious alter. The preference to avoid the infectious alter lowers not only the probability to get infected for the ego but also the probability for all of its susceptible connections. Consequently, even the connections that are less risk averse benefit from the ego’s decision to isolate from infectious alters. Additionally, our results suggest that the combination of network and actor properties shape disease spread in combination. That is, on the one hand, a cluster composed of mostly risk averse agents increases the chance for actors bridging two clusters to isolate the entire cluster from infectious alters. On the other hand, a cluster containing only a few risk averse agents increases the chance for actors bridging two clusters to maintain ties to infected alters from outside the cluster. This in turn increases the probability for the disease to spread into the cluster. Consequently, the overall large impact of homophilous clusters on epidemic size results from a comparably

large chance of entire clusters being isolated from the disease. In contrast to multiple clusters, isolation from infections in networks consisting of a single large cluster requires more ties to be cut and therefore such networks show a higher likelihood for the disease to spread to any part of the network (see also Watts & Strogatz, 1998).

These results are good news for two reasons. First, they support the notion that behavioral adaptation to health risks is a natural mechanism to not only reduce personal health risks, but also to mitigate disease spread. Dönges et al. (2021), for example, used a similar mechanism in a simulation study, concluding that non-pharmaceutical measures ought to leave sufficient room to exercise self-imposed measures resulting from risk perception, as strict measures of social distancing may result in rebound or compensatory effects when measures are lifted. Furthermore, social distancing may reinforce the mitigating effect of homophily. That is, limiting contacts needs to undergo an active selection process after which contacts with similar traits are more likely to remain (McPherson et al., 2001). Yamaguchi (1990), for example, showed that people with fewer friends show a stronger preference for similarity, such as regarding educational level, a trait that coincides with health behavior and risk perceptions (Bish & Michie, 2010). Second, we can use this mechanism to design targeted information campaigns addressing groups with lower risk perception to increase awareness of health risks. A remaining challenge is how to determine such groups. A promising way could be to find suitable proxies for scenarios of interest. From the literature, we know that risk perception often coincide with other personal characteristics, such as age, gender, ethnicity, educational level, or marital status (Bish & Michie, 2010). More specifically, a recent study among heterosexuals at high risk for HIV infection used individual and network data to determine predictors of self-perceived HIV risk (Koku & Felsher, 2020). The results show that men, individuals with lower education, and individuals in ethnically heterogeneous groups are on average more likely to perceive high risks.

Our study compares to earlier work in a few ways. Take Watts and Strogatz (1998), for example, who investigated the effects of clustering and path length on disease spread in static networks. Our findings also indicate that epidemic size decreases and epidemic duration increases with increasing values for clustering and average path length in static networks. In adaptive networks, however, we cannot confirm the effect of average path length on final size (when controlled for all other independent variables). As described earlier, this is a result of only a few bridges needing to be cut, thus causing the disease to die out quickly. Just as suggested by Kitchovitch and Li (2010), we observe that populations with on average less risk taking agents produce epidemics with smaller final size, shorter duration, and lower peak size. However, we considered homogeneous degree distributions, while Kitchovitch and Li (2010) used scale-free networks.

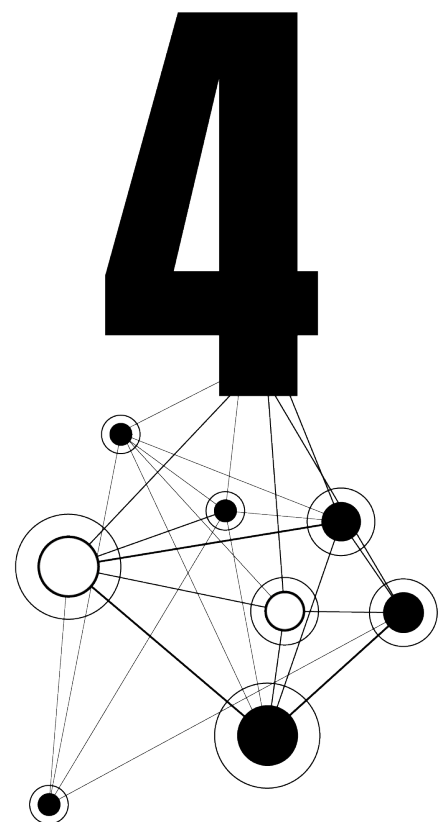
It is important to note that our study comes with a few limitations. First, it is not in and of itself reasonable to assume that agents know the disease state of others before creating social ties. Even at the time of evaluating a tie, a disease may be transmitted, affecting final size. Second, although research on COVID-19 has shown that infections are primarily transmitted through social network ties (e.g., Li et al., 2020; Shen et al., 2020), it is important to consider transmissions that are not based on cost-benefit considerations

(e.g., airborne diseases on public transport). This could be done in the form of a random component for spontaneous infections. Third, we only consider personal properties as a factor of risk perceptions. That is, agents are initialized to be either more or less risk averse throughout the entire course of a simulation run. Risk perceptions, however, may change according to changes in the prevalence of a disease or exchange of opinions with friends, which in turn may affect individual decision-making and thus disease spread (Lau et al., 2005; Teslya et al., 2022a). To account for such effects, our model could be extended so that risk perception reflects a personal tendency (base value) that is influenced by current events (e.g., variation according to prevalence). Fourth, in our model social relations are not “paused” but tried to be replaced. Although in some contexts this is unlikely to happen (e.g., families), in others it is to be expected (e.g., friendships break because one friend is not supportive in times of need). Fifth, the operationalization of homophily (in terms of tendency for social distancing) does not allow effects on epidemics in static networks. Other operationalizations, however, might be interesting extensions for future work.

To improve the model, other health behaviors lowering infectivity (e.g., handwashing, mask wearing) may be added. This would, most likely, also create an observable effect of health behavior homophily in static networks. Another approach to consider the effect of individual differences is to integrate a specific utility component expressing the preference to connect to others alike (*inbreeding* homophily). This has been neglected in the current work to ensure a better control of fixed average degree, and thus the isolated study of network components of interest (clustering, average path length). Furthermore, although our results show the general effect of health behavior homophily on a generic disease, a model could be informative for a specific scenario by fitting parameters to empirical data (e.g., disease severity, infectivity, degree distribution, clustering, recovery times) and refining compartments. Additionally, such an approach would allow considering additional factors that may affect individual susceptibility, treatment, and recovery, such as age, socio-economic status, or capacities of regional healthcare services. Finally, while the theories our model is based on have empirical grounding, our theoretical findings need to be corroborated by empirical experiments. This could be done by creating an experimental game-like study, in which human subjects take the role of the agents in our simulations. Another option is to fit the model to a specific scenario and study how well estimations fit to ongoing observations.

In conclusion, our model suggests that a combination of high network clustering and high health behavior homophily can mitigate the spread of infectious diseases. That is, adaptive agents can cause an epidemic to die out quickly by severing only a few bridging ties. The more risk averse a cluster is on average, the more likely bridging ties are severed, and the more likely it is that the entire cluster is isolated from the disease. Although clusters slow down disease spread in static networks, infections can still reach any part of the network, as agents are rendered incapable of reacting to their risk perceptions. Neglect of these fundamentally different dynamics may, therefore, lead to inaccurate estimations

when applied to a specific case. Considering the adaptive dynamics, however, can support the design of non-pharmaceutical interventions such as targeted information campaigns using proxies for risk perception.



Chapter 4

Disease avoidance may come at the cost of social cohesion: Insights from a large-scale social networking experiment

Abstract *It is known that people tend to limit social contacts during times of increased health risks, thus leading to disruption of social networks and changing the course of epidemics. It is, however, less known to what extent people show such avoidance reactions. To test the predictions and assumptions of an agent-based model on the feedback loop between avoidance behavior, social networks, and disease spread, we conducted a large-scale (2879 participants) incentivized experiment. The experiment rewards maintaining social relations and structures, and penalizes acquiring infections. We find that disease avoidance dominates networking decisions, despite relatively low penalties for infections; and that participants use more sophisticated strategies than expected (e.g., avoiding susceptible others with infectious neighbors), while they forget to maintain a profitable network structure. Consequently, we observe low numbers of infections, but also deterioration of network positions. These results imply that the focus on a more obvious signal (i.e., disease avoidance) may lead to unwanted side effects (i.e., loss of social cohesion).*

This chapter has been submitted to an international journal as Nunner, H., Buskens, V., Corten, R., Kaandorp, C., & Kretzschmar, M. (2023). *Disease avoidance may come at the cost of social cohesion: Insights from a large-scale social networking experiment* [unpublished manuscript]. Nunner and Buskens developed the mathematical model and designed the experiment. Kretzschmar contributed to the conceptualization of the model. Kretzschmar and Corten contributed to the conceptualization of the experiment. Nunner programmed a prototype of the experiment. Kaandorp programmed the experiment used for data collection. Nunner programmed the simulation, performed analyses, created visualizations, took care of data curation, wrote the original draft of the manuscript, and integrated the feedback of all co-authors. All authors reviewed and discussed the manuscript repeatedly.

4.1 Introduction

The literature contains numerous examples of individuals adapting behavior to lower their health risks (for reviews, see Davis et al., 2015; Verelst et al., 2016). Avoidance behavior, such as avoiding large crowds or public transport during an epidemic, or avoiding others who can be a source of infections, is a typical reaction to lower the personal probability of acquiring an infectious disease (Ferguson, 2007; Funk et al., 2010; Jones & Salathé, 2009). Additionally, the extent to which a person avoids others depends on individual risk perceptions (Bish & Michie, 2010; Ferguson, 2007; Leppin & Aro, 2009). That is, the higher people perceive the likelihood to catch a disease, and the more severe these people perceive the disease to be, the more likely they are to engage in avoidance behavior. Although risk perception is an individual characteristic, the composition of risk perceptions in a social network may influence disease spread (d’Andrea et al., 2022; Kitchovitch & Lì, 2010; Koku & Felsher, 2020). Furthermore, studies on epidemics in social networks have shown that the existence of clusters (densely connected areas within a network) can mitigate disease spread (Badham & Stocker, 2010; Keeling, 1999; Miller, 2009).

We believe that a better understanding of the interdependency between avoidance behavior and disease spread in social networks is crucial for the design of effective and efficient non-pharmaceutical interventions (*NPIs*), such as rules for authority imposed spatial distancing. For this purpose, we have developed the *Networking during Infectious Diseases Model* (*NIDM*; Nunner et al., 2021), an individual-based network model for infectious disease transmission. The NIDM defines network behavior, that is decisions to create, maintain, and break a social relation, as the trade-off between the benefits (e.g., affection, social capital, sense of belonging), costs (e.g., time, effort), and perceived risks of an infection (e.g., symptoms, hospitalization, absence from work) a social relation creates. Simulations with agents that myopically maximize the utility resulting from this trade-off produce non-linear dynamics that are hard to predict. That is, even small behavioral changes on the individual level can have large group-level effects, such as delaying an outbreak or preventing it entirely (Nunner et al., 2021). Furthermore, simulations with agents that differ in their perception of risk suggest that clusters composed of agents showing similar levels of avoidance behavior may enhance the mitigating effect of network clustering (Nunner et al., 2022a).

Although agent-based models have been powerful tools for modeling health behavior and disease spread (for a review, see Chang et al., 2020), their insights are usually based on computer simulations making specific assumptions on individual decision-making that need empirical scrutiny. For example, humans have cognitive limits and their decisions are typically influenced by social preferences (Camerer, 2003). Experiments can thus reveal whether human decision makers produce the same dynamics or whether dynamics change due to the models’ simplifying assumptions (Lunn & Ní Choisdealbha, 2018). Common types of experiments to address the aforementioned issues are so-called *incentivized experiments*. Participants of such incentivized experiments typically seek to optimize own or joint benefits by selecting the most rewarding actions in a given situation (Camerer, 2003).

Rewards, however, depend on the combined actions of the participants, so that decisions are made under uncertainty. To give an example, Woike et al. (2022) used a large-scale incentivized experiment to show that the effectiveness of behavioral interventions during an epidemic depends on the type of information shared with the participants (i.e., normative, informational). That is, reminding participants each round that they ought to adhere to the distancing rules to protect themselves and others was the most effective, while merely providing information on the actions of others were the least effective measures to minimize the number of infections.

4.1.1 The experiment

To study the feedback loop between avoidance behavior, network properties, and spread of infectious diseases, we developed the *Networking during Infectious Diseases Task (NIDT)*, the main task of a large-scale incentivized social networking experiment (see Figure 4.1). At its core, the NIDT is a round-based networking game (see Figure 4.2). Each participant is assigned to one of sixty nodes in a fictitious network with edges representing social relations. Time is modeled as discrete time steps (*rounds* of the game). Each round is composed of four consecutive stages: two interactive stages of decision-making, computation of point rewards, and simulation of disease transmissions. Participants alternate between stages 1 and 2, while stages 3 and 4 are performed by the software with the corresponding elements being updated during website reload.

4.1.1.1 Available actions

In stage 1 of each round, the NIDT selects per participant 12 nodes representing other participants, and offers each participant the opportunity to *create*, *maintain*, or *dissolve* the corresponding social relations. The type of decision depends on whether a relation exists to the selected node or not. That is, in stage 1 of each round, a relation to a neighboring node can be either maintained or dissolved. The creation of a new relation with a non-neighboring node can be proposed to the corresponding participant. Nodes in the network that are closer to the participant's node are prioritized by the NIDT. That is, selected nodes are on average 50% neighbors, 30% neighbors of neighbors, and 20% others.

In stage 2 of each round, the participants that received proposals to create new relations can decide whether to accept the proposals or not. Thus, dissolution of relations is a unilateral decision, while the creation of a relation requires the consent of both participants involved.

4.1.1.2 Point rewards

In stage 3 of each round, point rewards are awarded. Computation of points is based on theoretical considerations regarding the effects of social relations on well-being. That is, while social relations are beneficial for social well-being (e.g., affection, sense of belonging) (Ormel et al., 1999), they also produce costs in maintenance (e.g., effort, time) (Jackson, 2008). Additionally, infectious social relations constitute a risk to well-being, as they have

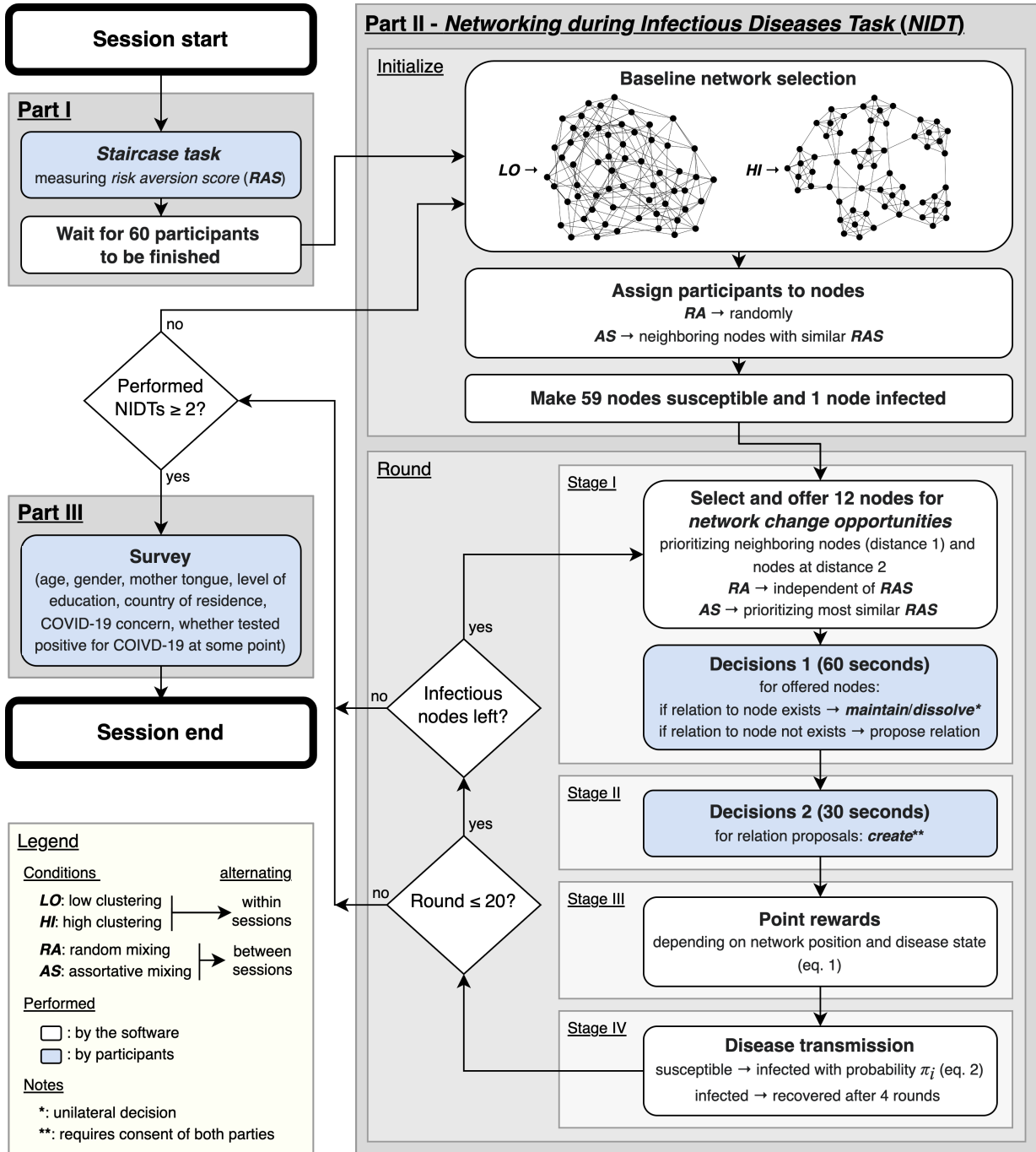


Figure 4.1: Flowchart of the experiment.

the potential to cause physical harm (Ormel et al., 1999). Consequently, social networking decisions in the context of infectious diseases are based on a trade-off between the social benefits, social costs, and perceived potential health costs a social relation creates (for more details on the theoretical background of our model see Nunner et al., 2021).

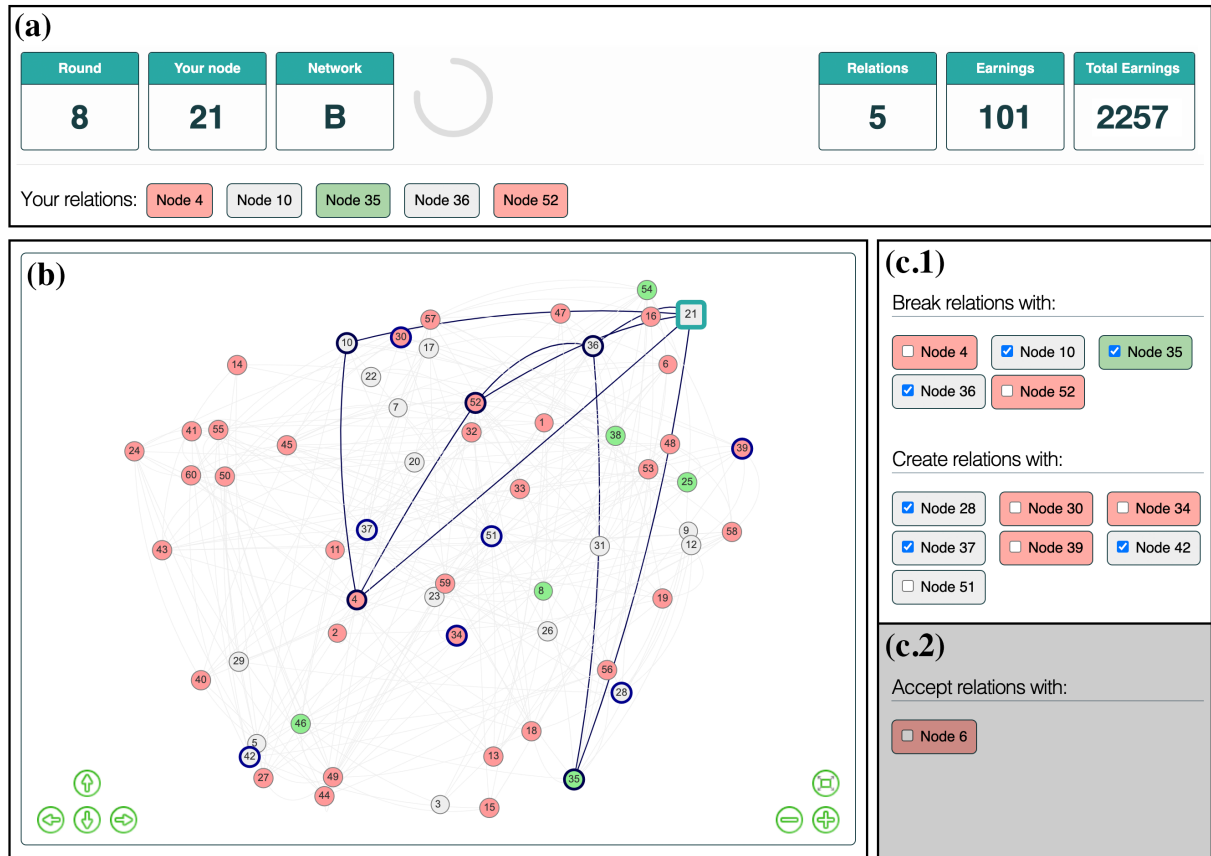


Figure 4.2: **Graphical user interface (GUI) of the *Networking during Infectious Diseases Task (NIDT)*.** The GUI consists of three parts: the information bar (a), the network viewer (b), and the interaction panel (c). On the left, the information bar (a) displays the current round of the game, the node ID of the participant, the network type that is being played (A: low clustering *LO*, B: high clustering *HI*), how much time is left in the form of a retracting circular line (60 seconds for breaking and proposing relations, 30 seconds for accepting relation proposals), and the node IDs of the current relations. On the right, the information bar (a) displays how many relations the participant has, how many points were awarded in the previous round, and how many points were awarded in total. The network viewer (b) displays the entire network with all nodes, disease states (gray: susceptible, red: infected, green: recovered), and social relations. The node of the participant is displayed as an enlarged square (here: node 21), while all other nodes are displayed as circles. Offered nodes are highlighted with a thick blue line. By clicking a node in the network, the relations of that node are highlighted with a blue line, while all other relations are shown in light gray. Positions of nodes can be changed by clicking and dragging nodes. The interaction panel (c) is separated into two sub-panels. The first sub-panel (c.1) is only visible in the first interactive stage of a round and shows the IDs of the nodes offered for decision opportunities. By unchecking a checkbox in the upper area of sub-panel c.1, participants can break an existing relationship, while not changing a checked checkbox corresponds to maintaining a relationship. By checking a checkbox in the lower area of sub-panel c.1, participants can propose to create a new relationship to another participant. The second sub-panel (c.2) is only visible in the second interactive stage of a round and shows the node IDs of the corresponding participants that proposed to create a relation in the first interactive stage of the round. Checking a box in this section implies that the proposal to create a relation is accepted.

Furthermore, costs and benefits for social relations depend on social context. That is, in a work environment, people may prefer relations between not connected others (A is connected to B and C, while B and C are not connected to each other) to be able to control information flow (Burt, 1992). In private settings, such as family and friends, people may prefer *triadic closure* (A is connected to B and C, while B and C are also connected to each other) (Simmel, 1950) for better social support or more leverage to enforce social norms (Coleman, 1994). These individual-level preferences exert scaling effects on the structure of the entire social network, by either giving rise to a large sparsely connected component or several densely connected clusters.

Based on the aforementioned considerations, the NIDT defines point rewards (or *utility*, U) of an individual i as the composition of three terms:

$$U_i = \left[b_1 \cdot t_i + b_2 \cdot \left(1 - 2 \cdot \frac{|x_i - \alpha|}{\max(\alpha, 1 - \alpha)} \right) \right] - \left[c_1 \cdot t_i + c_2 \cdot t_i^2 \right] - \left[\sigma \right]. \quad (4.1)$$

First, there are rewards (b_1) for the number of social relations (t_i) and the weighted (b_2) proportion of closed triads (x_i) i belongs to, while α is the preferred proportion of closed triads. This operationalization leads to a single peaked utility component on the proportion of closed triads. Second, there are marginally increasing costs (c_1 , c_2) for the number of social relations (t_i). Third, there are costs for being infected (σ). The utility function allows us to control two important properties: the preferred number of relations and the preferred proportion of closed triads. Since such a complex reward system is not only difficult to understand but also too artificial to be used in an experiment, we have chosen parameter settings resulting in easy to understand point rewards. Regarding the number of relations, we implement parameters for six relations to be most beneficial. That is, having six relations awards 100 points per round. Having more or fewer relations results in progressively lower point rewards (see Table 4.1). Negative rewards for number of relations were set to 0, so that having more than 11 relations did not result in a loss of points. To incentivize the maintenance of clustering of the initial networks (see Figure 5.2), the NIDT awards up to 20 points depending on the setting for clustering (for details on parameter settings and conditions, see below). That is, in the low clustering setting (*LO*; $\alpha = 0.0$) 20 points are rewarded for not being part of a closed triad. In the high clustering setting (*HI*; $\alpha = 0.67$) 20 points are rewarded if two thirds of the neighbors form closed triads. The further away participants are from the optimal clustering, the fewer points they receive. Node positions in the initial networks are optimal in terms of network benefits. That is, in the absence of infection, benefits cannot be increased by adding or removing social relations. The experiment is thus focused on changing network structure under the threat of infection. Finally, 14 points are deducted (possibly resulting in negative rewards) for every round a participant is infected.

Table 4.1: Properties and corresponding point rewards per round.

	Point rewards [*]	
<i>Number of relations</i> [†]		
6	100	
5 or 7	97	
4 or 8	89	
3 or 9	75	
2 or 10	56	
1 or 11	31	
0 or 11+	0	
<i>Proportion of neighbors forming closed triads</i> [‡]	<i>LO</i> [§]	<i>HI</i> [¶]
0.00	20	0
0.33	7	7
0.67	0	20
1.00	0	7
<i>Disease state</i> [#]		
Susceptible	0	
Infected	−14	
Recovered	0	

Notes: ^{*}: Based on Equation 4.1, parameter settings (see below), and a scaling factor of 41.55. [†]: $b_1 = 1.0$, $c_2 = 0.067$. [‡]: The proportions shown are incomplete and serve only to illustrate the difference between the *LO* and *HI* settings. [§]: $b_2 = 0.5$, $\alpha = 0.0$. [¶]: $b_2 = 0.5$, $\alpha = 0.67$. [#]: $\sigma = 0.34$.

4.1.1.3 Disease transmission

In stage 4 of each round, disease transmissions from infectious to susceptible nodes via social relations are computed. The probability for a susceptible node to get infected depends on the number of infected (and thus infectious) neighbors:

$$\pi_i = 1 - (1 - \gamma)^{t_{iI}}, \quad (4.2)$$

with γ denoting the probability to get infected per neighbor and t_{iI} the number of infected neighbors of participant i . We use $\gamma = 0.15$. Once infected, nodes recover after 4 rounds and cannot get infected again.

The NIDT concludes either after 20 rounds, or when no more infectious nodes are left in the network.

4.1.2 Hypotheses and conditions

Hypotheses and conditions were derived based on a simulation study (Nunner et al., 2022a). Furthermore, parameter settings for the conditions were determined using additional simulations (for details see Section C.2.1 of the Supplements to Chapter 4). In all simulations, we used artificial agents that myopically optimized rewards provided by Equation 4.1. To

realize individual risk perceptions, we replaced the costs of being infected (σ) with perceived costs of being infected (σ^{r_i}) depending on perceived risks of getting infected ($\pi_i^{2-r_i}$):

$$\left[\sigma \right] \rightarrow \sigma^{r_i} \cdot \pi_i^{2-r_i}. \quad (4.3)$$

That is, the risk perception parameter r_i ($0.0 < r_i < 2.0$) transforms severity of the disease (σ) and probability of getting infected (π_i) into subjective versions of the same. Consequently, higher values for r_i lead to higher perceived costs of being infected and higher perceived risks of contagion, and therefore corresponds to higher risk aversion. A value of $r_i = 1.0$ corresponds to risk neutrality.

Simulations showed that susceptible agents avoid infectious agents, higher risk aversion leads to stronger avoidance of infectious agents, and stronger avoidance of infectious agents leads to smaller epidemics (Nunner et al., 2022a). We therefore hypothesize:

H4.1: Infectious alters are avoided by susceptible egos.

H4.2: Higher risk aversion causes stronger avoidance of infectious alters.

H4.3: Higher risk aversion lowers the individual probability to get infected.

To test Hypotheses **H4.1-H4.3**, we measured risk aversion in participants using the *staircase task* (Andersen et al., 2006; Falk et al., 2016). The staircase task consists of a series of five binary choices between a guaranteed reward and a 50:50 chance to win a higher reward. The more often a person opts for the guaranteed reward, although rewards for gambling increase, the more risk averse that person is considered to be. Based on the threshold when a participant prefers gambling over a guaranteed return, we defined a *risk aversion score* per participant ranging from 0.0 to 2.0 with the same interpretation as in our simulations (< 1.0 : risk seeking, 1.0 : risk neutral, > 1.0 : risk averse).

In line with earlier studies (e.g., Badham & Stocker, 2010; Keeling, 1999; Miller, 2009), our simulations suggested that epidemics are smaller in networks with multiple densely connected clusters than in networks with a more open structure. That is because relations that bridge two clusters pose a bottleneck that impedes the further spread of a disease. In addition, our simulations suggested that assortative mixing in terms of health risk perceptions, and thus the local aggregation of individuals who respond similarly to health risks, may further reduce epidemic size. Finally, our simulations suggested that the combination of assortative mixing in networks with multiple clusters produces the smallest epidemics (Nunner et al., 2022a). That is because a cluster composed of predominantly risk averse agents is likely to quickly dissolve bridges to infectious alters, protecting the entire cluster from getting infiltrated by the disease. We therefore hypothesize:

H4.4: Epidemics are smaller in networks with a higher degree of clustering.

H4.5: Epidemics are smaller in networks with a higher degree of assortative mixing regarding risk aversion.

To test Hypotheses **H4.4** and **H4.5**, we defined four conditions (*LO:RA*, *LO:AS*, *HI:RA*, *HI:AS*) varying network clustering and assortative mixing in a two-by-two design. Network clustering is either low (*LO*) or high (*HI*) and social mixing is either random (*RA*) or assortative (*AS*) regarding risk aversion. Clustering settings were realized with two different baseline networks. That is, one network was composed of a single large cluster (*LO*), while the other contained several densely connected clusters (*HI*; see Figure 5.2). To eliminate unwanted structural side effects, both networks were similar regarding average degree (both 5.93) and closeness (*LO*: 0.975, *HI*: 0.955).

While settings for clustering were communicated to the participants, settings for social mixing were not. To realize differences in social mixing, the NIDT first initialized networks with either low or high degrees of assortative mixing. That is, risk aversion scores of participants were first arranged in ascending order. Thereafter, the corresponding participants were assigned to fixed orders of nodes to attain either networks with neighboring nodes possessing mostly different risk aversion scores (*RA*) or networks with neighboring nodes possessing mostly similar risk aversion scores (*AS*). Furthermore, to maintain the level of assortative mixing, nodes for decision opportunities (creation and dissolution) were also selected based on risk aversion score. That is, to inhibit assortative mixing (*RA*), the NIDT selected nodes independent of risk aversion score ($\omega = 0.0$, with ω denoting the probability of selecting the node most similar regarding risk aversion score). To foster assortative mixing (*AS*), the NIDT prioritized nodes most similar regarding risk aversion score ($\omega = 0.8$).

Finally, we predefined the index cases (the initially infected node) for each condition so that the epidemic started with the node that is closest to average degree, average local clustering coefficient, and average risk aversion score. We used neutral, not emotionalizing language to explain the rules of the NIDT (see Section C.1.3 of the Supplements to Chapter 4). That is because we intend to study risk perceptions as an analytical heuristic in decision-making. Furthermore, although the experiment was conducted during the COVID-19 pandemic, we did not use COVID-19 as context or motivation. We believe that this facilitates the repeatability of the experiment and the comparability of the results for other conditions or parameter settings.

4.2 Results

Our results are based on data collected from 48 experimental sessions, with each session composed of one staircase task, two NIDTs (clustering alternating within each session, social mixing alternating between sessions), and a final survey (age, gender, mother tongue, level of education, country of residence, COVID-19 concern, whether tested positive for SARS-CoV2 at some point). Consequently, we analyze a total of 96 NIDTs, with 24 NIDTs per experimental condition, and a total number of 711,159 networking decisions made by 2,879 participants.

Table 4.2: Proportions of specific decisions made given specific decision opportunities, and the proportion of those decisions with rewards higher than the rewards of the opposite decision.

		S		Ego I		R	
<i>Create</i>							
Alter	S	0.24	(0.33)	0.50	(0.82)	0.26	(0.60)
	I	0.04	(0.20)	0.47	(0.74)	0.24	(0.50)
	R	0.47	(0.36)	0.74	(0.80)	0.41	(0.52)
<i>Not create</i>							
Alter	S	0.76	(0.80)	0.50	(0.27)	0.74	(0.75)
	I	0.96	(0.84)	0.53	(0.28)	0.76	(0.75)
	R	0.53	(0.79)	0.26	(0.24)	0.59	(0.81)
<i>Dissolve</i>							
Alter	S	0.09	(0.59)	0.07	(0.43)	0.15	(0.89)
	I	0.69	(0.56)	0.19	(0.33)	0.21	(0.63)
	R	0.04	(0.73)	0.02	(0.18)	0.06	(0.82)
<i>Maintain</i>							
Alter	S	0.91	(0.61)	0.93	(0.75)	0.85	(0.47)
	I	0.31	(0.49)	0.81	(0.79)	0.79	(0.59)
	R	0.96	(0.54)	0.98	(0.82)	0.94	(0.55)

Notes: Numbers without parentheses give the proportion of decisions that were made for all decision opportunities in that cell; numbers within parentheses give the proportion of the decisions that were made that led to higher rewards than the opposite decision. For example, when susceptible egos had the opportunity to create a relation to susceptible alters, they were willing to make that decision in 24% of all opportunities. In 33% of these decisions to create a relation, the rewarded points were higher than if they had not decided to create that relation. Furthermore, *Create* and *Not create* as well as *Dissolve* and *Maintain* are opposite decisions for the same opportunities: *Create* → the combination of accepted opportunities to propose a relation in stage 1 and accepted relation proposals in stage 2 of a round; *Not create* → the combination of declined opportunities to propose a relation in stage 1 and accepted relation proposals in stage 2 of a round; *Dissolve* → accepted opportunity to dissolve a relation in stage 1 of a round; *Maintain* → declined opportunity to dissolve a relation in stage 1 of a round.

4.2.1 The role of disease avoidance in decision-making

Table 4.2 shows the proportion of decisions made (accepted decision opportunities) to decision opportunities (nodes offered by the NIDT in stage 1 and relation proposals by other participants in stage 2 of a round). Furthermore, Table 4.2 shows the proportion of decisions made that were rewarding in terms of increasing or maintaining point rewards to decisions made that lowered point rewards (in parentheses). The data is divided by decision type (*create*, *not create*, *maintain*, *dissolve*) and disease states (**S**usceptible, **I**nfected, **R**ecovered) of the ego (decision maker) and alter (subject of the decision).

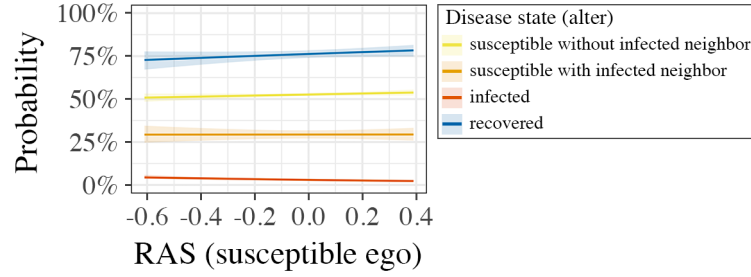


Figure 4.3: **Marginal effects of (i.) risk aversion score of susceptible egos and (ii.) disease states of alters on the probability of a relation being desirable.** Risk aversion score is mean centered at $M = 1.22$ ($SD = 0.46$), denoting that the average participant was slightly risk averse (< 1.0 : risk seeking, 1.0 : risk neutral, > 1.0 : risk averse). Risk aversion score refers to the ego, the node, to describe the degree of desire for a relation. Colors denote the disease state of the alter, the subject of the desire for a relation.

The data reveal that infected alters are avoided the most. That is, relations to infected alters were created the least, while existing relations to infected alters were dissolved the most. Although this effect can be found independent of disease states of the ego, it is strongest for egos that are susceptible and thus at risk of getting infected themselves. That is, susceptible egos dissolved 69% of the relations if the offered node was infected, compared to 9% and 4% if the offered node was susceptible or recovered. The decision of susceptible egos to avoid infectious alters, however, was rewarding for only about half of the decisions (56%). The average decision of infected egos to create and maintain relations, in contrast, was among the most rewarding. This is likely an effect of the lowered average degree of infected nodes, which dropped from an average of more than seven relations in the round before acquiring an infection ($M = 7.4$, $SD = 3.1$) to less than 3 relations for the entire period of being infected ($M = 2.8$, $SD = 1.99$; see also Figure C.5 in the Supplements to Chapter 4). Creating and maintaining relations therefore allows compensating for points lost as a result of social isolation. Thus, these data suggest that while susceptible egos make decisions in order to avoid the disease at a potential cost of point rewards, infected egos make decisions in order to minimize losses due to being socially isolated.

As Table 4.2 suggests that disease states influence how much one relation is preferred over another, we performed logistic regressions on the factors contributing to the attractiveness of a relation (Table 4.3). A relation being attractive is represented by a binary variable composed of decisions to create a relation (proposals in stage 1 and accepted proposals in stage 2 of a round) and declined opportunities to dissolve a relation. Based on the regression analysis, Figure 4.3 reveals that susceptible egos favor relations to recovered alters, followed by susceptible alters without infected neighbors, then susceptible alters with infected neighbors, and finally infected alters. The interaction effect between susceptible egos, infected alters, and risk aversion score (also visible by the negative slope of the red line in Figure 4.3), reveals that the more risk averse the ego is, the less attractive an infected alter gets. Furthermore, Table 4.3 shows that recovered alters are the most attractive, while infected alters are among the least attractive relations for all egos.

In summary, we find support for Hypotheses **H4.1** and **H4.2**. That is, infected alters are the least desirable nodes for susceptible egos and higher risk aversion causes stronger avoidance of infectious alters. Our results, however, go beyond what could be expected from our model. That is, on the one hand, avoidance behavior towards infected alters can also be observed for infected and recovered egos. On the other hand, susceptible nodes that have infectious neighbors are avoided by susceptible egos. As a result, infected nodes are getting more strongly isolated in the experiment than in our simulations. That is, infected nodes in the experiment have on average 2.8 relations, while infected nodes in the simulations have on average 5.6 relations (see also Figure C.5 (b) in the Supplements to Chapter 4). This stronger isolation of infected nodes in the experiment has direct implications for disease spread on the group-level.

4.2.2 Group- and individual-level effects

Figure 4.4 shows that, on average, final epidemic size (the proportion of infected nodes during a single installment of the NIDT) did not differ between settings for clustering (*LO*: $Mdn = 0.05$, *HI*: $Mdn = 0.03$) and mixing (*RA*: $Mdn = 0.04$, *AS*: $Mdn = 0.04$). Wilcoxon rank-sum tests confirmed that the differences were not statistically significant for clustering ($Z = -1.31$, $p = 0.19$) and mixing ($Z = -0.06$, $p = 0.95$). Because these results do not show any discernible differences between the different settings, we can neither confirm nor reject our group-level hypotheses regarding clustering and assortative mixing (Hypotheses **H4.4** and **H4.5**). In all conditions, the disease hardly ever spread to more than ten percent of the entire population (six nodes), which was also considerably less than expected by our simulations (see Figure C.4 in the Supplements to Chapter 4).

Furthermore, Table 4.4 shows the results of logistic regression analyses for factors that contribute to getting infected. The strongest predictor for getting infected is the number of rounds a participant had more than 12 relations. Having many relations creates many potential transmission routes, but on the other hand, having many relations decreases the chance that an infected neighbor is selected by the NIDT for the opportunity to break that relation. Although Table 4.4 suggests that risk aversion does not affect the probability to get infected, the overall absence of disease spread does not allow a final judgment whether higher risk aversion affects the probability to get infected in general (Hypothesis **H4.3**).

Table 4.3: Factors contributing to the attractiveness of a relation.

Constant	5.866***	(0.074)
<i>Disease states</i>		
Ego ^S - Alter ^S without infected neighbor	-0.103*	(0.059)
Ego ^S - Alter ^S with infected neighbor	-1.070***	(0.061)
Ego ^S - Alter ^I	-3.728***	(0.067)
Ego ^S - Alter ^R	1.070***	(0.063)
Ego ^I - Alter ^S without infected neighbor	0.200***	(0.062)
Ego ^I - Alter ^S with infected neighbor	-0.014	(0.079)
Ego ^I - Alter ^I	-1.347***	(0.094)
Ego ^I - Alter ^R	1.949***	(0.114)
Ego ^R - Alter ^S without infected neighbor	-0.886***	(0.057)
Ego ^R - Alter ^S with infected neighbor	-0.813***	(0.083)
Ego ^R - Alter ^I	-1.560***	(0.084)
Ego ^R - Alter ^R (reference)		
<i>Risk aversion</i>		
Risk aversion score (RAS)	0.185	(0.172)
Ego ^S - Alter ^S without infected neighbor \times RAS	-0.037	(0.161)
Ego ^S - Alter ^S with infected neighbor \times RAS	-0.130	(0.179)
Ego ^S - Alter ^I \times RAS	-0.833***	(0.235)
Ego ^S - Alter ^R \times RAS	0.182	(0.199)
<i>Network properties</i>		
Degree	-0.492***	(0.007)
Degree ²	0.020***	(0.001)
Degree of alter	-0.043***	(0.001)
Clustering (<i>HI</i>)	0.026	(0.032)
<i>HI</i> \times ECC [†]	-10.258***	(0.106)
<i>LO</i> \times ECC [†]	6.655***	(0.137)
<i>Decision opportunity type</i>		
Creation opportunity in stage 1 of a round	-4.126***	(0.010)
Creation opportunity in stage 2 of a round	-3.299***	(0.015)
Number of offers by opportunity type	-0.098***	(0.003)
Log Likelihood	-286,474	
AIC	573,006	
BIC	573,339	
Observations	711,159	
Number of groups (rounds)	52,650	
Number of groups (nodes)	5,758	
Number of groups (games)	96	
Variance (rounds)	0.28	
Variance (nodes)	0.87	
Variance (games)	0.01	

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, SEs in parentheses. [†]*ECC*: Expected Change in Clustering. The numbers describe a four-level random intercept logistic regression model (level 4: games, level 3: nodes, level 2: rounds, level 1: decisions) of whether a relation is attractive (0: declined opportunities to create a relation in stage 1 and 2 of a round, and accepted opportunities to break a relation, 1: accepted opportunities to create a relation in stage 1 and 2 of a round, and declined opportunities to break a relation). Refer to Table C.10 in the Supplements to Chapter 4 for a comparison of models by different decision opportunities.

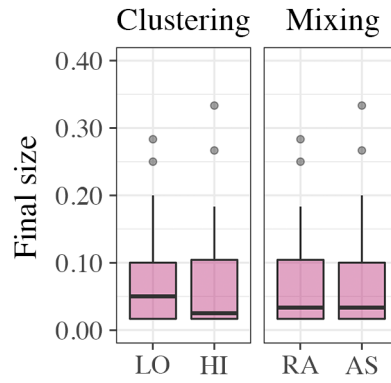


Figure 4.4: **Final epidemic size by clustering and social mixing.** Box-and-whisker plots show the median, interquartile range, minimum, maximum, and outliers of final size by settings for clustering and social mixing. Final size describes the proportion of nodes that got infected at some point during a single installment of the NIDT.

Table 4.4: Risk factors for acquiring infection during one installment of the NIDT.

Constant	−2.004***	(0.333)
<i>Risk aversion</i>		
Risk aversion score	−0.006	(0.008)
<i>Parameter settings</i>		
Clustering (<i>HI</i>)	−0.117	(0.092)
Mixing (<i>AS</i>)	0.025	(0.115)
<i>Individual-level controls</i>		
Age	−0.013	(0.008)
Gender (female)	−0.130	(0.124)
Gender (other)	0.423	(0.236)
Mother tongue (English)	−0.073	(0.173)
Education	−0.018	(0.043)
COVID-19 concern	−0.035	(0.064)
COVID-19 positive	0.024	(0.166)
Rounds relations 12+	0.148***	(0.025)
Log Likelihood	−1,452.84	
AIC	2,930	
BIC	3,010	
Observations	5,758	

Notes: *** $p < 0.01$. SEs in parentheses were made robust using clusters of 2,879 participants and two observations (NIDTs) per participant. The numbers describe a logistic regression model of whether a node acquired an infection during one installment of the NIDT (0: no, 1: yes).

4.2.3 Determinants of decision-making beyond avoidance behavior

Table 4.3 reveals further insights into the decision-making process. We observe that alters with fewer relations are more desirable. This might be a result of the assumption that alters with lower number of relations, especially below six, are more likely to accept opportunities to create a relation to attain the optimal number of six relations. We also find significant interaction effects for expected change in clustering and the two clustering conditions. However, contrary to what we would expect, in the high clustering setting decisions were favored that were expected to decrease clustering. In the low clustering setting, decisions were favored that were expected to increase clustering. Considering that in both clustering settings, the baseline networks ensured that all nodes started at an optimal position regarding clustering, a decision that did not consider clustering for decision-making would therefore be likely to work against the initial optimum. Consider, for example, a node with five neighbors and none of them share relations with each other (*LO*). Furthermore, the node gets offered six opportunities of which four refer to neighbors of the node's neighbors. A random decision that does not consider the expected change in clustering would likely lead to an increase of clustering, resulting in a less optimal network position, and thus suboptimal point rewards.

To delve deeper into the decision-making process, we ran additional simulations with parameters set according to the empirical data (for details see Section C.2.2 of the Supplements to Chapter 4). That is, we set risk perception (parameter r in Equation 4.3) to the risk aversion score observed in our experiment ($M = 1.22$, $SD = 0.46$), which was slightly less risk averse than the initial setting ($M = 1.27$, $SD = 0.45$) based on Vriens and Buskens (2020). Furthermore, we found that 65.61% of all decisions in the experiment were rewarding in terms of increasing or maintaining point rewards, while there was a clear preference to avoid infected alters. We therefore adjusted agents so that all decisions that involved infected alters were made in order to increase point rewards, while two out of three decisions that did not involve infected alters were made to increase point reward and one out of three decisions was made to decrease point reward. We also found that 14% of the participants tried to maximize the number of relations rather than point rewards. We, therefore, initialized 14% of the agents with lower marginal costs for relations so that they sought to attain twenty ($c_2 = 0.02$) rather than six ($c_2 = 0.067$) relations. Finally, to achieve a similarly strong avoidance reaction towards infected alters, we increased perceived disease severity and perceived infectivity by a factor of 2.5 (see Figures C.5 (a) and C.7 in the Supplements to Chapter 4).

Figure 4.5 shows a comparison of how average utility, average degree, average clustering, and disease states progress over time between the experiment (pink), the initial simulation to determine parameter settings for the experiment (*simulation 1*, green), and the additional simulated experiment with parameters set according to the data from the experiment (*simulation 2*, purple). Furthermore, we divided the plots into the two settings for clustering (left: *LO*, right: *HI*). The plots show that participants did not act like the

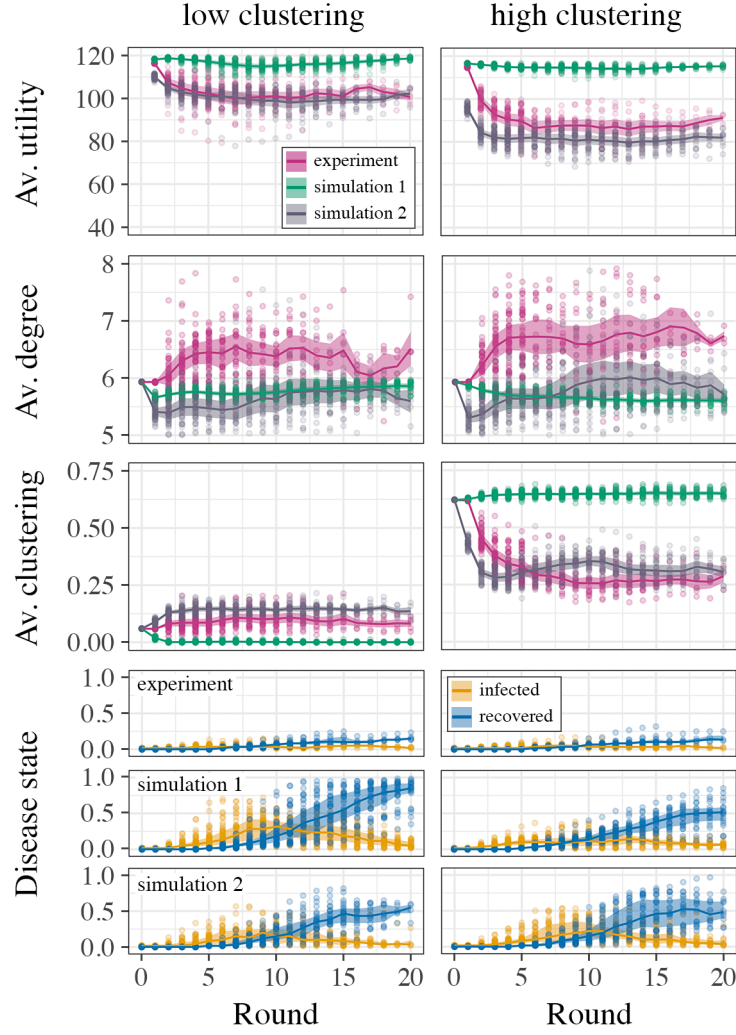


Figure 4.5: **Progression of properties over time.** Plots are divided by different properties (rows) and clustering settings (columns). Average utility (row 1), average degree (row 2), and average clustering (row 3) depict the results for the experiment (pink), simulation 1, which has been performed to determine parameter settings for the experiment (green), and experiment 2 with parameter settings resembling the experiment (purple). Progression of disease states has been divided into individual rows per data source (experiment: row 4, simulation 1: row 5, simulation 2: row 6). Each plot shows the progression of the corresponding property over time (max. 20 rounds). Dots within the plots show the mean of all nodes per single NIDT. Lines depict the mean and ribbons the standard deviation for NIDT means. Refer to Figure C.6 in the Supplements to Chapter 4, for an extended version, including settings for social mixing and degree of health behavior homophily.

agents in simulation 1. That is, average degree was on average higher for both settings, while clustering was on average higher in the low clustering setting and lower in the high clustering setting. As a consequence, average utility was on average lower. Although 14% of agents that sought to maximize the number of relations, the average degree did not

deviate much from the optimal average of six relations.¹ Similar to simulation 1, average degree of infected nodes remained at a comparatively high level in simulation 2, and thus infected nodes were not as strongly avoided as in the experiment (see Figure C.5, plot (b), in the Supplements to Chapter 4). Furthermore, we can see that simulation 2 was very close to the experiment in terms of average clustering. That is, due to two out of three decisions being made randomly if no infected neighbor is involved, we see a rise of clustering in the low clustering setting and a drop of clustering in the high clustering setting. Although the parameter adjustments in simulation 2 produced better matches for average clustering and average utility, they did not affect the spread of the disease.

4.3 Discussion

To study the extent of avoidance behavior in people and how the corresponding level of avoidance affects disease spread in social networks, we developed the *Networking during Infectious Diseases Task* (NIDT). Data collected from a large-scale (2,879 participants) online experiment revealed that infected alters are the least desirable relations for susceptible egos and that higher risk aversion causes stronger avoidance of infectious alters (supporting Hypotheses **H4.1** and **H4.2**). The extent of avoidance behavior was, however, stronger than what we expected from simulations, resulting in a fictitious disease to hardly spread to more than ten percent of the nodes in our networks. We could therefore neither confirm nor reject whether higher risk aversion affects the probability to get infected (**H4.3**), and whether networks composed of multiple clusters and networks with neighbors similar in risk perception have mitigating effects on disease spread (**H4.4** and **H4.5**).

Analyses on decision-making revealed that participants deviated from the agents in our simulations in a few, but important aspects. That is, especially susceptible participants prioritized disease avoidance over social benefits. Clustering, for example, did not play a role in networking decisions, although it could create higher rewards (up to 20 points per round) than the penalty of being infected (-14 points per round). As a result, clustering deteriorated for both clustering settings (increasing in the low clustering setting, decreasing in the high clustering setting). Although additional simulations with parameters set according to the data collected in the experiment revealed that neglect of clustering in networking decisions can reproduce the observations from the experiment regarding clustering, we were unable to reproduce the observed disease spread. Furthermore, we increased the perceived disease severity and infectivity to achieve similarly strong avoidance reactions. The disease, however, still spread much further in the simulation than in the experiment. The reason is probably that in the experiment the decisions that resulted in suboptimal rewards were not random, but tended towards avoiding infected alters. We found, for example, that participants integrated additional information into networking decisions, information that is not incorporated into the decision-making rules of the simulated agents. Participants in

¹Refer to Figure C.7 in the Supplements to Chapter 4, for an overview of how variations of the input parameters informed by the experiment data in simulation 2 (i.e., marginal costs for additional relations, probability of random decisions, and disease overestimate) affect outcomes (i.e., average degree, average clustering, final size).

the experiment not only avoided infectious nodes, but also nodes with infectious neighbors. On the other hand, infected alters were avoided irrespective of the disease state of egos. That is unexpected because infected and recovered egos were not at risk of experiencing any (additional) disadvantages when neighboring infected alters. Consequently, infected individuals were more isolated in the experiment than in the simulations, causing disease spread to stop quickly in the experiment.

We identify three main reasons for why participants prioritize disease avoidance in networking decisions. First, infections create salient signals (red nodes) with obvious consequences (14 points deducted as penalty for every round of being infected). In comparison to infections, clustering requires more effort to keep track of. That is, determining the number of closed triads requires identifying all neighbors including their relations between each other. Furthermore, changes in clustering may occur stepwise over several rounds, so that corresponding changes in point rewards are less noticeable. In the literature, we typically assume that people act parsimonious regarding cognitive resources (Gigerenzer & Todd, 1999), and myopically regarding the optimization of rewards (Halevy, 2016). If a person chooses to optimize only one element, the choice is thus likely in favor of disease avoidance because it is easier to evaluate.

Second, from *Prospect Theory* (Kahneman & Tversky, 2013; Tversky & Kahneman, 1992) we know that people do not act rationally in the sense of maximizing rewards purely based on expected economic utility. In particular, people appear to be more motivated by avoiding losses than by maximizing gains. Gains and losses, however, are not absolute outcomes but are relative to a reference point. In addition, people are generally bad at estimating risks or probabilities for the occurrence of events (Slovic & Peters, 2006) and tend to overestimate the probability of rare events (Slovic & Weber, 2002). As a result, people tend to reject tangible benefits in order to avoid low probabilities of losses (Denes-Raj & Epstein, 1994). If we consider point rewards from the previous round as reference point and interpret getting infected as a loss, a deviation of behavior compared to the model is in line with Prospect Theory. That is, participants seem to favor the aversion of a loss (-14 points per rounds of being infected) above a potential higher benefit (up to 20 points per round for maintaining social structure), although the risk of getting infected is low (15% for one infectious neighbor per round).

Third, the COVID-19 pandemic was ongoing at the time we conducted our experiment in July 2021. Although we did not mention the pandemic in our instructions and kept to a neutral, not emotionalizing language (see Section C.1.3 of the Supplements to Chapter 4), Slovic and Weber (2002) argue that hazardous events, such as the outbreak of a disease, can trigger *social amplification of risk* (Kasperson et al., 1988). That is, behavioral responses are subject not only to individual risk assessment, but also to *social amplification stations*, such as the media, cultural groups, or interpersonal networks. A topic that passes through the filter of these social amplification stations and is backed up with additional information, such as expert commentary, can trigger responses that go beyond the miti-

gation of immediate harm. In the context of disease avoidance, the consistent presence of the COVID-19 pandemic in the public mind may therefore have amplified the avoidance of infected nodes in our experiment.

Although our experiment was conducted in an abstract environment, and its outcomes cannot easily be generalized to the real world, the results can be considered a warning signal that loss of social cohesion might occur during epidemics. Just like in our experiment, visible symptoms of infection, or shared information about personal health are salient signals for behavioral adaptation. Furthermore, the consumption of disease related information in the media may increase perceived disease severity and perceived susceptibility to infectious diseases (Tagini et al., 2021). Both symptoms of others and media consumption have the potential to trigger or reinforce avoidance reactions in the short term. Corrosion of social cohesion as a result of disease avoidance, however, is a comparatively less salient signal for behavioral adaptation, while it has the potential to lower well-being in the long term. The COVID-19 pandemic has already shown that the implementation of distancing measures can lead to social isolation and consequently, to anxiety, depression, sleep deprivation, substance abuse, etc. (Banerjee & Rai, 2020; Heape, 2021; Kim & Jung, 2021; Pietrabissa & Simpson, 2020; Sepúlveda-Loyola et al., 2020). An innate preference to avoid infections may reinforce social isolation, lead to the corrosion of social coherence, and thus the loss of individual well-being in the long run.

It is furthermore important to point out the limitations of our study. First, although we provide simple rules and give examples of how clustering affects point rewards, keeping track of clustering in the experiment might have been too abstract or tedious. Second, we acknowledge that risk perception is a combination of analytic and affective processes that heavily rely on each other (Slovic & Peters, 2006; Slovic & Weber, 2002). Our experiment, however, is focused on the analytic processes of risk perception. That is, we used neutral, not emotionalizing language without framing of personal relationships.

As disease spread in the experiment was not sufficient to test our group-level hypotheses, future work could adjust the experiment to overcome this issue. The number of decision opportunities, for example, could be limited, or some relations could be fixed, so that relations with infected neighbors persist longer. This would reflect relations that are maintained independent of the circumstances (e.g., family). Additionally, the visibility of disease states could be modified. That is, disease states could depend on the distance to the ego, with disease states of immediate neighbors the most likely to be revealed. Another option could be the implementation of a delay between acquiring an infection and revealing the infectiousness of a node to reflect that transmission may occur before symptom onset. Finally, the experiment might produce random transmission events independent of the network, resembling disease transmissions through short-term, non-personal contacts (e.g., public transport, grocery stores).

In conclusion, we find that despite the similarity in test environments, available actions, and behavioral mechanisms, small but significant differences between human participants and simulated agents have a strong impact on the course of our hypothetical epidemics. However, we do not consider a shortcoming of either method. It rather demonstrates

that the combined application of theoretical models and empirical studies provides better insights than either method alone, and should therefore find greater implementation in scientific research.

4.4 Methods

4.4.1 Experiment

4.4.1.1 Participants

A total of 48 experimental sessions were conducted between July 7 and July 22, 2021. All but one experimental sessions consisted of 60 participants resulting in a total of 2,879 participants. In the one session that could not be filled up completely, one node remained non-responsive and thus did not perform any actions. All but five participants, who reported technical difficulties (i.e., connection loss, glitches in the user interface), completed the entire study (2,874). Unresponsive nodes remained in the network, but did not initiate any relational changes. We assume that the nodes for this small number of unresponsive participants do not substantially affect behavior of other participants, nor the main dynamics of the game. Another 2,972 persons signed up for the study, but either did not show up for the corresponding session, or could not be assigned to a session of 60 participants. A session lasted between 24 and 80 minutes ($M = 47$, $SD = 12$). Participants earned between £5.00 and £8.79 ($M = \$5.29$, $SD = \$0.45$). Demographic data of participants is summarized in Table C.8 in the Supplements to Chapter 4.

4.4.1.2 Recruitment and compensation

Recruitment was carried out via *Prolific* (<https://www.prolific.co/>), an online participant recruitment platform for surveys and market research. We used a short registration study (<5 minutes) in which we described the design and background of the experiment, what was expected from participants, possible advantages and disadvantages of participating, the confidentiality of data processing, and provided contact information of the main researcher and independent contacts for comments and complaints about the study (see Section C.1.1 of the Supplements to Chapter 4).

Furthermore, participants were informed that they earn a minimum of £5.00 per 60 minutes of their participation as required by Prolific. Points earned during the experiment (see Section 4.4.1.3) were converted at the exchange rate of 500 points = £1.00, and paid as bonus if the amount exceeded the compensation for the time. Due to complaints about missing or incorrect bonus payments, we extended the instructions with a specific example after the first three sessions. The added example describes that a participant who earned 3,000 points during a 60 minutes long session earned £6.00: £5.00 for the minimum payment and £1.00 in addition as a bonus.

By clicking the “Sign up” button, participants declared that they had read all information about the experiment, that they agreed to participate at a specified time, and that they may quit the study at all times without explanation or consequences. Registered participants were invited personally to the experiment at the specified time using their anonymous Prolific ID. The experiment was approved by the Faculty Ethics Review Board (FERB) of the Faculty of Social and Behavioral Sciences of Utrecht University on June 14, 2021 (reference number 21-0210).

4.4.1.3 Design

The experiment followed a two-by-two mixed design. That is, two settings for clustering (low - *LO*; high - *HI*) that were performed within the same experimental session, and two settings for social mixing (random - *RA*; assortative regarding risk aversion - *AS*) that were performed between different experimental sessions. To avoid order bias, we alternated the initial clustering settings every experimental session and the social mixing settings every second experimental session. Consequently, the order of settings per session was as follows: session 1.1: *LO:RA*; session 1.2: *HI:RA*; session 2.1: *HI:RA*; session 2.2: *LO:RA*; session 3.1: *LO:AS*; session 3.2: *HI:AS*; session 4.1: *HI:AS*; session 4.2: *LO:AS*; session 5.1: *LO:RA*; etc.

Settings for clustering differed in the degree of clustering of the baseline network structures, while other properties were kept the same. That is, one network contained a single large cluster (*LO*: global clustering coefficient = 0.06, average degree = 5.93, closeness = 0.975), while the other contained several densely connected clusters (*HI*: global clustering coefficient = 0.62, average degree = 5.93, closeness = 0.955; see Figure 4.1).

Settings for social mixing differed in two aspects. First, participants were assigned to nodes depending on individual risk aversion score. To achieve randomly mixed networks (*RA*), participants were assigned to nodes so that neighbors mostly possessed different risk aversion scores. To achieve assortatively mixed networks (*AS*), participants were assigned to nodes so that neighbors mostly possessed similar risk aversion scores. Second, nodes for decision opportunities were selected based on risk aversion score. That is, to inhibit assortative mixing we offered nodes independent of risk aversion score (*RA*: $\omega = 0.0$); while to foster assortative mixing, we prioritized offered nodes that were the most similar regarding risk aversion score (*AS*: $\omega = 0.8$). As Figure C.6 in the Supplements to Chapter 4 shows, this approach was effective in both simulations and experiment.

Furthermore, we predefined the index cases (the initially infected node) for each condition so that the epidemic started with the most average node regarding degree, local clustering coefficient, and risk aversion score.

We then performed a series of simulated experiments (*simulation 1*) with the same setup as intended for the online experiment (48 sessions, 96 games, networks with 60 nodes) to determine parameter settings for homophily (*RA*: $\omega = 0.0$, *AS*: $\omega = 0.8$), infectivity ($\gamma = 0.15$), disease severity ($\sigma = 0.34$), recovery time ($\tau = 4$ rounds), number of offered nodes per round ($\phi = 0.2$), probability for each offered node to be a neighbor ($\psi = 0.5$), and probability for each offered node to be a neighbor of a neighbor ($\xi = 0.3$). Parameters were

selected so that differences in clustering and homophily produce epidemics differing in the number of (simultaneously) infected nodes and duration (see *simulation 1* in Figure C.4 in the Supplements to Chapter 4). Based on empirical data previously collected using the same task (Vriens & Buskens, 2020), risk perception parameters of the agents were set randomly using a probability distribution, based on data reported by a study using the same task ($M = 1.27$, $SD = 0.45$; Vriens & Buskens, 2020).² Based on these settings, agents were more likely to perceive high risks of infection and thus more likely to behave risk averse.

Finally, we rescaled utility to make point rewards more comprehensible. That is, rather than awarding $b_1 \cdot t_i - (c_1 \cdot t_i + c_2 \cdot t_i^2) = 1.0 \cdot 6 - (0.2 \cdot 6 + 0.067 \cdot 6^2) = 2.388$ points for the optimal number of 6 relations, we used a factor of 41.55 to award 100 points. The same applies to the reward for proportion of closed triads (e.g., optimal proportion for LO : $b_2 \cdot \left(1 - 2 \cdot \frac{|x_i - \alpha|}{\max(\alpha, 1 - \alpha)}\right) = 0.5 \cdot \left(1 - 2 \cdot \frac{|0.0 - 0.0|}{\max(0.0, 1.0 - 0.0)}\right) = 0.5 * 41.88 = 21$ points) and the penalty for being infected ($\sigma = 0.34 * 41.88 = 14$ points per round being infected).

4.4.1.4 Procedure

An experimental session was composed of three parts: a risk-aversion assessment using the *staircase task* (Part I), two installments of the NIDT (Part II), and a survey (Part III; see Figure 4.1). All parts were programmed in Elixir (Elixir Core Team, 2021), displayed using Phoenix (McCord, 2020), and performed using a web browser (see Figure 4.2). Participants started immediately with Part 1 after clicking a link sent via a personal message through the Prolific platform. Part 2 started when 60 participants have finished Part 1 and finished reading the instructions for Part 2. If after 20 minutes there were not enough participants ready to fill up a network game (≤ 50), all waiting participants were released and paid a show-up fee of £5.00.

The staircase task (Andersen et al., 2006; Falk et al., 2016) is a series of binary choices to determine the individual risk aversion score. In five consecutive rounds, participants were asked to choose between either a 50:50 chance to win 300 points vs. 0 points, or a guaranteed reward of a certain number of points.³ In the first round, the guaranteed reward was 160 points. The guaranteed reward for round two depended on whether a participant opted for the safe choice (80 points) or the gamble (240 points) in round one. According to the staircase task, people who repeatedly opt for the safe choice, although the rewards are dropping, are more risk averse than others who start to gamble at some threshold. After five rounds, participants end up at a position between 1 and 32 on the staircase, with 1 being the most and 32 the least risk averse. For an easier comparison with the theoretical model, we inverted and recoded the risk aversion score to a range between 0.0 and 2.0, with higher values denoting higher risk aversion and 1.0 representing risk neutrality.

²Note that the in Vriens and Buskens (2020) reported scores of $M = 19.7$ ($SD = 6.98$) were rescaled to the range used in the simulations (0.0-2.0).

³For the complete decision tree refer to Falk et al. (2016, p. 49, Figure 6)

After completing Part 1, participants were asked to read the instructions for the NIDT (see Section C.1.3 of the Supplements to Chapter 4). Once 60 participants finished reading the instructions and answered three questions to test their understanding of the rules, participants were assigned randomly to nodes in the network, and played three test rounds to familiarize themselves with the user interface (see Figure 4.2) and how to perform the task. Thereafter, the first NIDT started by assigning participants to nodes in the network according to the setting for social mixing and their risk aversion score (see Section 4.4.1.3).

Each round of the NIDT consists of four stages: decisions 1 (60 seconds to maintain/dissolve existing relations and propose new relations to other participants), decisions 2 (30 seconds to accept proposals to create new relations), computation of point rewards, and computation of disease transmissions. During the entire time, participants could click a link to open a pop-up window with condensed explanations of the reward system (see Sections C.1.4 and C.1.5 of the Supplements to Chapter 4). Part 2 concluded after two installments of the NIDT. Each NIDT concluded either once no more infected nodes were left, or after a maximum of 20 rounds. For more details on the NIDT, see Section 4.1 and Figure 4.1).

In Part 3, participants were asked to fill in a survey asking for age, gender, mother tongue, level of education, country of residence, COVID-19 concern, and whether tested positive for COVID-19 at some point (see Section C.1.6 of the Supplements to Chapter 4). By clicking a link to finish the experiment, participants were redirected to Prolific. After all participants finished, the session concluded.

4.4.2 Data and analysis

Data can be divided into three categories: participant data, network data, and decision data. Participant data contain data for each participant collected only once, consisting of the node IDs in the two networking games, risk aversion score, age, gender, mother tongue, education, residence, COVID-19 concern, and whether being tested positive for COVID-19 at some point. Network data describe the entire network at the beginning of each round in the form of an edge list (session ID, game ID, settings for clustering and social mixing, round number, ID of a node, disease state of that node, ID of a connected node). Decision data contain decisions on whether participants wanted to change relations with offered nodes or not (round number, node IDs receiving offers, node IDs of offered nodes, offer type (relation creation or dissolution), decisions on the relations). From these data, all further data (degree of clustering, homophily, etc.) were computed at the time of analysis.

All data were collected to prevent attribution to individual persons. That is, we received anonymized user IDs from Prolific for compensation purposes only. Furthermore, we stored participant data with our own anonymized user IDs that allowed to match them with network and decision data. All personal information, such as age, gender, education, etc. was provided by the participants actively by filling in the questionnaire at the end of the experiment.

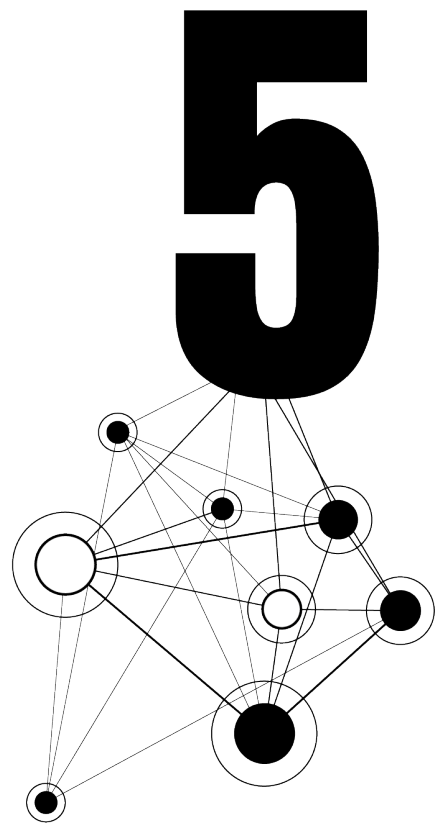
We define the *final size* of an epidemic as the proportion of infected and recovered nodes at the end of a network game. Due to violation of the normality assumptions, we used Wilcoxon rank-sum and Kruskal-Wallis tests to test whether final size differed significantly between the settings for clustering and social mixing. Clustering was computed as the average proportion of closed triads over all possible closed triads per node. In accordance with the concept of degree-based assortative mixing (Newman, 2002), homophily was defined as the Pearson correlation coefficient of risk aversion score between all pairs of connected nodes. Closeness is defined as the reversed and normalized average distance between any two nodes in the network (Buechel & Buskens, 2013, p.163).

To understand the decision-making process, we compare the proportions of decisions made to decision opportunities. That is, a node offered (*alter* – either by the system in stage 2 or by another participant in stage 3 of a round) constitutes an opportunity for the receiving node (*ego*) to make a networking decision (create, maintain, dissolve). The decision by the ego to *Create* a relation to an alter is therefore defined as an accepted opportunity to create the corresponding not yet existing relation in both stages 2 and 3. The decision by the ego to *Maintain* a relation to an alter is a declined opportunity to dissolve that relation in stage 2. Finally, the decision by the ego to *Dissolve* a relation to an alter is an accepted opportunity to dissolve that relation in stage 2.

Furthermore, we perform a four-level random intercept logistic regression (level 4: 96 games, level 3: 5,758 nodes, level 2: 52,650 rounds, level 1: 711,159 decisions) on whether a relation is desirable or not. That is, a relation is desirable if the opportunities to create and maintain a relation are accepted. Consequently, a relation is undesirable if the opportunity to create a relation is declined and the opportunity to dissolve a relation is accepted. Additionally, we performed a logistic regression analysis on whether a node got infected during one installment of the NIDT. Thus, results indicate risk factors for acquiring infections. Standard errors were made robust using clusters of 2,879 participants and two observations (NIDTs) per participant. We mean centered risk aversion score and consider linear relationships for predictors in all regression analyses.

4.4.3 Availability of data and materials

The data, the Java 8 source code to generate the data (including an executable program and an easy to use graphical user interface), and the R scripts to analyze the data during the current study are available under the GPLv3 license on GitHub (Nunner, 2020). The experiment data are available upon request.



Chapter 5

Prioritizing high-contact occupations raises effectiveness of vaccination campaigns

Abstract *A twenty-year-old idea from network science is that vaccination campaigns would be more effective if high-contact individuals were preferentially targeted. Implementation is impeded by the ethical and practical problem of differentiating vaccine access based on a personal characteristic that is hard-to-measure and private. Here, we propose the use of occupational category as a proxy for connectedness in a contact network. Using survey data on occupation-specific contact frequencies, we calibrate a model of disease propagation in populations undergoing varying vaccination campaigns. We find that vaccination campaigns that prioritize high-contact occupational groups achieve similar infection levels with half the number of vaccines, while also reducing and delaying peaks. The paper thus identifies a concrete, operational strategy for dramatically improving vaccination efficiency in ongoing pandemics.*

This chapter has been published as Nunner, H., van de Rijt, A., & Buskens, V. (2022b). Prioritizing high-contact occupations raises effectiveness of vaccination campaigns. *Scientific Reports*, 12(1), 1–13. Nunner and Buskens developed the mathematical model. Nunner programmed the simulation, performed model calibration, data generation, data analyses, and produced the figures. Nunner and van de Rijt wrote the manuscript, while all authors provided feedback.

5.1 Introduction

Today, two years after its global outbreak, sustained propagation of COVID-19 continues to kill thousands of people a day, inflict economic damage and allows mutations to emerge that may require the development of new vaccines. The challenge this poses is how to effectively control the virus with limited means. In this article, we draw on social network analysis to propose an easily implementable strategy for more effectively vaccinating a population.

Two decades ago, network scientists showed that in theory, the targeting of highly connected individuals should be an effective vaccination strategy when propagation networks exhibit high variability in connectivity across nodes (Dezső & Barabási, 2002; Pastor-Satorras & Vespignani, 2002). Recent studies that calibrate models using data on close-range contact frequencies have demonstrated that also in the case of COVID-19, the prioritizing of individuals with many close-range contacts would dramatically increase the effectiveness of vaccination campaigns (Herrmann & Schwartz, 2020; Manzo & van de Rijt, 2020). This is because short-range physical contact is highly unequally distributed across individuals. “Hubs” have many more contacts than other individuals, and according to contact diaries, these contacts are not shorter-lived (Manzo & van de Rijt, 2020). As a result, hubs are not just more likely to get infected, but once infected, they also pass it on to more others.

Despite its promise, this strategy has remained mostly a theoretical idea (Rocha et al., 2011). A key obstacle to implementation is the ethical and practical problem of differentiating vaccine access based on a personal characteristic that is hard-to-measure and private. How does one identify high-contact individuals so that they can be targeted in vaccination campaigns? We propose the use of occupational groups as a proxy for the number of close-range contacts in a contact network. Differentiating COVID-19 policy interventions on the basis of individuals’ occupations is executable. Indeed, it has already been part of public policy in many countries, both in social distancing legislation and vaccine access, except that prioritization was not based on network analysis.

For this approach to be effective, there must be significant variability in close-range exposure between individuals working in different occupations. Clearly, occupational group is imperfect as a proxy, as people with the same job can still vary greatly in the number of short-range contacts they have. In this paper, we draw on data from a recent survey conducted at the beginning of the COVID-19 pandemic in early 2020 (Belot et al., 2020) that combines detailed occupational codes with measures of close-range contact. The survey covers six countries – China, South Korea, Japan, Italy, UK, and US – and is nationally representative of each by age, gender, and income. Information is available on contact at *under 1 meter distance* prior to the COVID-19 pandemic, as well as on such short-range contact during the first lockdown in Spring 2020. The data reveal substantial occupational differences, with teachers and cashiers being among the most connected and computer programmers among the least connected. To investigate whether this variability can produce significant gains when exploited in targeted vaccination programs, we used the data

in two ways. First, we generated networks that have degree distributions calibrated with occupational contact data. Second, we simulated epidemics and compared the effectiveness of vaccination campaigns targeting individuals randomly or targeting occupational groups with the highest average number of social contacts.

5.1.1 Related work

Since the outbreak of COVID-19 in early 2020, a plethora of scientific studies has been published from a wide variety of disciplines, such as medical sciences (Hoffmann et al., 2020; Singh & Singh, 2021; Sun et al., 2020), artificial intelligence and machine learning (Laguarta et al., 2020; Pang et al., 2021), social sciences (Brethouwer et al., 2021; Elmer et al., 2020; Lo Iacono et al., 2021; Settersten et al., 2020), psychology (Elbay et al., 2020; Ustun, 2021), economy (Altig et al., 2020; Borio, 2020; McKibbin & Fernando, 2020), and food and agriculture (Altieri & Nicholls, 2020; Siche & Siche, 2020). Our model builds in particular on a long strand of so-called compartmental models. The origins of compartmental models in epidemiology (Brauer, 2008), for example, go back more than a century. These models divide a population into different compartments representing disease states (e.g., susceptible, infected, recovered) and define how to progress from one compartment to another. They are a powerful tool for predicting the possible course of epidemics and the effectiveness of countermeasures (Flaxman et al., 2020; Kretzschmar et al., 2020; Kumari et al., 2021; Matrajt & Leung, 2020; Teslya et al., 2020). Compartmental models have also been used for the simulation of diffusion processes on social networks, such as disease spread (Danon et al., 2011; Nunner et al., 2021), information spread (Liu et al., 2020), and their interplay (Zhan et al., 2018). Our model likewise explicitly simulates the diffusion of an infectious disease in a social network.

Recent studies have proposed numerous network interventions (Valente, 2012) for reducing the propagation of COVID-19. Some interventions seek to strategically restrict close-range contact to occur only within predetermined interaction structures so that the speed and reach of COVID-19 spread can theoretically be greatly reduced (Block et al., 2020; Nishi et al., 2020). However, even severe social distancing policies such as full-scale lock-downs can only temporarily reduce infections and hospitalizations (Guzzetta et al., 2020; Hyafil & Moriña, 2020; Kaur et al., 2020; Lau et al., 2020; Santamaría & Hortal, 2020; Thu et al., 2020; VoPham et al., 2020), leaving large-scale vaccination as the primary vehicle for sustainable control over the SARS-CoV-2 virus. Highly effective vaccines are being mass-distributed and evidence is mounting that vaccinations do not just prevent severe cases but also greatly reduce infection (Amit et al., 2021; Goldberg et al., 2021; Hall et al., 2021; Leshem & Lopman, 2021; Levine-Tiefenbrun et al., 2021; Mallapaty, 2021; Tande et al., 2021; Thompson et al., 2021; Tran Kiem et al., 2021). Nonetheless, global vaccine roll-out has logistical and financial limits. We study a different kind of network intervention that seeks to minimize resources needed to achieve a certain level of epidemic control by strategically making use of network properties. Specifically, we research the prioritizing of occupational groups with workers exposed to close-range contact with large numbers of individuals.

For many diseases spreading through close-range contacts, evidence has accumulated that a small fraction of source individuals is responsible for most infections (James et al., 2007; Lin et al., 2020; Little et al., 2014; Lloyd-Smith et al., 2005; Stein, 2011; Sun et al., 2014; Wong et al., 2015). It is estimated that for COVID-19, between 10% and 20% of infected individuals produce 80% to 90% of new cases (Adam et al., 2020; Bi et al., 2020; Endo et al., 2020; Hamner, 2020; Kay, 2020; Miller et al., 2020). This suggests that if one could somehow identify and protect the minority of spreaders, the virus may be controlled through focused interventions at lower overall cost. While the mechanisms that underlie interpersonal variability in infectiousness are poorly understood (Cho et al., 2016; Galvani & May, 2005; Woolhouse et al., 1997), it is self-evident that the more others one exposes to a given intensity and duration of short-range contact, the larger the number of new cases that one generates. One may suspect a trade-off between the number of close-range contacts and the length of such contact. Then, if the infection probability were increasing in contact length, hubs would not play a relatively less critical role. However, data from contact diaries suggest that those who meet only a handful of people on one day do not expose these others for a longer period of time than those who meet dozens of distinct people on one day (Manzo & van de Rijt, 2020), reinforcing the strategic value of targeting hubs in contact networks for vaccination and other forms of infection prevention.

5.2 Simulation model

Networks of 10,000 nodes were generated using a network formation model (Nunner et al., 2021) that allows control of degree (for details, see Section 5.5). A genetic algorithm was used to fit the average degrees per major occupational group (according to the SOC codes from the US bureau of labor statistics) reported for times prior to the epidemic. Table 5.1 shows the numbers for reported mean degrees by major occupational group (Belot et al., 2020). Recent US labor market numbers were taken to set occupational group size (U.S. Bureau of Labor Statistics, 2019). *Office and Administrative Support Occupations*, for example, were the largest group containing 12.74% of the entire labor market, and thus our generated networks included the same percentage of nodes for this occupational group. A cross-sectional survey study on social contacts in GB among 5000 respondents (Danon et al., 2013) estimated the average proportion of closed triads in contact networks to be about 0.46, while reducing with age. Accordingly, we varied clustering in a range around that value, at 0.3, 0.4, and 0.5. Occupational group homophily was varied to cover scenarios with no (0.0), medium (0.4), and high (0.8) probabilities of ties between nodes from the same occupational group.

For each combination of clustering (3 values) and homophily (3 values), we selected the 10 best fitting networks. Based on these 90 normal networks, another 90 lockdown networks were generated by severing ties between nodes. Ties were severed based on the average contact number reduction reported for the two connecting nodes' occupational groups (Table 5.1). This procedure lasted until the empirical average degrees of the occupational groups were achieved. Consequently, we ended up with a total number of 180

Table 5.1: Mean degrees per major occupational group for empirical networks at time points prior to (normal) and during the first COVID-19 lockdown in Spring 2020.

	Normal	Lockdown
Healthcare Practitioners and Technical Occupations	20.17	5.19
Personal Care and Service Occupations	12.82	4.58
Educational Instruction and Library Occupations	12.76	2.61
Legal Occupations	8.28	0.92
Management Occupations	5.84	1.32
Sales and Related Occupations	5.59	3.3
Healthcare Support Occupations	5.44	3.7
Food Preparation and Serving Related Occupations	5.32	1.57
Transportation and Material Moving Occupations	5.31	2.69
Life, Physical, and Social Science Occupations	4.63	3.55
Office and Administrative Support Occupations	4.42	1.94
Building and Grounds Cleaning and Maintenance Occupations	3.93	1.26
Installation, Maintenance, and Repair Occupations	3.76	2.83
Business and Financial Operations Occupations	3.66	1.52
Construction and Extraction Occupations	3.6	1.99
Architecture and Engineering Occupations	3.47	1.88
Arts, Design, Entertainment, Sports, and Media Occupations	3.23	2.51
Production Occupations	2.87	2.91
Computer and Mathematical Occupations	2.85	1.21
Community and Social Service Occupations	2.84	1.08
Unemployed	2.34	0.96
Farming, Fishing, and Forestry Occupations	2.15	1.76
Retired	2.13	0.87
Protective Service Occupations	1.12	1.05

networks. Detailed descriptive statistics on network composition and fitting of average degrees to occupational groups can be found in respectively Tables D.1 and D.2 – D.5 in the Supplements to Chapter 5.

To assess the effectiveness of different vaccination campaigns, we simulated epidemics under three different conditions shown in Figure 5.1. In the *baseline* condition (a), no vaccinations were given, and thus all nodes remained susceptible. The baseline condition therefore provides a benchmark for judging the effect of vaccination campaigns on epidemics. The two vaccination campaigns differ in the way nodes are selected for the administration of vaccinations. In the *random* condition (b), randomly selected nodes were vaccinated irrespective of occupational group membership. In the *targeted* condition (c), nodes were vaccinated based on occupational group membership and in descending order of the reported average number of social contacts (i.e., 1. Health Practitioners and Technical Occupations, 2. Personal Care and Service Occupations, 3. Educational Instruction and Library Occupations, etc. in the normal networks; and 1. Health Practitioners and Technical Occupations, 2. Personal Care and Service Occupations, 3. Healthcare Support, etc. in the lockdown networks; see Table 5.1).

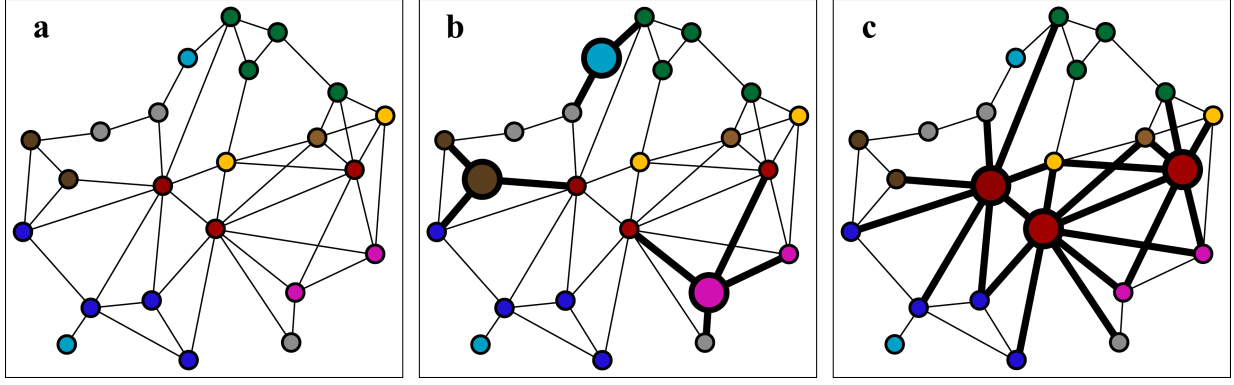


Figure 5.1: **Vaccination campaign scenarios.** (a) Baseline scenario without vaccinations, (b) random distribution of vaccines, and (c) targeted distribution of vaccines based on occupational group membership in descending order of average number of social contacts. Nodes represent individuals and colors represent occupational group membership (e.g., red: *Healthcare Practitioners and Technical Occupations* with high average degree, brown nodes: *Office and Administrative Support Occupations* with low average degree). Enlarged nodes represent immunized individuals. Thick ties represent social connections of immunized nodes, and are therefore ruled out as possible transmission routes.

A flowchart for how we simulated disease transmission and intervention is given in Figure 5.2. In addition to varying the input networks (normal vs. lockdown, clustering, occupational group homophily), we varied the availability of vaccines for the population (5%, 10%, 20%, 30%, 40%, 50%) and vaccine effectiveness in terms of probability of immunization (0.60, 0.75, 0.90). Note that the latter entails both the probability to prevent sickness and the probability that the disease is spread further. 20 simulation runs were performed for each of the 180 networks, 3 vaccination campaign conditions, and 6 vaccine availability percentages as well as 3 vaccine effectiveness controls (non-baseline conditions only), resulting in a total number of 133200 simulated epidemics.

Each simulation run was initiated by distributing vaccinations to an entirely susceptible population. Whether a node was immunized depended on whether the node was selected for vaccination (vaccination campaign condition and vaccine availability) and whether the vaccination was successful (vaccine effectiveness). In contrast to immunized nodes, unsuccessfully vaccinated nodes remained susceptible. Note that we neglect the possible spread of the disease through vaccinated people who do not develop symptoms.

In a second step, a randomly selected node was infected (*index case*). The remainder of a simulation run consisted of discrete time steps to compute disease transmission between infectious and susceptible nodes and recovery events of infected nodes. That is, whether a node i got infected depended on the probability of disease transmission per single contact ($\gamma = 0.15$) and the number of infectious contacts of node i (n_{i_I}):

$$\pi_i = 1 - (1 - \gamma)^{n_{i_I}}. \quad (5.1)$$

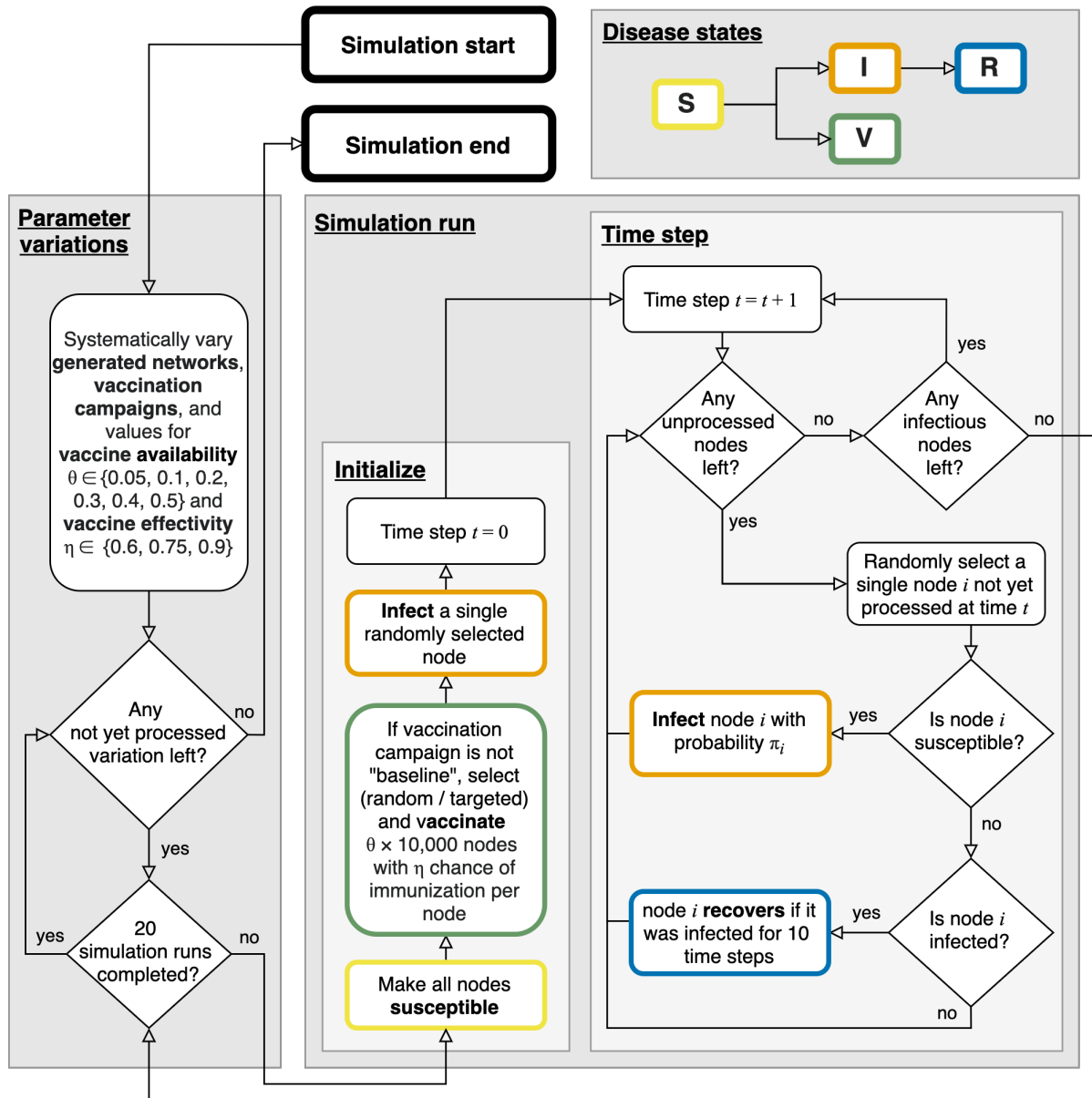


Figure 5.2: Flowchart of the simulation.

Infectious nodes recovered after 10 time steps and could not get infected a second time. A simulation run ended when no infectious nodes were left. Detailed descriptive statistics on network compositions, index cases, epidemics, and counter measures can be found in Table D.1 in the Supplements to Chapter 5.

5.3 Results

5.3.1 Comparison of vaccination campaigns

Figure 5.3 shows across all simulation runs, separately for each vaccination scenario, the distributions of two commonly studied measures of epidemic control: (a) final size and (b) peak size. Final size is the percentage of nodes that have been infected over the entirety of a simulated epidemic. Peak size describes the maximum percentage of simultaneously infected nodes per epidemic. To increase the resolution of differences between the conditions, the inset of plot (b) shows peak sizes in epidemics involving a minimum of at least two simultaneously infected nodes. Each plot shows the relative frequencies of one measure per vaccination campaign (baseline – red, random – yellow, targeted – blue). Dashed lines depict mean values, and solid lines median values per vaccination campaign.

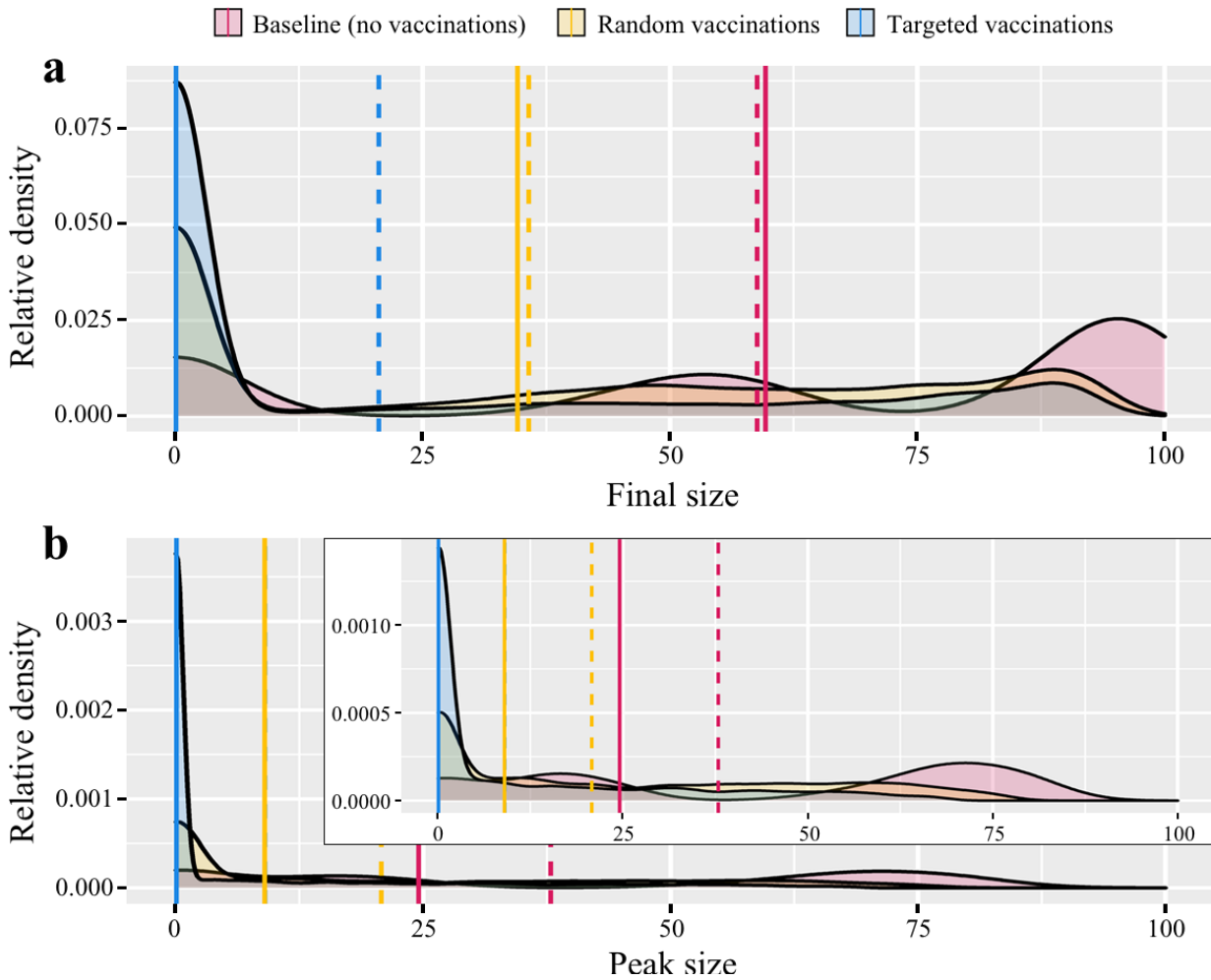


Figure 5.3: **Densities of final and peak size of epidemics.** Relative densities of final size (a) and peak size (b) of epidemics by vaccination campaign. Both measures are reported as percentage of the population (final size: percentage of cumulative infected nodes; peak size: maximum percentage of simultaneously infected nodes). Dashed lines show mean, solid lines show median values.

Our simulations suggest that independently of network composition, vaccine availability, and vaccine effectiveness, targeting high contact occupations is much more effective than random distribution of vaccines. While the random campaign produces average final sizes of about 35% infected nodes (both mean and median), the targeted campaign produces mean final sizes of only 20% infected nodes and even prevents most epidemics entirely (median final size close to 0%). A similar picture is drawn regarding peak size. While random distribution of vaccines produces epidemics with about 10% (median) to 20% (mean) simultaneously infected nodes, the majority of epidemics in the targeted distribution do not show notable peaks at all (median close to 0%).

Table 5.2 shows that, independent of input parameter variations, targeted distribution of vaccines is substantially more effective in reducing infection numbers than random distribution. Specifically, targeted distribution of vaccines achieve on average the same reduction of infection numbers with only *half* the number of vaccines. That is, targeted vaccination of 5% of the population reduces final size by 11%, while the same reduction in the random condition requires 10% vaccine coverage. The same applies to 10% and 20% vaccine availability in the targeted condition, which require 20% and 40% vaccine coverage in the random condition, respectively.

The effect of vaccine availability is even more striking for the lockdown networks (lower half of Table 5.2). That is, targeted vaccination of 20% of the population can bring the epidemic to a halt (reduction of 26.02 percent points from 26.64 in the baseline condition). To achieve the same with random distribution of vaccines, two-and-a-half times the number of vaccines is required, as about half of the population would need to be vaccinated. In sum, with or without lockdown, targeted vaccination greatly improves campaign effectiveness.

A similar picture is drawn for vaccine effectivity. For both network types (no lockdown, lockdown), and all other parameters considered equal, targeted distribution of vaccines reduces the number of infections significantly stronger than random distribution of vaccines.

These main results are broadly robust across network parameters. Table 5.2 shows effects of network composition on the final size of epidemics. Clustering has been divided into three categories: networks at the lower end (0.4), the upper end (0.6), and in the middle (0.5) of the clustering range reported for contact networks (Danon et al., 2013). The simulations show that while targeted intervention is more effective at all clustering levels, the more clustered the network, the greater the relative gains vis-à-vis random. Occupational group homophily is also divided into three categories. The *no homophily* category contains the networks without consideration of occupational group for tie creation. The *medium homophily* category contains networks that used a probability of 0.4, while the *high homophily* category contains networks that used a probability of 0.8 for the creation of ties between nodes from the same occupational group. Because of the redundancy of occupation-targeted intervention under high occupational homophily, we find that when homophily is increased the effectiveness of targeted versus random intervention is somewhat reduced, by 7 percent points.

Table 5.2: Mean final size of baseline condition (2nd column) and difference by test condition in percent points.

	Baseline	Random to Baseline	Targeted to Baseline	to Random
<i>No lockdown</i>				
Overall	90.97	−30.19	−52.39	−22.20
Vaccine availability 5%		−5.73	−11.27	−5.54
Vaccine availability 10%		−11.28	−25.25	−13.97
Vaccine availability 20%		−22.50	−47.71	−25.21
Vaccine availability 30%		−35.27	−67.31	−32.04
Vaccine availability 40%		−47.13	−78.67	−31.54
Vaccine availability 50%		−59.23	−84.13	−24.90
Vaccine effectivity 60%		−23.75	−40.32	−16.57
Vaccine effectivity 75%		−30.17	−53.80	−23.62
Vaccine effectivity 90%		−36.65	−63.05	−26.40
Clustering - low	95.11	−28.91	−47.88	−18.97
Clustering - medium	91.98	−29.67	−50.79	−21.12
Clustering - high	83.59	−32.84	−61.31	−28.47
Homophily - low	89.16	−30.77	−56.05	−25.27
Homophily - medium	95.16	−27.17	−45.84	−18.67
Homophily - high	92.90	−29.87	−48.17	−18.30
<i>Lockdown</i>				
Overall	26.64	−15.98	−24.11	−8.13
Vaccine availability 5%		−3.95	−16.01	−12.06
Vaccine availability 10%		−7.51	−22.87	−15.36
Vaccine availability 20%		−14.91	−26.02	−11.10
Vaccine availability 30%		−20.14	−26.55	−6.42
Vaccine availability 40%		−23.77	−26.59	−2.82
Vaccine availability 50%		−25.58	−26.61	−1.03
Vaccine effectivity 60%		−13.97	−22.77	−8.80
Vaccine effectivity 75%		−16.21	−24.39	−8.18
Vaccine effectivity 90%		−17.75	−25.16	−7.42
Clustering - low	35.33	−20.46	−30.70	−10.23
Clustering - medium	25.43	−15.36	−23.48	−8.13
Clustering - high	17.60	−11.30	−16.81	−5.51
Homophily - low	23.62	−14.03	−21.47	−7.44
Homophily - medium	33.57	−18.89	−28.82	−9.93
Homophily - high	29.06	−18.19	−26.79	−8.60

Note: Raw numbers are provided in Table D.6 in the Supplements to Chapter 5.

5.3.2 Dynamics of epidemics

Figure 5.4 visualizes the effects of vaccine availability and type of distribution on the temporal progression of epidemics. Depicted is a comparison of the average course of epidemics between the three vaccination campaign conditions (columns: baseline, random, targeted) and the availability of vaccines (rows). Each plot shows colored lines for the median proportion

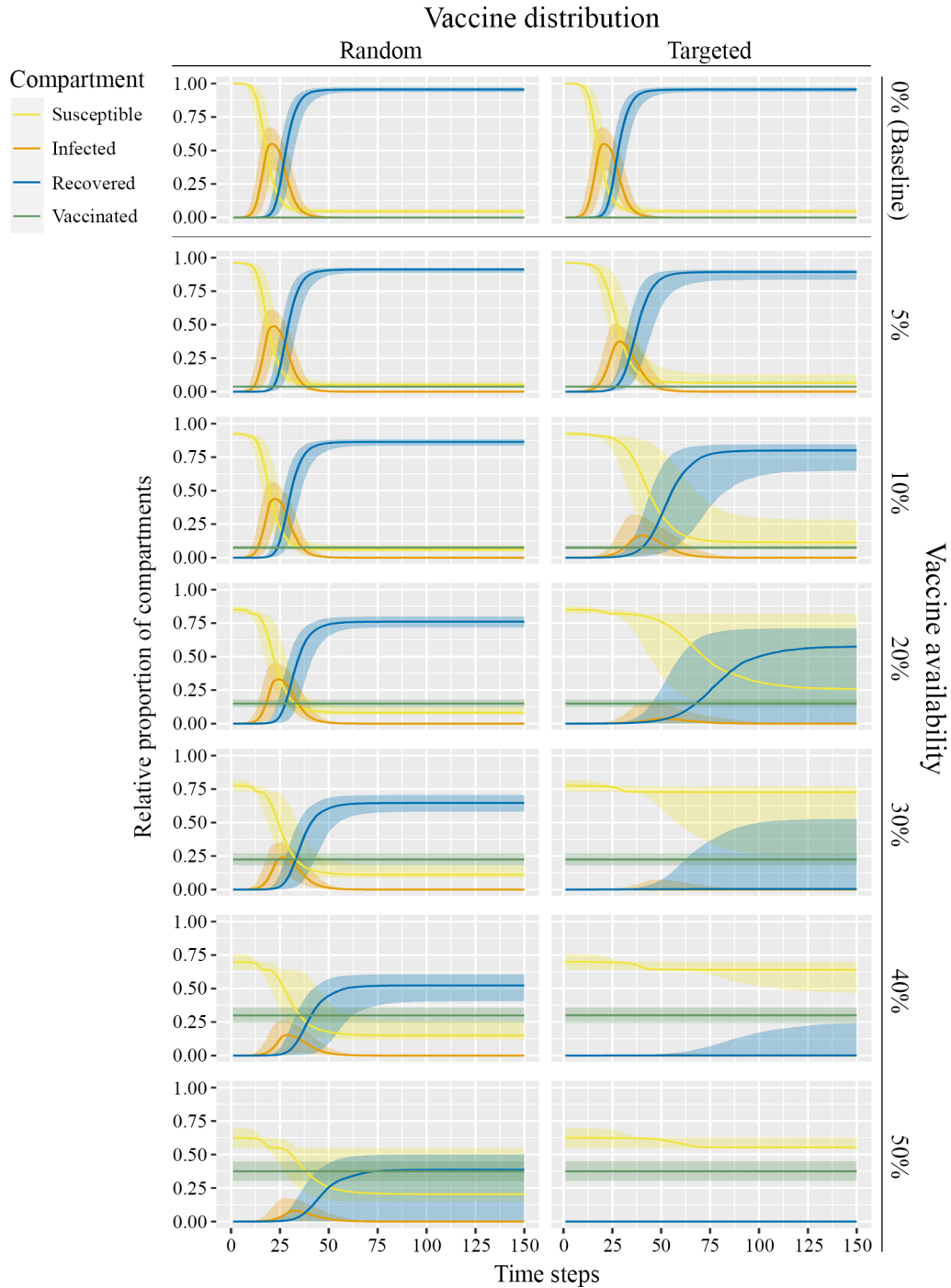


Figure 5.4: **SIRV plots for vaccine availability per vaccination campaign condition.** Solid lines show median proportion (y-axis) of susceptible (yellow), infected (orange), recovered (blue), and successfully vaccinated / immunized (green) agents over the first 150 simulated time steps. Ribbons show interquartile range.

portion of susceptible (yellow), infected (orange), recovered (blue), and vaccinated (green) nodes over time. Ribbons show the variability (interquartile ranges) of these results. As before, we find that with only half of the vaccines the same reduction of final size can be achieved through prioritization of high-contact occupations: When comparing the plots for 20% vaccine availability in the targeted distribution condition (column 3, row 3) and 40% vaccine availability in the random distribution condition (column 2, row 5) we observe a similar end point for the number of recovered nodes (i.e., final size). Additionally, we can see that the shape of the curve of infected nodes differs. That is, the peak in the targeted distribution condition is at an average of 15.22%, while the peak in the random distribution condition is at an average of 23.60% (see Table D.7 in the Supplements to Chapter 5). Furthermore, the epidemic in the random distribution condition requires on average 71.81 time steps from the first infection until the last infected node recovers. In the targeted distribution condition, a similar number of infection and recovery events occur in 101.85 time steps (see Table D.9 in the Supplements to Chapter 5), thus creating a flatter curve of infected nodes.

In summary, our simulations suggest that distribution of vaccines prioritizing societal groups with high contact rates (here: occupations) effectively and consistently reduces final size, peak size, and slow down the spread of a disease, largely independently of input parameter variations. That is, while similar numbers of infections can be achieved with only half the number of vaccines, targeted vaccinations can also *flatten the curve* more effectively than random distribution of vaccines.

5.4 Discussion

These results bring the theoretical strategy of targeting people with many close-range contacts in vaccination campaigns one step closer to real-world implementation. We have long known from theoretical diffusion studies that targeting interventions at hubs in social networks should reduce the spread of infectious diseases that are passed on through person-to-person contact (Dezső & Barabási, 2002; Pastor-Satorras & Vespignani, 2002). Furthermore, recent model studies suggest that prioritizing individuals with many close-range contacts in the case of COVID-19 would dramatically increase the effectiveness of vaccination campaigns (Herrmann & Schwartz, 2020; Manzo & van de Rijt, 2020). Although all these studies, including our own, promise large efficiency gains through network targeting, they are hardly used for disease control. This is because implementation of network targeting has been impeded by the issue of how to identify people with many contacts in a practical manner. Deriving contact data from digital trace data poses privacy concerns. Random neighbor sampling (Cohen et al., 2003; Kim et al., 2015), whereby randomly nominated friends are vaccinated, may also work because of the statistical tendency for one’s friends to have more friends than oneself has (Feld, 1991). However, the lack of familiarity with selecting recipients of a life-saving public health intervention through such unorthodox methods may raise practical, legal, and ethical challenges.

By contrast, policies that intervene based on target individuals' occupations are already used widely. During the COVID-19 pandemic in various countries, certain occupations saw mandatory closures, while other sectors of the economy were allowed to remain open. In the vaccination campaigns in most Western countries, certain occupations, such as doctors and teachers, were given preferential access to vaccines. The Centers for Disease Control and Prevention (CDC) already recommends the prioritization of *essential workers* who "ensure the continuity of critical functions in the United States" with those higher in order who have higher risks of exposure to SARS-CoV-2 ("Vaccines and immunizations - Interim list of categories of essential workers mapped to standardized industry codes and titles", 2021). Similar prioritization can be found in other countries' vaccination strategies around the world. Many strategies, however, do not consider prioritization of occupational groups beyond essential occupations or healthcare occupations.

In this paper, we claim that expanding vaccination strategies to focus more on the number of contacts per occupational group, can reduce the final size of an epidemic by about 20% compared to strategies neglecting contact rates (everything else equal). In some scenarios, targeted vaccination requires only half the number of vaccines of non-targeted vaccination to achieve similar final sizes with later and lower epidemic peaks. Furthermore, the paper has shown that the positive effect of targeted vaccination is independent of vaccine effectivity and can be increased by stronger network clustering and lower occupational group homophily. Thus, maintaining social distancing policies to limit social contacts to close family and friends (increasing clustering) and working remotely (decreasing occupational group homophily) during targeted vaccination roll-out may help to bring an epidemic to a quicker end.

Currently, the choice for prioritizing people with certain occupations is often made based on how important these occupations are for society and the vulnerabilities individuals in these occupations are exposed to. Furthermore, age and preexisting conditions that increase the vulnerability of specific people play an important role in the planning of COVID-19 vaccination strategies. We do not claim that the maintenance of essential societal functions and the protection of the most vulnerable should not be given the highest priority. We do claim, however, that it is important to also consider the number of relations people in an occupation have and thus their role in spreading an epidemic further. That is, people with many relations are not only more likely to get infected themselves, but are responsible for causing larger numbers of secondary cases. Immunizing persons in high contact occupations therefore has the potential to increase effectiveness of vaccination campaigns, as more people can be (indirectly) saved from infections. This becomes all the more important when vaccine availability is limited. Moving forward, an important question is how to integrate contact numbers into existing vaccination strategies. Based on our findings, we propose that occupational groups can function as a reasonably effective proxy for such an extension.

Our study has several important limitations. First, while we were able to reproduce the mean measured contacts per occupation based on the available data on these occupations, these are only estimates and for some occupations based on relatively few observations.

Moreover, small sample sizes for single occupations and minor occupational groups forced us to use major occupational groups. While this approach allows an easy-to-implement strategy for policy design, it neglects, however, that the number of relations may vary between occupations within the same group. In particular, we see that the variance in numbers of relations within occupations is larger in the empirical data than in the simulated networks. The reason is that we focused on matching average numbers of relations with our network generation procedure. Furthermore, matching variance is not straightforward. In addition, given the numbers of observations and the numbers of the relations reported in the data, it can be inferred that the variances are not well estimable based on the empirical data and outliers might cause variances to seem larger than they actually are.

The data we used also showed that differences in contact numbers between countries exist. Average contact numbers in the UK before the pandemic (9.48 contacts per day), for example, were almost 3.5 times as large as for Korea (2.77 contacts per day). Our results therefore show an average effect that may differ between countries. Furthermore, we do not consider age structures or household compositions. Thus, results are focused on occupational networks only, neglecting the role children may play in the continuation of an epidemic. Lastly, we do not take advantages or disadvantages of strategies into account that go beyond the studied network effects. Prioritizing specific job categories, for example, may have additional effects on the course of an epidemic (e.g., medical sector, tourism industry).

While we are confident that the main result of a large efficiency gain from targeted vaccination is robust, this gain may differ in magnitude when applied to specific diseases. To adapt the model to COVID-19, for example, some details, such as probability of transmission per contact, recovery times, and finer resolution of compartments, need to be modelled more precisely. Furthermore, we assume that successfully vaccinated people cannot get, nor spread the disease anymore. Although it is still under discussion to what extent vaccinated people remain potential transmitters of COVID-19 and thus reduce the effect of targeted vaccine distribution, it is now clear that vaccinations do not just prevent severe cases but greatly reduce infection (James et al., 2007; Lin et al., 2020; Little et al., 2014; Lloyd-Smith et al., 2005; Stein, 2011; Sun et al., 2014; Wong et al., 2015). These effects, however, will depend on viruses and variants that emerge and the vaccines that are developed to combat them.

In conclusion, when severe social distancing measures do not suffice, vaccination remains the most effective weapon against an epidemic outbreak. Although we have long known that immunizing high contact individuals can reduce the spread of infectious diseases, these measures have hardly been used for disease control. Our study suggests that using high-contact occupations as a readily available proxy for targeted vaccination campaigns can significantly increase the effectiveness of vaccine roll-out, while avoiding some pitfalls impeding implementation (e.g., privacy concerns, practical, legal, and ethical challenges).

5.5 Methods

The methods presented here provide a higher-level overview. The aim is to promote understanding of the simulation procedures and to enable putting the results into context. For a more detailed and formal description of the methods, including all equations, pseudocode of algorithms, and parameter settings, please refer to Section D.2 of the Supplements to Chapter 5.

5.5.1 Network formation model and simulation

The networks used as input for the simulation of epidemics and the simulation of epidemics are based on a specific model case (Nunner et al., 2022a) of the *Networking during infectious diseases model (NIDM)* (Nunner et al., 2021). The network formation model is based on the idea that social ties provide utility (Jackson, 2008; Jackson & Wolinsky, 1996). Furthermore, this utility can be maximized by changing the position someone takes in the network. Utility in the NIDM describes personal well-being from the perspective of each node i and is the difference between benefits of social ties (social capital, comfort, sense of belonging, etc.) and costs to maintain these ties (time, effort, etc.):

$$U_i = b_1 \cdot t_i + b_2 \cdot \left(1 - 2 \cdot \frac{|x_i - \alpha|}{\max(\alpha, 1 - \alpha)} \right) - c_1 \cdot t_i + c_2 \cdot t_i^2. \quad (5.2)$$

b_1 is the immediate benefit for the number of ties t_i and is discounted by the immediate costs c_1 and the marginal costs c_2 for the number of ties. Note that variations of c_2 , while keeping parameters constant (here: $b_1 = 1.0$ and $c_1 = 0.2$), allow controlling the optimal number of ties per node. A setting of $c_2 = 0.05$, for example, translates into an optimum of 8 ties per node, while $c_2 = 0.1$ creates an optimum of 4 ties per node. The benefit of network positions is furthermore dependent on whether the actual proportion of closed triads x_i matches the optimal proportion of closed triads α and how much weight b_2 there is on this part of the equation. Consider, for example, node A having three ties (AB, AC, AD), which implies three possible closed triads node A is part of (AB-AC-BC, AB-AD-BD, AC-AD-CD). Furthermore, consider α set to a value of 0.33. It follows that the optimum for node A would be to have only one of the three possible closed triads.

Agent-based simulations (Nunner, 2020) were used to generate the networks based on the data with normal (prior to lockdown) degree distribution. Starting from an empty network, agents maximize individual utility based on Equation 5.2 by either creating or severing ties to other agents. Simulation parameter ω allows furthermore to control the proportion of ties between agents from the same occupational group. That is, corresponding to the concept of baseline homophily (McPherson et al., 2001), the simulation provides more opportunity to meet for agents sharing similar traits the higher the setting for ω .

5.5.2 Calibration of network structure with empirical data

Networks of 10,000 nodes were generated using the NIDM simulation and fitted to empirical data using a genetic algorithm and a lockdown generation algorithm. Empirical data to define target values of the generated networks consisted of three sources. First, employment numbers reported by the U.S. Bureau of Labor Statistics (U.S. Bureau of Labor Statistics, 2019) for major occupational groups according to the *Standard Occupational Classification (SOC)* system (U.S. Bureau of Labor Statistics, 2018). These data were used to assign a major occupational group to each agent, with a probability according to the group's proportional size. Second, mean degree per major occupational group collected in a six-country survey on COVID-19 and reporting contact numbers before the pandemic and during the first lockdown in Spring 2020 (Belot et al., 2020) (see Table 5.1). Third, network clustering ($\alpha \in \{0.3, 0.4, 0.5\}$) as collected in a cross-sectional study of social contacts in England, Scotland, and Wales (Danon et al., 2013). Due to lack of empirical data, occupational group homophily was varied to realize scenarios without ($\omega = 0.0$), with medium ($\omega = 0.4$) and with strong ($\omega = 0.8$) assortative mixing.

A genetic algorithm was used to find initial settings for the average number of ties per occupational group (c_2) and the degree of network clustering (α) that match the target values for mean degree and clustering best. Note that occupational group homophily was not considered, but varied explicitly due to the lack of empirical data. The algorithm consisted of six generations for each of the nine systematically varied parameters (3 for clustering, 3 for homophily). Each generation consisted of a number of model realizations with varying parameter settings for marginal costs per occupational group (c_2) and optimal proportion of closed triads (α). The initial generation consisted of 4 simulations with parameter settings according to the empirical data. At the beginning of each of the following five generations, the four best fitting model realizations were selected as parents (the lowest percentage error to the target values). For each of the six possible pairs of parents, two offspring model realizations were created. That is, parameter settings for marginal costs per occupational group and clustering were randomly selected from one of the parents (gene selection). If the parameter settings created outcomes that deviated more than 2% from the target values, the parameter settings were varied randomly within a range between 0.0 and the percentage error for the according value (gene mutation). The new offspring was used to generate a network in a subsequent NIDM simulation. Each simulation lasted until the fitness (cumulative percentage errors between target and realized values for average degrees per occupational group and network clustering) did not improve for 5 consecutive rounds. A single combination of the explicitly varied parameters (prior to lockdown contacts, 3 settings for clustering, 3 settings for homophily) thus created $4 + 5 \cdot 6 \cdot 2 = 64$ networks, while the entire procedure resulted in $3 \cdot 3 \cdot 64 = 576$ networks. Finally, we selected the 10 best fitting networks for each parameter combination of clustering and homophily, resulting in 90 networks with normal (prior to lockdown) degree distributions.

Lockdown generation was realized by pruning the 90 previously generated networks. That is, network ties were severed based on the reduction of contacts during lockdown per occupational group. For every network tie, the probability of severing the tie depended

on two aspects. First, both nodes have not reached the target lockdown degree of their corresponding occupational group. Second, the probability to sever a tie depended on the reduction of contacts during lockdown (Belot et al., 2020). Consider two nodes are connected by a tie. Node 1 belongs to the group *Legal Occupations*, which showed an average reduction of contacts from 8.28 to 0.92 (88.89%). Node 2 belongs to the group *Healthcare Support Occupations*, which showed an average reduction of contacts from 5.44 to 3.70 (31.99%). Node 1 has 6 ties, while Node 2 has 4 ties. Thus, both ties have more ties than the average node in their occupational group during lockdown, and the tie between the nodes is severed with a probability of $\frac{88.89+31.99}{2} = 60.44\%$. Lockdown generation stopped when, for every pair of nodes, at least one node reached the lockdown degree.

In summary, a total number of 180 networks were used as input for the simulation of epidemics. This number consists of the 10 best fitting networks for each of the 9 systematic parameter variations (3 for clustering, 3 for homophily) and average degrees reported prior to the first COVID-19 lockdown in Spring 2020; and the same number of lockdown networks generated from these networks.

5.5.3 Technical setup and runtimes

Simulations were run on a MacBook Pro 13", early 2013, with an Intel Core i5 Dual-Core processor running at 2.6 GHz, using 8 GB DDR3 RAM at 160 MHz. At the time of the simulations, the computer ran on macOS Catalina version 10.15.7 (19H524). The simulation was programmed in Java 8 with GraphStream v1.3 (Pigné et al., 2008) used for graph handling. The Java code was executed using Eclipse v4.18 and Java compiler v1.8.0_91. For analyses, we used R v4.0.4 (R Core Team, 2019) with *ggplot2* (Wickham, 2016) for data visualization.

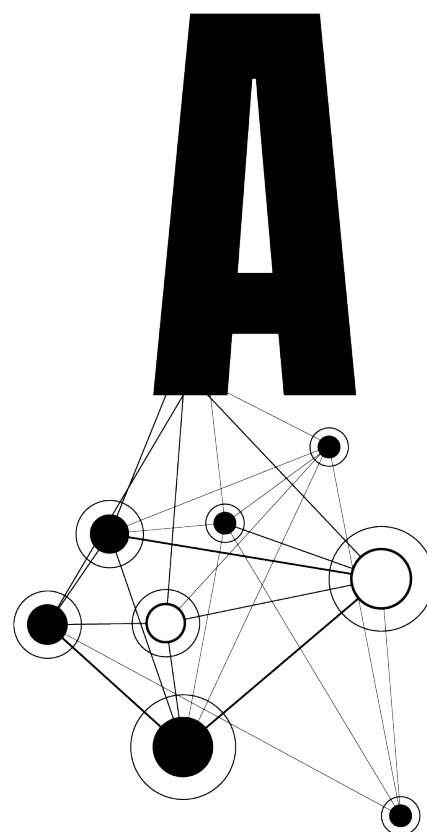
All simulations ran in 8 parallel threads. Runtimes are reported as combined totals. The network generation and fitting process for the normal (prior to lockdown) networks took 222 hours and 24 minutes. Network pruning to generate the lockdown networks took 239 hours and 41 minutes. Simulation of epidemics took 212 hours and 27 minutes.

5.5.4 Code availability

The Java source code to generate the data, and the R script to analyze the data are available under the GPLv3 license in the GitHub repository, <https://github.com/hnunner/NIDM-simulation> (version: v4.2.1., commit: 1707a0b, DOI: 10.5281/zenodo.5257528). Detailed instruction how to reproduce all simulations and analyses can be found in the release notes of version v4.2.1 on GitHub and Zenodo.

5.5.5 Data availability

The simulated network and epidemics data are available in the GitHub repository, <https://github.com/hnunner/NIDM-simulation> (version: v4.2.1., commit: 1707a0b, DOI: 10.5281/zenodo.5257528). The six country survey data of close-range contact including occupational codes (Belot et al., 2020) can be publicly accessed at <https://osf.io/aubkc/>.



Appendix A

Supplement to Chapter 2

Figures A.1 and A.2 show the results of Figure 2.2 divided between simulation runs with network changes and without network changes. A comparison of the two figures shows that epidemics in networks where ties are being cut (Figure A.1) have a significantly lower final size and mostly shorter duration across all network sizes. Furthermore, epidemics without network changes show hardly any variance in final size or duration (Figure A.2). Variance in epidemic measures is therefore not merely an effect of average degree in different network sizes, but sensitive to network changes of agents trying to distance themselves from the disease. An integrated discussion of these results can be found in Section 2.7.3 of Chapter 2.

Table A.1 shows descriptive statistics over the whole data for epidemics, network dynamics during and before epidemics, and network measures at the last time step before the first infection (pre-epidemic) and the first time step after the last infected agent has recovered (post-epidemic). It turns out that in a considerable portion of the simulation runs, agents do not see reasons to change their ties given their anticipated benefits and costs of ties even though some of their neighbors are infected by the diseases. By comparing these simulations with the ones in which agents do indeed change ties to distance themselves from others (see *Epidemics II.* and *III.* in Table A.1), we observe that in the runs with network changes attack rate is on average 34.28% lower and duration is on average 2.83 time steps shorter. Furthermore, considering that networks were pairwise stable before the initial infection, the average number of 3.41 network changes per agent (sum of dissolution and creation of ties) is caused and mainly driven by the presence of infectious agents. Additionally, it is shown that even more network changes were initiated by agents trying to connect to others (on average 18.00 tie requests per agent). These tie requests, however, were accepted only in 9% of the cases. This is a vast difference to the network initialization stage prior to the epidemics, where 40% tie requests were accepted from an average of 7.97 tie requests per agent. Despite the many changes in network ties, the overall network measures - with regard to degree at distance 1 and 2, and summary statistics of the density, clustering, and closeness - are the same before and after the epidemics.

Tables A.2 and A.3 show the *Interaction effects* models of Tables 2.2 and 2.3 divided by network size. The relatively large effect of degree of patient-0 (deg_{start}^0) in the model for networks of size 50 in Table A.2 (and the accompanying significance reduction of the other parameters) suggests that final size of epidemics in large networks is mostly driven by average degree. The reversal of many main effects between small and large networks for duration (Table A.3) shows that in larger and thus denser networks provide more potential transmission routes, ultimately allowing the disease to travel faster. Furthermore, most interaction effects vanish in the $N = 50$ model. Thus, network size drives the entire epidemic dynamics in large networks, while other parameters only have an effect in smaller networks. Regression models of epidemic duration split by network size (Table A.2) show that many main effects reverse between small and large networks. Consider, for example, benefits at distance 2 (β). In the models for networks with $N < 50$ the effect is positive, while in models for networks with $N = 50$ the effect is negative. That is, in large networks increased benefits create even more ties per agent, ultimately providing large numbers of potential transmission routes, and thus allowing the disease to travel faster.

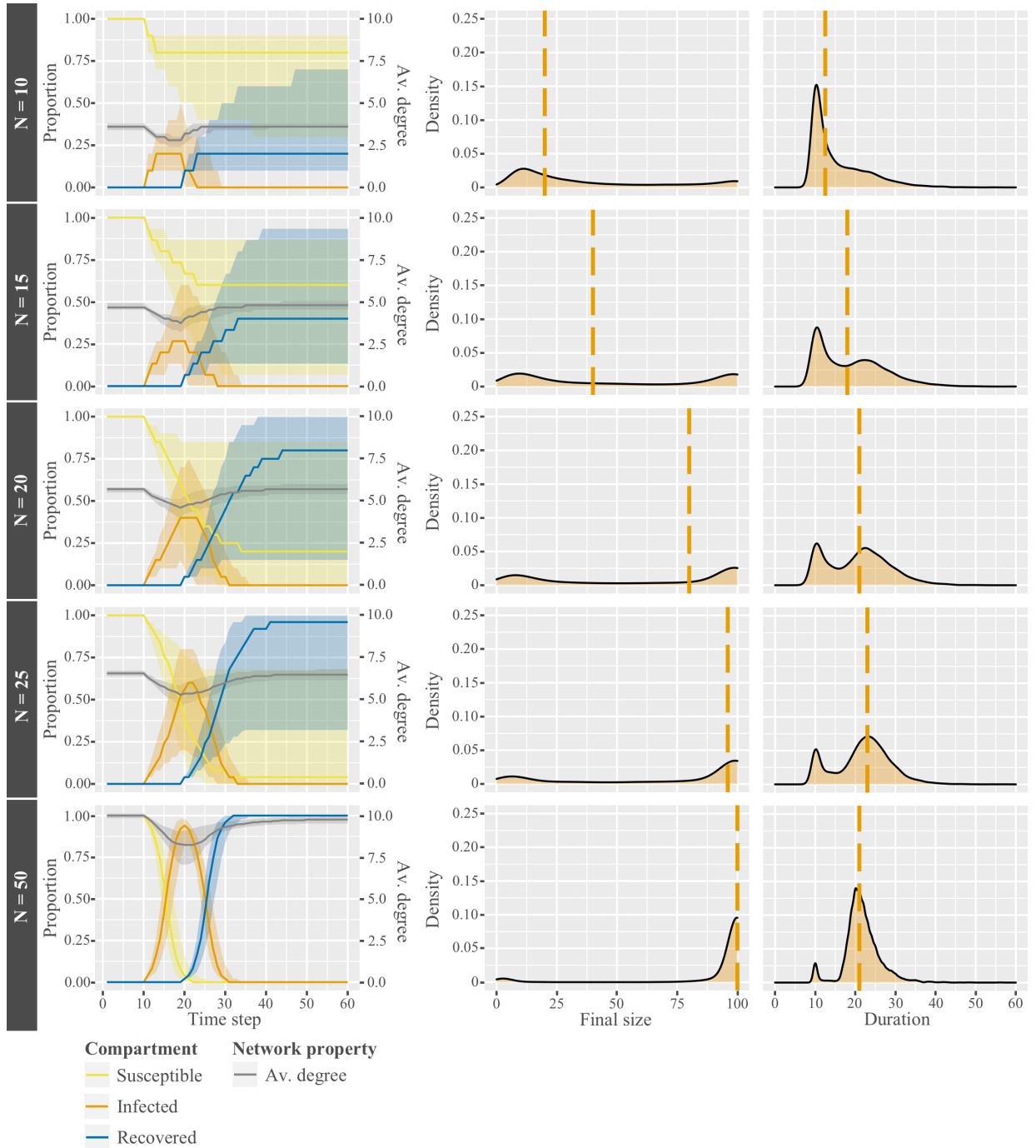


Figure A.1: **Disease dynamics by network size in networks with network changes.** Interplay of network and epidemic dynamics by comparison of median proportions of disease states and average degree over time (column 1), distribution of final size of the epidemic (column 2), and distribution of duration (column 3), divided by network size. The dashed orange line indicates the median of the respective distributions.

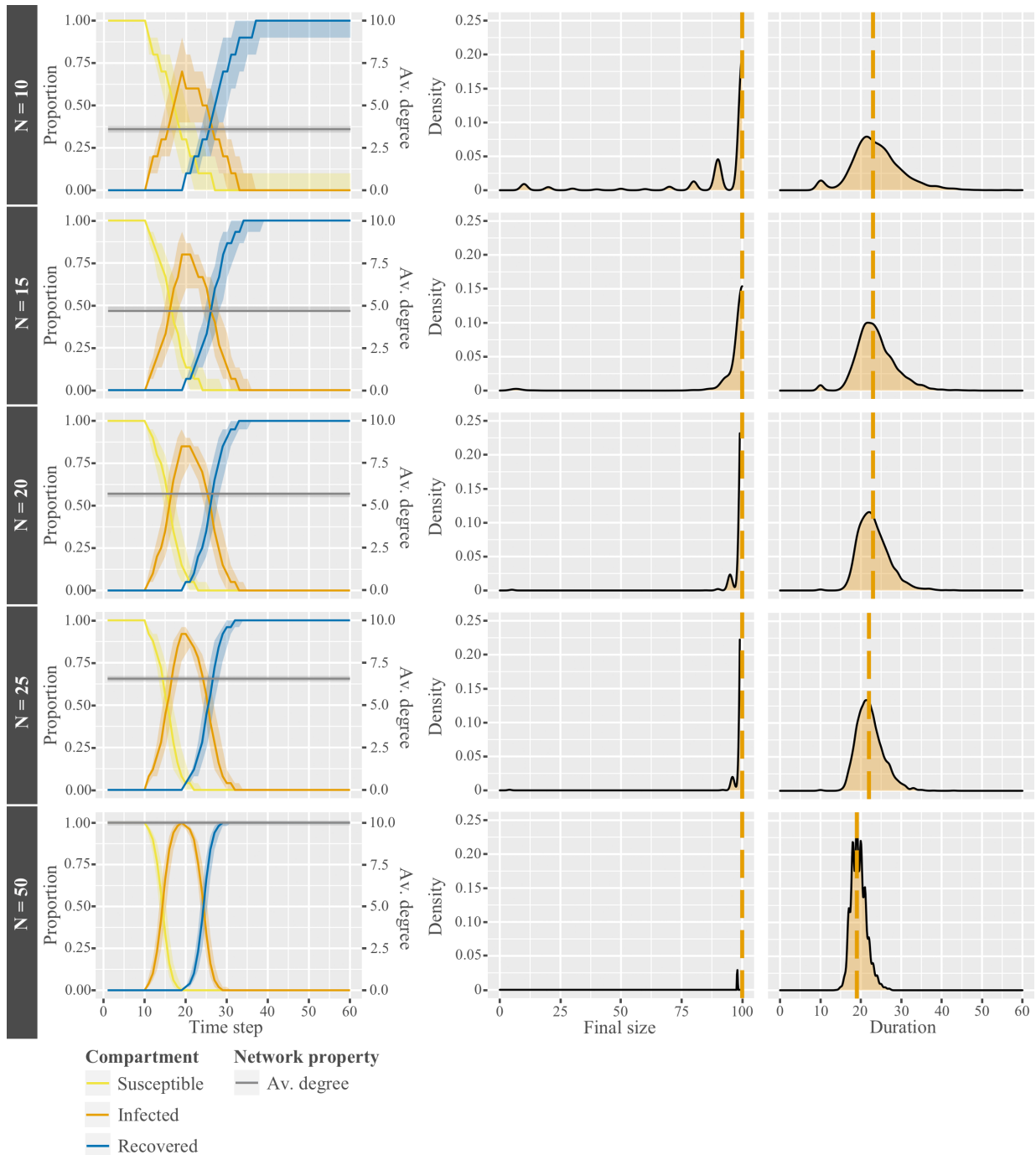


Figure A.2: **Disease dynamics by network size in networks without network changes.** Interplay of network and epidemic dynamics by comparison of median proportions of disease states and average degree over time (column 1), distribution of final size of the epidemic (column 2), and distribution of duration (column 3), divided by network size. The dashed orange line indicates the median of the respective distributions.

Table A.1: Descriptive statistics of network dynamics, epidemics, and network measures.

	Mean	SD	Min	Max	Skew
<i>Epidemics</i>					
<i>I. All simulation runs (36,000)</i>					
Final size	74.75	36.54	2.00	100.00	-1.00
Duration	20.66	6.68	10	60	0.30
<i>II. Simulation runs with network changes (23,930)</i>					
Final size	63.26	39.34	2.00	100.00	-0.40
Duration	19.71	7.34	10	59	0.43
<i>III. Simulation runs without network changes (12,070)</i>					
Final size	97.54	11.57	4.00	100.00	-6.49
Duration	22.54	4.58	10	60	1.10
<i>Network dynamics</i>					
<i>I. During epidemic</i>					
Av. no. ties broken / agent	1.80	2.61	0.00	22.86	2.46
Av. no. ties created / agent	1.81	2.61	0.00	22.80	2.46
Av. no. tie requests / agent	18.95	25.27	0.00	231.82	2.37
% tie requests accepted	0.09	0.03	0.01	1.00	6.21
<i>II. Pre-epidemic</i>					
Av. no. tie requests / agent	7.15	6.48	0.00	34.94	1.59
% tie requests accepted	0.41	0.18	0.07	1.00	0.59
<i>Network measures</i>					
<i>I. Pre-epidemic</i>					
Degree	6.12	2.18	3.00	10.80	0.77
Distance 2 degree	16.88	11.78	4.00	39.68	1.08
Density	0.30	0.07	0.19	0.56	0.17
Clustering	0.01	0.01	0.00	0.27	4.65
Closeness	0.96	0.02	0.93	0.98	-0.31
<i>II. Post-epidemic</i>					
Degree	6.09	2.11	3.00	10.92	0.75
Distance 2 degree	16.91	11.86	4.00	40.12	1.08
Density	0.30	0.07	0.18	0.56	0.24
Clustering	0.01	0.01	0.00	0.27	4.62
Closeness	0.96	0.02	0.93	0.98	-0.31

Table A.2: Comparison of two-level random-intercept logistic regression models for final size of the epidemic by network size.

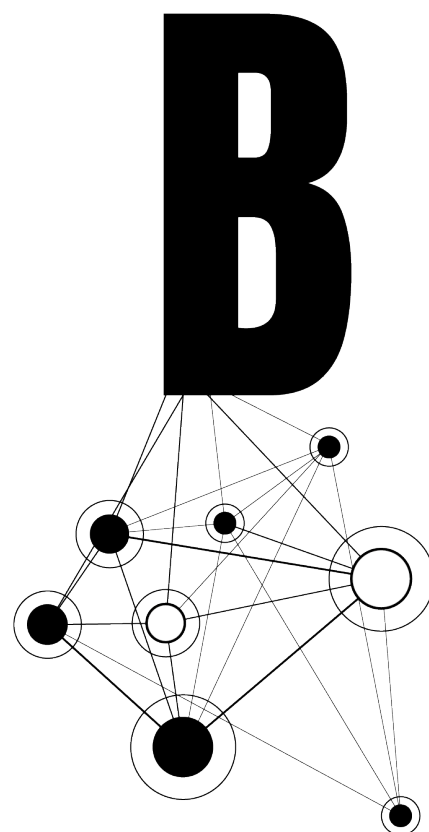
	Combined	N=10	N=15	N=20	N=25	N=50
Fixed/observed effects						
Intercept	2.59*** (0.07)	0.05 (0.04)	1.46*** (0.11)	2.40*** (0.16)	3.17*** (0.19)	4.75*** (0.27)
Main effects						
<i>I. CIDM parameters</i>						
Benefit distance 2 (β)	0.23*** (0.02)	0.23*** (0.01)	0.25*** (0.03)	0.27*** (0.04)	0.26*** (0.04)	0.20*** (0.04)
Cost increase for infected ties (μ)	-5.30*** (0.21)	-4.89*** (0.17)	-5.85*** (0.42)	-6.71*** (0.62)	-6.53*** (0.76)	-4.44*** (0.94)
Disease severity ($\frac{\sigma}{50}$)	-1.75*** (0.11)	-1.86*** (0.10)	-1.91*** (0.23)	-2.50*** (0.30)	-2.11*** (0.31)	-1.59*** (0.35)
Risk perception (r)	-3.16*** (0.13)	-2.51*** (0.11)	-3.16*** (0.26)	-3.76*** (0.36)	-4.28*** (0.43)	-3.57*** (0.51)
Network size ($\frac{N}{50}$)	6.22*** (0.40)					
<i>II. Network properties</i>						
Density (den_{start})	-7.35*** (0.98)	-4.16** (1.27)	-11.91*** (2.36)	-10.00** (3.44)	-13.64** (5.24)	-5.72 (20.02)
Degree of patient-0 (deg_{start}^0)	6.46*** (0.54)	2.73*** (0.45)	5.79*** (0.67)	2.88*** (0.82)	6.40*** (1.07)	17.29*** (3.33)
Interaction effects						
$\beta \times \mu$	0.52*** (0.06)	0.46*** (0.05)	0.48*** (0.13)	0.54*** (0.16)	0.50** (0.16)	0.30 (0.16)
$\mu \times \frac{\sigma}{50}$	3.16*** (0.44)	1.74*** (0.40)	3.80*** (0.92)	4.70*** (1.18)	3.99** (1.21)	2.39* (1.17)
$\mu \times r$	6.41*** (0.49)	4.33*** (0.42)	7.60*** (1.04)	8.34*** (1.44)	9.60*** (1.71)	7.24*** (1.91)
$\frac{\sigma}{50} \times r$	-2.04*** (0.27)	-3.14*** (0.27)	-1.93*** (0.57)	-1.65* (0.69)	-1.88** (0.73)	-0.74 (0.71)
$r \times den_{start}$	9.39*** (1.52)	-5.29 (3.03)	13.86* (5.38)	21.26* (8.43)	23.66 (13.10)	32.70 (37.99)
$\frac{N}{50} \times deg_{start}^0$	13.79*** (2.30)					
Random/unobserved effects						
s^2	0.47	0.02	0.44	0.57	0.52	0.25
Log Likelihood (ℓ)	-11,581.31	-3,092.16	-2,717.52	-2,507.48	-2,277.77	-926.14
Intraclass correlation (ρ)	0.125	0.007	0.118	0.149	0.136	0.072
Observations	36,000	7,200	7,200	7,200	7,200	7,200
Groups: parameter combinations	360	72	72	72	72	72

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

Table A.3: Comparison of two-level random-intercept linear regression models for duration of epidemics by network size.

	Combined	N=10	N=15	N=20	N=25	N=50
Fixed/observed effects						
Intercept	21.23*** (0.19)	18.41*** (0.29)	20.43*** (0.23)	21.48*** (0.18)	22.00*** (0.16)	20.97*** (0.08)
Main effects						
<i>I. CIDM parameters</i>						
Benefit distance 2 (β)	0.36*** (0.04)	0.59*** (0.10)	0.71*** (0.08)	0.45*** (0.06)	0.19*** (0.05)	-0.16*** (0.03)
Cost increase for infected ties (μ)	-3.42*** (0.44)	-10.68*** (1.15)	-6.26*** (0.94)	-2.83*** (0.73)	0.11 (0.65)	2.58*** (0.31)
Disease severity ($\frac{\sigma}{50}$)	-1.81*** (0.26)	-3.64*** (0.68)	-3.08*** (0.56)	-2.42*** (0.44)	-1.38*** (0.39)	1.50*** (0.19)
Risk perception (r)	-2.81*** (0.27)	-5.06*** (0.70)	-4.35*** (0.57)	-3.62*** (0.45)	-2.46*** (0.40)	1.45*** (0.19)
Network size ($\frac{N}{50}$)	-0.11 (1.15)					
<i>II. Network properties</i>						
Density (den_{start})	-19.93*** (3.20)	-22.55*** (2.81)	-33.53*** (4.87)	-25.41*** (6.60)	-18.31* (8.19)	-41.03*** (11.45)
Degree of patient-0 (deg_{start}^0)	3.31*** (0.62)	2.24* (0.98)	7.35*** (1.38)	1.71 (1.66)	2.57 (1.88)	-1.67 (2.11)
Interaction effects						
$\beta \times \mu$	1.05*** (0.15)	1.71*** (0.38)	2.01*** (0.31)	1.45*** (0.24)	0.56** (0.22)	-0.47*** (0.10)
$\beta \times N$	-1.06*** (0.13)					
$\mu \times den_{start}$	-38.17*** (5.00)	2.64 (10.54)	-1.11 (18.65)	30.21 (25.49)	31.33 (31.70)	-27.52 (45.03)
$\frac{\sigma}{50} \times r$	-4.16*** (0.64)	-6.74*** (1.67)	-6.84*** (1.37)	-6.38*** (1.07)	-3.98*** (0.95)	3.12*** (0.45)
$\frac{\sigma}{50} \times N$	6.48*** (0.93)					
$r \times N$	8.20*** (0.96)					
$\frac{N}{50} \times den_{start}$	32.97*** (9.35)					
Random/unobserved effects						
s^2	3.96	5.59	3.57	2.05	1.60	0.30
Log Likelihood (ℓ)	-11,3274.74	-23,109.20	-23,355.84	-23,203.31	-22,761.55	-19,558.51
Intraclass correlation (ρ)	0.114	0.138	0.087	0.054	0.048	0.022
Observations	36,000	7,200	7,200	7,200	7,200	7,200
Groups: parameter combinations	360	72	72	72	72	72

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.



Appendix B

Supplements to Chapter 3

B.1 Additional model definitions

B.1.1 Utility

Utility for an agent i is defined in the SWIDM as the trade-off between the social benefits (B_i), social maintenance costs (C_i), and (potential) harm (D_i) of infections:

$$U_i(\mathbf{G}, \mathbf{d}, \mathbf{R}) = B_i(\mathbf{G}) - C_i(\mathbf{G}) - D_i(\mathbf{G}, \mathbf{d}, \mathbf{R}). \quad (\text{B.1})$$

We consider networks to be undirected, unweighted, and non-reflexive graphs:

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \dots & g_{1N} \\ g_{21} & g_{22} & \dots & g_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ g_{N1} & g_{N2} & \dots & g_{NN} \end{bmatrix}, \quad (\text{B.2})$$

with $g_{ij} \in \{0, 1\}$, $g_{ii} = 0$, and $g_{ij} = g_{ji} = 1$ if a tie exists. Disease states (susceptible, infected, or recovered) are defined for each agent in the network:

$$\mathbf{d} \in \{S, I, R\}^N. \quad (\text{B.3})$$

Risk perception for each agent is composed of perceived risk to get infected (r_π) and perceived severity of the disease (r_σ):

$$\mathbf{R} : \mathbf{r}_\pi \in \mathbb{R}, \mathbf{r}_\sigma \in \mathbb{R}. \quad (\text{B.4})$$

Social benefits are defined as the weighted sum of the benefits for ties and the benefits for the proportion of closed triads:

$$B_i(\mathbf{G}) = b_1 \cdot t_i + b_2 \cdot \left(1 - 2 \cdot \frac{|x_i - \alpha|}{\max(\alpha, 1 - \alpha)} \right), \quad (\text{B.5})$$

with x_i denoting the proportion of closed triads i belongs to in i 's ego-network, α the preferred proportion of closed triads of agent i , and t_i the number of ties agent i possesses:

$$t_i = \sum_j g_{ij}. \quad (\text{B.6})$$

Social maintenance costs are assumed to be quadratic in the number of ties t_i to model increasing marginal costs of additional ties:

$$C_i(\mathbf{G}) = c_1 \cdot t_i + c_2 \cdot t_i^2. \quad (\text{B.7})$$

This allows to control for the number of ties agents want to establish in the network. (Potential) harm of infections is a combination of perceived probability to get infected and perceived severity of the disease:

$$D_i(\mathbf{G}, \mathbf{d}, \mathbf{R}) = p_i(\mathbf{G}, \mathbf{d}, \mathbf{R}) \cdot s_i(\mathbf{d}, \mathbf{R}). \quad (\text{B.8})$$

Perceived probability depends on the disease state of an agent:

$$p_i(\mathbf{G}, \mathbf{d}, \mathbf{R}) = \begin{cases} \pi_i(\mathbf{G}, \mathbf{d})^{2-r_\pi}, & \text{if } d_i = S \\ 1, & \text{if } d_i = I \\ 0, & \text{if } d_i = R, \end{cases} \quad (\text{B.9})$$

while actual probability to get infected

$$\pi_i(\mathbf{G}, \mathbf{d}) = 1 - (1 - \gamma)^{t_{iI}}, \quad (\text{B.10})$$

depends on the probability to get infected per single contact (γ) and on the number of infected ties:

$$t_{iI} = \sum_{j, d_j=I} g_{ij}. \quad (\text{B.11})$$

Again, perceived severity of a disease depends on the disease state of an agent:

$$s_i(\mathbf{d}, \mathbf{R}) = \begin{cases} \sigma^{r_\sigma}, & \text{if } d_i = S \\ \sigma, & \text{if } d_i = I \\ 0, & \text{if } d_i = R. \end{cases} \quad (\text{B.12})$$

Consequently, agents with risk perception values below 1 underestimate, while agents with risk perception values above 1 overestimate the probability of infections and/or disease severity.

B.1.2 Process overview and scheduling

Time is modeled as discrete time steps. Within each time step two distinct, interdependent processes are simulated: (i) disease dynamics and (ii) social network dynamics. Disease dynamics simulate the transmission of infectious diseases:

- Repeat until all agents have been processed:
 - Randomly select an unprocessed agent i :
 - If i is infected, compute whether agent recovers: passed time steps since infection $\geq \tau$.
 - If i is susceptible, compute whether i gets infected from infected direct ties (see Equation 3.1).

Social network dynamics simulate the formation and dissolution of ties:

- Create an empty set of unprocessed agents A^1 and repeat until all agents have been processed:
 - Randomly select an unprocessed agent i , either:
 - from A if $A \neq \{\}$ and not all agents from A have been processed yet, or
 - from the entire population N if $A = \{\}$.
 - Repeat until agent i has evaluated $\phi \cdot (N - 1)$ ties to other agents j :
 - Create a new set J of agents to be evaluated through one of the three following options:
 1. With probability ψ : J becomes all agents at distance 1 that have not yet been evaluated by i in the current time step.
 2. With probability ξ : J becomes all agents at distance 2 that have not yet been evaluated by i in the current time step.
 3. With probability $1 - \psi - \xi$: J becomes all agents from the entire population that have not yet been evaluated by i in the current time step.²
 - Select a single agent j from J either:
 - The agent most similar regarding risk perception (with probability ω):

$$\min_{\forall j \in J} |(r_{\sigma,i} + r_{\pi,i}) - (r_{\sigma,j} + r_{\pi,j})|$$
, or
 - a randomly selected agent irrespective of risk perception similarity (with probability $1 - \omega$).
 - If j is directly tied to i :
 - Dissolve tie ij , if i 's utility excluding ij exceeds the utility including ij .
 - else:
 - Form tie ij , if both i 's and j 's utilities including ij exceed the utilities excluding ij .
 - Add j to A if j is directly tied to i .

B.1.3 Entities, state variables and scales

Table B.1 shows an extended version of Table 3.1, including parameters introduced in Section B.2.

¹The use of A (in combination with parameters ψ and ξ that are introduced later) and thus the consecutive processing of an agent's neighbors allows producing high levels of clustering and homophilous mixing in networks too large for a purely randomized approach.

²Allowing each agent to alter social ties to a number of other agents that are selected from neighbors at distance 1, distance 2, and randomly selected from the entire network realizes the concept of focused interaction in an agent's social vicinity (Feld, 1981).

Table B.1: State variables, ranges, and initial settings.

State variable	Scale	Setting
Agent		
<i>I. Parameters</i>		
Benefit per social tie*	$b_1 \in \mathbb{R}_0^+$	$b_1 = 1.0$
Benefit for triadic closure	$b_2 \in \mathbb{R}_0^+$	$b_2 = 0.5$
Preferred proportion of closed triads [†]	$0 \leq \alpha \leq 1$	$\alpha \sim U[0, 1]$
Simple cost per tie*	$c_1 \in \mathbb{R}_0^+$	$c_1 = 0.2$
Marginal cost per tie*	$c_2 \in \mathbb{R}_0^+$	$c_2 = 0.05$
Risk perception (disease severity)	$0 \leq r_\sigma \leq 2$	$r \sim U[r_{min}, r_{max}]$
Risk perception (susceptibility)	$0 \leq r_\pi \leq 2$	
Proportion of agents to evaluate per time step	$0 \leq \phi \leq 1$	$\phi = 0.2$
Proportion of ϕ as distance 1 ties	$0 \leq \psi \leq 1$	$\psi = 0.4$
Proportion of ϕ as distance 2 ties	$0 \leq \xi \leq 1$	$\xi = 0.2$
<i>II. Outcomes</i>		
Number of social ties	$t_i \in \mathbb{N}_0$	
Actual proportion of closed triads	$0 \leq x \leq 1$	
Disease state	$d \in \{S, I, R\}$	
Index case (first infected agent)	$\mathcal{I}_i \in \{0, 1\}$	
Network		
<i>I. Parameters</i>		
Number of agents	$N \in \mathbb{N}_0$	$N = 80$
Dynamic network during epidemic	$\delta \in \{0, 1\}$	$\delta = \{0, 1\}$
Minimum risk perception	$0 \leq r_{min} < 2$	$r_{min} \sim U[0, 1]$
Maximum risk perception	$0 < r_{max} \leq 2$	$r_{max} \sim U[1, 2]$
Likelihood of ties similar in risk perception [‡]	$0 \leq \omega \leq 1$	$\omega \sim U[0, 1]$
<i>II. Outcomes</i>		
Number of susceptible agents	$ S \in \mathbb{N}_0$	
Number of infected agents	$ I \in \mathbb{N}_0$	
Number of recovered agents	$ R \in \mathbb{N}_0$	
Number of broken ties	$t_G^- \in \mathbb{N}_0$	
Number of created ties	$t_G^+ \in \mathbb{N}_0$	
Clustering	$0 \leq \mathcal{C}_G \leq 1$	
Average path length	$\mathcal{L}_G \in \mathbb{R}_0^+$	
Homophily	$-1 \leq \mathcal{H}_G \leq 1$	
Average degree	$\mathcal{D}_G \in \mathbb{R}_0^+$	
Infectious disease, parameters		
Disease severity	$\sigma > 1$	$\sigma \sim U(1, 100]$
Infectivity [§]	$0 \leq \gamma \leq 1$	$\gamma \sim U[0.01, 0.20]$
Recovery time in time steps	$\tau \in \mathbb{Z}^+$	$\tau = 5$

Notes: *: The combination of $b_1 = 1.0$, $c_1 = 0.2$, and $c_2 = 0.05$ sets the preferred number of ties to 8. [†]: The effect of α on network clustering and average path length is depicted in Figure B.1. [‡]: The effect of ω on homophily is depicted in Figure B.1. [§]: Infectivity is operationalized as transmission probability per contact and time step.

B.1.4 Emergence

Network structures, social mixing patterns, and course of epidemics are the result of the cumulative networking behavior of agents and disease transmission between connected agents. Heterogeneity in network structures between simulation runs (i.e., clustering, path length) is the result of a stochastic process with (homogeneous) preferences for proportion of closed over open triads (α). Heterogeneity in social mixing patterns (i.e., random, homophilous) is the result of (homogeneous) likelihoods of ties similar in risk perception (ω , r). Consequently, the combination of triadic closure and homophily facilitate the creation of collectives (i.e., clusters of agents similar in risk perception). In addition to network structure, the course of epidemics is affected by the infectivity (γ), disease severity (σ), and risk perceptions of the agents (r).

B.1.5 Stochasticity

Stochasticity in the model is a consequence of two things. First, parameters sampled at the start of each simulation run to control either network properties (preferred proportion of closed triads (α) and preferred proportion of ties with similar risk perception (ω)), individual behavior (risk perception per agent (r)), or disease properties (severity (σ) and probability of transmission per contact (γ)) are varied randomly. Second, the procedure contains random processes, consisting of random selection of the initially infected agent (*index case*) and random order of agents to compute disease transmission and social network dynamics.

B.2 Additional analyses

B.2.1 Network and disease properties

Part 1 of Table B.2 shows detailed descriptive statistics for independent variables of the simulated epidemics. That is, network properties resulting from the randomized network formation process and disease properties resulting from parameter randomization. We see that all network properties of interest (homophily, network clustering, average path length) show considerable variances. As intended, average degree remains fairly stable at just below 8 ties per node. Although we do find correlations between the remaining variance in average degree and network properties of interest (average degree - homophily: $r_\tau = 0.09$, $p < 0.01$; average degree - clustering: $r_\tau = -0.64$, $p < 0.01$; average degree - average path length: $r_\tau = -0.67$, $p < 0.01$), we made sure that average degree does not play a significant role for the course of epidemics (see Section B.2.2). Hence, we excluded average degree from all further analyses. Finally, we confirm large variances in the varied disease properties (severity, infectivity) and average risk perception. Simulated epidemics are therefore based on a wide range of different small-world networks, infectious diseases, and health behaviors.

Table B.2: Descriptive statistics.

Variable	Mean	SD	Min	Max	Skew
<i>Independent variables</i>					
Homophily* (\mathcal{H}_G)	0.35	0.27	-0.19	0.98	0.27
Network clustering* (\mathcal{C}_G)	0.32	0.26	0	0.84	0.59
Average path length* (\mathcal{L}_G)	2.68	0.48	2.28	6.17	1.31
Average degree* (\mathcal{D}_G)	7.76	0.34	6.47	8.05	-1.13
Disease severity (σ)	49.86	28.89	1	100	0.01
Infectivity (γ)	0.11	0.05	0.01	0.2	0
Average risk perception ($r_{\sigma,\pi}$)	1	0.21	0.44	1.59	0
<i>Dependent variables</i>					
<i>I. Epidemics, adaptive networks</i>					
Final size [†]	46.33	44.61	1.25	100	0.2
Duration [‡]	15.93	10.15	5	74	0.91
Duration [‡] (final sizes $\leq 10\%$)	7.2	2.97	5	27	1.59
Duration [‡] (final sizes $\geq 90\%$)	22.28	7.52	9	74	1.23
Peak size [†]	21.85	23.05	1	80	0.75
Peak size [†] (final sizes $\leq 10\%$)	2.18	1.52	1	8	1.4
Peak size [†] (final sizes $\geq 90\%$)	47.97	14.4	11	80	-0.11
Number of broken ties (t_G^-)	119.67	116.33	0	901	0.73
Number of created ties (t_G^+)	116.69	117.38	0	896	0.71
<i>II. Epidemics, static networks</i>					
Final size [†]	72.96	39.62	1.25	100	-1.06
Duration [‡]	18.18	9.32	5	84	1.01
Duration [‡] (final sizes $\leq 10\%$)	7.65	3.62	5	30	1.46
Duration [‡] (final sizes $\geq 90\%$)	18.76	6.47	9	83	1.64
Peak size [†]	39.5	26.44	1	80	-0.2
Peak size [†] (final sizes $\leq 10\%$)	1.97	1.35	1	8	1.55
Peak size [†] (final sizes $\geq 90\%$)	56.47	15.02	12	80	-0.48
Number of broken ties (t_G^-)	0	0	0	0	0
Number of created ties (t_G^+)	0	0	0	0	0

Notes: *: at the first time step of epidemics. †: reported as percentages. *Final size* = percentage of the population that got infected at some point during a simulation run. *Peak size* = maximum percentage of agents that were infected at the same time during a simulation run. ‡: reported as number of time steps.

B.2.2 Effect of degree on final size

We find that both parameters to control homophily (ω), clustering and average path length (both α) affect average degree of the generated networks (Fig. B.1). While there is a strong effect of ω on homophily (**a**: approx. 0.0–0.8) there is a comparably small effect on average degree (**d**: approx. 7.7–7.9). Furthermore, compared to the strong effect of α on clustering (**f**: approx. 0.0–0.75) and average path length (**g**: approx. 2.4–3.75), the effect on average degree (**h**: approx. 7.2–8.0) is comparably small.

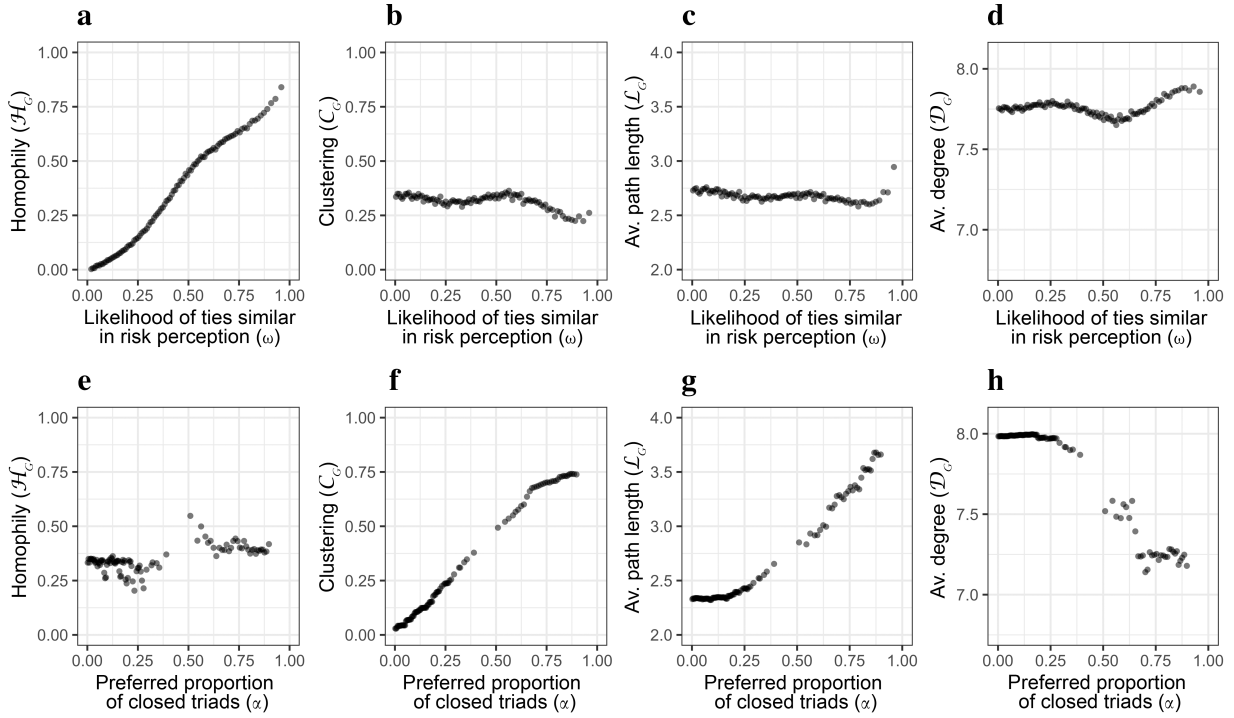


Figure B.1: **Effects of model parameters on network properties.** Row 1 shows the effect of likelihood of ties similar in risk perception (ω), row 2 shows the effect of preferred proportion of closed triads (α) on homophily (column 1), clustering (column 2), average path length (column 3), and average degree (column 4). Due to the large number of observations, outcome variables are sorted and binned into 100 points on the x-axis. Each point, therefore, shows the mean of $\frac{80,305}{100} \approx 803$ observations.

To ensure that the determining factors for the course of epidemics are the varied network properties of interest (homophily, clustering, average path length) rather than average degree, we compared their predictive power on final size (Tables B.3 and B.4). Although we find a small effect of average degree in adaptive and static networks (*Average degree* models), the improvement of explained variance and goodness of fit becomes negligible when controlled for the varied network parameters of interest (no increase for adjusted R^2 and minor increase for AIC between the *Network properties* and *Combined* models). Furthermore, average degree is correlated with homophily ($r_\tau = 0.09$, $p < 0.01$), clustering ($r_\tau = -0.64$, $p < 0.01$), and average path length ($r_\tau = -0.67$, $p < 0.01$). The effect of average degree on final size is, therefore, in principle expressed by homophily, clustering, and average path length, and therefore discarded from further analyses.

B.2.3 Additional effects on epidemics

Figure B.2 shows an extended version of Figure 3.2 in Section 3.5.2, including additional effects (average degree, average risk perception, disease severity, infectivity). Effects of all parameters on epidemics in static networks can be found in B.3.

Table B.3: Logistic regression models for the effect of average degree on final size in adaptive networks.

	Network changes	Average degree	Network properties	Combined
Intercept	0.08*** (0.01)	0.07*** (0.01)	-0.12*** (0.01)	-0.12*** (0.01)
Number of network changes ($t_G^{+/-}$)	2.94*** (0.02)	2.89*** (0.02)	2.79*** (0.02)	2.80*** (0.02)
Average degree (\mathcal{D}_G)		0.20*** (0.01)		-0.42*** (0.04)
Homophily (\mathcal{H}_G)			-0.65*** (0.01)	-0.68*** (0.01)
Clustering (\mathcal{C}_G)			-0.48*** (0.03)	-0.70*** (0.04)
Path length (\mathcal{L}_G)			0.28*** (0.03)	0.13*** (0.03)
Adjusted R ²	0.76	0.77	0.80	0.80
AIC	30,441	30,026	26,208	26,047
Number of observations	80,305	80,305	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

Table B.4: Logistic regression models for the effect of average degree on final size in static networks.

	Average degree	Network properties	Combined
Intercept	1.01*** (0.01)	1.01*** (0.01)	1.01*** (0.01)
Average degree (\mathcal{D}_G)	0.28*** (0.01)		0.09*** (0.02)
Homophily (\mathcal{H}_G)		-0.01 (0.01)	-0.01 (0.01)
Clustering (\mathcal{C}_G)		-0.08*** (0.02)	-0.03 (0.02)
Path length (\mathcal{L}_G)		-0.22*** (0.02)	-0.19*** (0.02)
Adjusted R ²	0.02	0.02	0.02
AIC	90,273	90,119	90,102
Number of observations	80,305	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

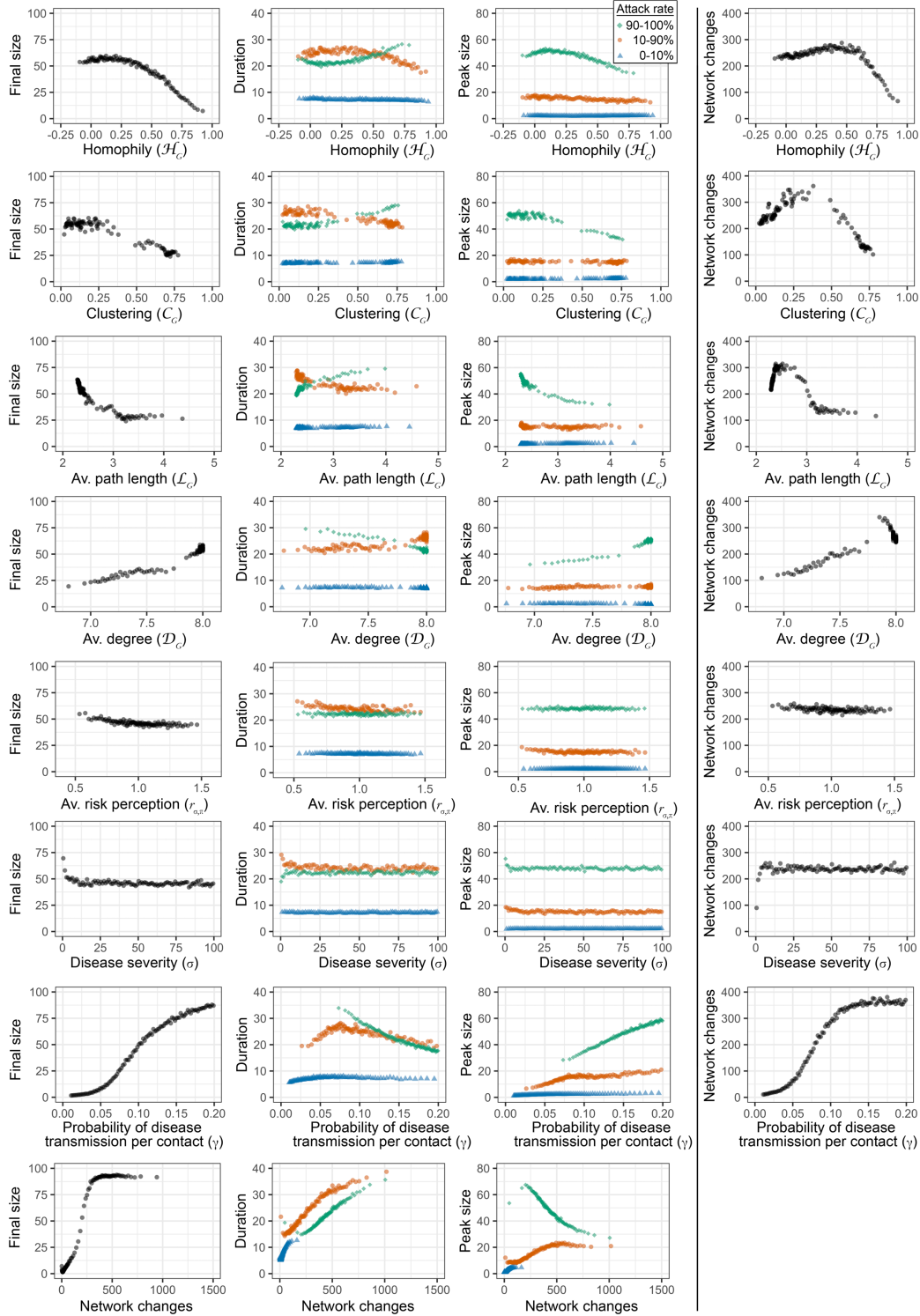


Figure B.2: **Effect of network properties on epidemics in adaptive networks.** Plots show bivariate effects of homophily (row 1), network clustering (row 2), and average path length (row 3), average degree (row 4), average risk perception (row 5), disease severity (row 6), infectivity (row 7), and network changes (row 8) on final size (column 1), duration (column 2), and peak size (column 3) of epidemics, as well as number of network changes (column 4) in adaptive networks. There are 100 points per group and plot, with each point showing the average of ≈ 803 ($\frac{80,305}{100}$) observations. Measures of variance are omitted, as they merely exceed the size of the dots. Data for duration and peak size have been grouped by final size: 0 – 10% (blue), 10 – 90% (orange), and 90 – 100% (green).

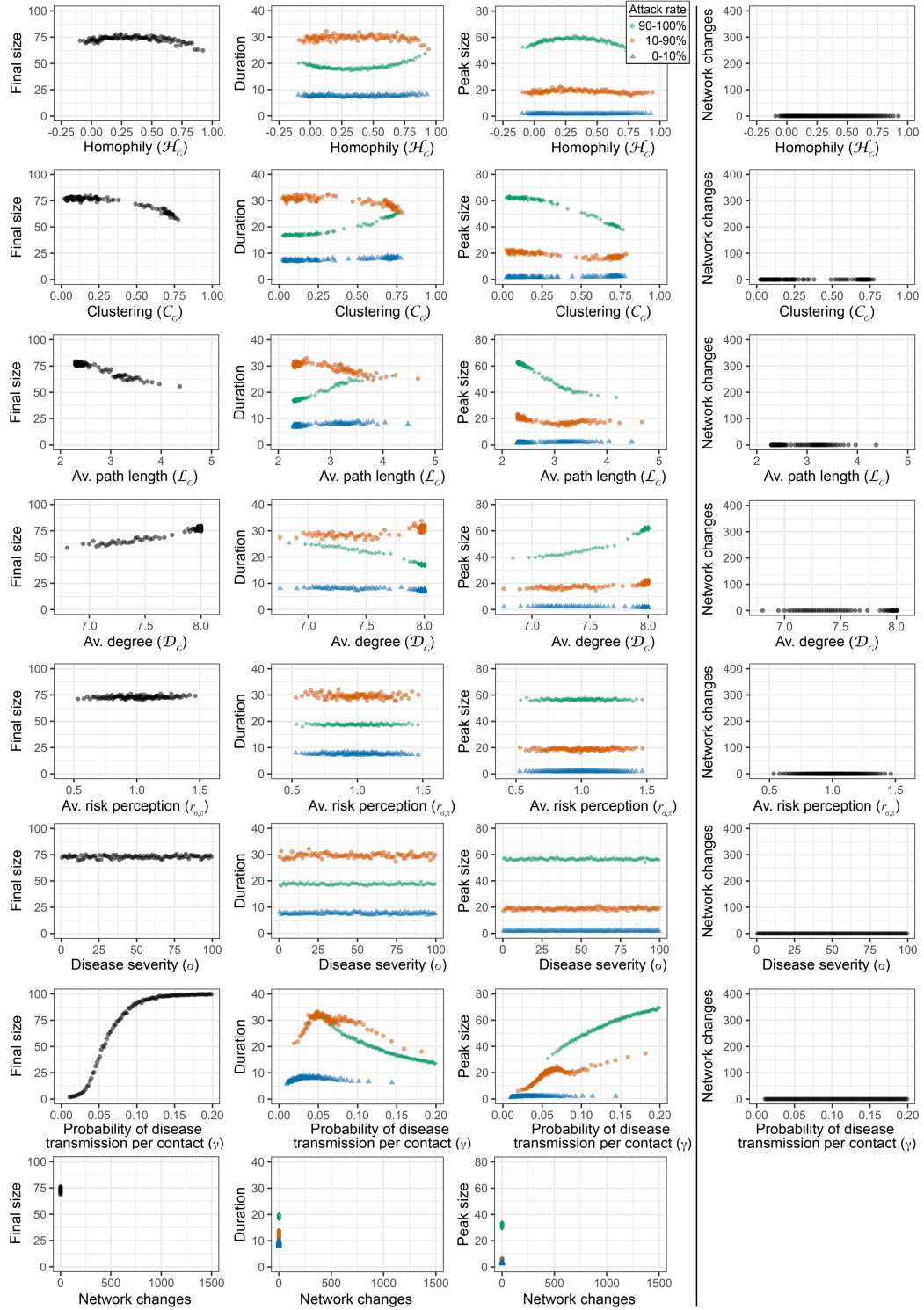


Figure B.3: **Effect of network properties on epidemics in static networks.** Plots show bivariate effects of homophily (row 1), network clustering (row 2), and average path length (row 3), average degree (row 4), average risk perception (row 5), disease severity (row 6), infectivity (row 7), and network changes (row 8) on final size (column 1), duration (column 2), and peak size (column 3) of epidemics, as well as number of network changes (column 4) in adaptive networks. There are 100 points per group and plot, with each point showing the average of ≈ 803 ($\frac{80,305}{100}$) observations. Measures of variance are omitted, as they merely exceed the size of the dots. Data for duration and peak size have been grouped by final size: 0 – 10% (blue), 10 – 90% (orange), and 90 – 100% (green).

B.2.4 Regression models

Tables B.5-B.11 contain all regression models used to generate Table 3.2.

Table B.5: Logistic regression models of final size with network dynamics.

	Network dynamics	Main effects	Interaction effects
Intercept	0.08*** (0.01)	-0.42*** (0.01)	-0.57*** (0.02)
Number of network changes ($t_G^{+/-}$)	2.94*** (0.02)	2.07*** (0.02)	2.03*** (0.02)
Clustering (\mathcal{C}_G)		-0.47*** (0.04)	-0.54*** (0.04)
Path length (\mathcal{L}_G)		-0.05 (0.04)	0.10** (0.04)
Homophily (\mathcal{H}_G)		-0.75*** (0.01)	-0.85*** (0.02)
Average risk perception ($r_{\sigma,\pi}$)		-0.16*** (0.01)	-0.19*** (0.01)
Disease severity (σ)		-0.20*** (0.01)	-0.24*** (0.01)
Infectivity (γ)		1.33*** (0.02)	1.48*** (0.02)
$t_G^{+/-}$ x \mathcal{C}_G			-0.46*** (0.02)
$t_G^{+/-}$ x \mathcal{H}_G			-0.44*** (0.02)
$t_G^{+/-}$ x γ			0.77*** (0.03)
\mathcal{C}_G x \mathcal{H}_G			0.22*** (0.02)
\mathcal{H}_G x γ			-0.41*** (0.02)
Adjusted R ²	0.76	0.88	0.92
Log Likelihood	-15,219	-9,480	-8,195
Number of observations	80,305	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

Table B.6: Linear regression models of epidemic duration with network dynamics.

	Network dynamics	Main effects	Interaction effects
Intercept	15.93*** (0.02)	15.93*** (0.02)	18.46*** (0.02)
Number of network changes ($t_G^{+/-}$)	8.33*** (0.02)	9.38*** (0.02)	10.83*** (0.02)
Clustering (\mathcal{C}_G)		-1.10*** (0.05)	-2.00*** (0.02)
Path length (\mathcal{L}_G)		2.48*** (0.05)	4.08*** (0.05)
Homophily (\mathcal{H}_G)		-0.95*** (0.02)	-0.53*** (0.02)
Average risk perception ($r_{\sigma,\pi}$)		-0.38*** (0.02)	-0.32*** (0.02)
Disease severity (σ)		-0.39*** (0.02)	-0.35*** (0.02)
Infectivity (γ)		-1.47*** (0.02)	-2.67*** (0.02)
$t_G^{+/-} \times \mathcal{C}_G$			-2.89*** (0.05)
$t_G^{+/-} \times \mathcal{L}_G$			4.72*** (0.05)
$t_G^{+/-} \times \gamma$			-3.64*** (0.02)
$\mathcal{C}_G \times \mathcal{H}_G$			-0.12*** (0.02)
Adjusted R ²	0.67	0.71	0.80
RMSE	5.79	5.42	4.54
Number of observations	80,305	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

Table B.7: Linear regression models of epidemic peak size with network dynamics.

	Network dynamics	Main effects	Interaction effects
Intercept	21.85*** (0.06)	21.85*** (0.04)	17.81*** (0.05)
Number of network changes ($t_G^{+/-}$)	15.04*** (0.06)	6.32*** (0.05)	5.93*** (0.04)
Clustering (\mathcal{C}_G)		-3.48*** (0.11)	-5.88*** (0.09)
Path length (\mathcal{L}_G)		-1.86*** (0.12)	-0.08 (0.09)
Homophily (\mathcal{H}_G)		-5.44*** (0.04)	-7.35*** (0.03)
Average risk perception ($r_{\sigma,\pi}$)		-0.72*** (0.04)	-0.79*** (0.03)
Disease severity (σ)		-0.70*** (0.04)	-0.76*** (0.03)
Infectivity (γ)		13.07*** (0.05)	14.07*** (0.04)
$t_G^{+/-} \times \mathcal{C}_G$			-3.01*** (0.05)
$t_G^{+/-} \times \mathcal{H}_G$			-1.90*** (0.04)
$t_G^{+/-} \times \gamma$			5.38*** (0.05)
$\mathcal{C}_G \times \mathcal{H}_G$			2.54*** (0.03)
$\mathcal{L}_G \times \gamma$			-2.91*** (0.04)
$\mathcal{H}_G \times \gamma$			-3.97*** (0.04)
Adjusted R ²	0.43	0.73	0.85
RMSE	17.46	11.98	8.82
Number of observations	80,305	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

Table B.8: Linear regression models of number of network changes.

	Main effects	Interaction effects
Intercept	236.37*** (0.66)	239.19*** (0.66)
Clustering (\mathcal{C}_G)	28.68*** (1.78)	37.32*** (1.78)
Path length (\mathcal{L}_G)	-75.68*** (1.78)	-86.14*** (1.79)
Homophily (\mathcal{H}_G)	-14.77*** (0.67)	-6.65*** (0.71)
Average risk perception ($r_{\sigma,\pi}$)	-4.50*** (0.66)	-4.43*** (0.66)
Disease severity (σ)	1.92** (0.66)	1.94** (0.66)
Infectivity (γ)	127.84*** (0.66)	128.04*** (0.66)
$\mathcal{L}_G \times \mathcal{H}_G$		-20.02*** (0.63)
$\mathcal{L}_G \times \gamma$		-12.15*** (0.66)
$\mathcal{H}_G \times \gamma$		16.83*** (0.66)
Adjusted R ²	0.35	0.37
RMSE	187.91	185.73
Number of observations	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

Table B.9: Logistic regression models of final size without network dynamics.

	Network dynamics	Main effects	Interaction effects
Intercept	0.99*** (0.01)	2.88*** (0.02)	3.08*** (0.03)
Clustering (\mathcal{C}_G)		-0.13 (0.03)	-0.37*** (0.03)
Path length (\mathcal{L}_G)		-0.65*** (0.03)	-0.69*** (0.03)
Homophily (\mathcal{H}_G)		-0.02 (0.01)	-0.04** (0.01)
Average risk perception ($r_{\sigma,\pi}$)		0.01 (0.01)	0.01 (0.01)
Disease severity (σ)		0.01 (0.01)	0.01 (0.01)
Infectivity (γ)		3.57*** (0.03)	3.71*** (0.03)
$\mathcal{L}_G \times \gamma$			-0.46*** (0.02)
Adjusted R ²	0.00	0.76	0.76
Log Likelihood	-45,973	-16,216	-16,080
Number of observations	80,305	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

Table B.10: Linear regression models of epidemic duration without network dynamics.

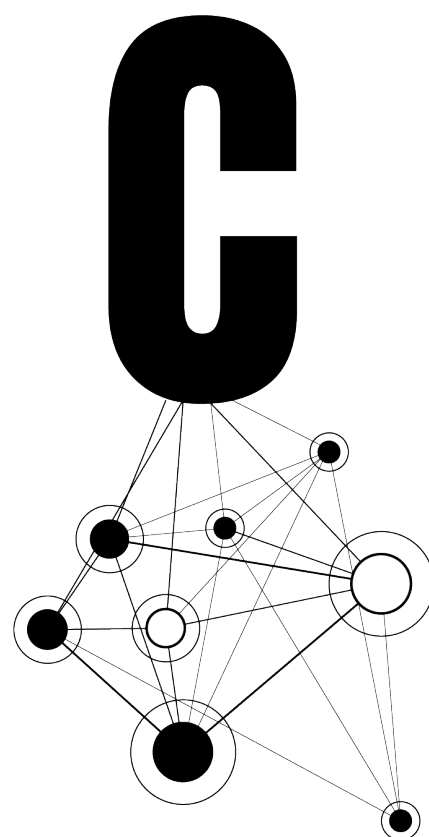
	Network dynamics	Main effects	Interaction effects
Intercept	18.18*** (0.03)	18.18*** (0.03)	18.18*** (0.03)
Clustering (\mathcal{C}_G)		1.08*** (0.09)	1.10*** (0.09)
Path length (\mathcal{L}_G)		0.61*** (0.09)	0.59*** (0.09)
Homophily (\mathcal{H}_G)		0.03 (0.03)	0.03 (0.03)
Average risk perception ($r_{\sigma,\pi}$)		0.03 (0.03)	0.02 (0.03)
Disease severity (σ)		0.01 (0.03)	0.01 (0.03)
Infectivity (γ)		-0.32*** (0.03)	-0.32*** (0.03)
$\mathcal{L}_G \times \gamma$			1.84*** (0.03)
Adjusted R ²	0.00	0.03	0.07
RMSE	9.32	9.16	8.98
Number of observations	80,305	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.

Table B.11: Linear regression models of epidemic peak size without network dynamics.

	Network dynamics	Main effects	Interaction effects
Intercept	39.50*** (0.09)	39.50*** (0.04)	39.51*** (0.03)
Clustering (\mathcal{C}_G)		-4.10*** (0.10)	-4.14*** (0.09)
Path length (\mathcal{L}_G)		-4.92*** (0.10)	-4.88*** (0.09)
Homophily (\mathcal{H}_G)		-0.31*** (0.04)	-0.30*** (0.03)
Average risk perception ($r_{\sigma,\pi}$)		0.01 (0.04)	0.03 (0.03)
Disease severity (σ)		-0.01 (0.04)	0.00 (0.03)
Infectivity (γ)		22.66*** (0.04)	22.66*** (0.03)
$\mathcal{L}_G \times \gamma$			-3.41*** (0.03)
Adjusted R ²	0.00	0.85	0.87
RMSE	26.44	10.29	9.71
Number of observations	80,305	80,305	80,305

Note: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, SEs in parentheses.



Appendix C

Supplements to Chapter 4

C.1 Communication with participants

C.1.1 Registration

Dear Sir, Madam,

we would like to invite you to participate in the study “Progression of epidemics in a networking game”. This study will help to expand our knowledge of networking behavior in the context of infectious diseases. The study will be conducted on *DATE*, at

- *TIME* BST/IST/WEST (Ireland, Portugal, UK)
- *TIME* CEST (Austria, Belgium, Croatia, Czechia, Denmark, France, Germany, Hungary, Italy, Luxembourg, Malta, Netherlands, Poland, Slovakia, Slovenia, Spain, Sweden)
- *TIME* EEST (Bulgaria, Cyprus, Estonia, Finland, Greece, Latvia, Lithuania, Romania).

You will receive a link to participate in the study about 10 minutes before the specified date and time via Prolific. Since this is an interactive study, you cannot join later than 10 minutes after the specified time.

Design/execution of the study

The form of this experiment is a large-scale interactive browser game imitating the spread of infectious diseases in social networks.

Background of the study

The COVID-19 pandemic has shown that our globally connected world is highly vulnerable to the spread of infectious diseases. This study will help us to understand how social networks change over the course of an epidemic.

What is expected of you as a participant

To participate in the study, you will need a desktop computer or laptop (no smartphones or tablets). With regard to the browser, we advise you to use Google Chrome or Mozilla Firefox. Using Internet Explorer is not recommended, as it may not support all features. Furthermore, we expect you to be available for about 1 hour and 40 minutes. Note that during the experiment, some other members may be waiting until you have completed a given task, before they can proceed. So, please be mindful of other participants and only sign up if you can complete all tasks in a timely fashion. Please only sign up if you can commit to completing it.

Possible advantages and disadvantages of the study

The study is an interactive browser game. All interactions and diseases are only virtual and do not have any consequences in the real world.

Remuneration/compensation

During the experiment, you can earn points. These points will be converted at the end of the experiment at the exchange rate of:

$$500 \text{ points} = \text{£}1.00 \text{ (€}1.17\text{)}$$

The number of points you earn depends on your own choices and the choices of other participants. Depending on your performance in the experiment, you will earn maximally £10.00 (€11.65). We will guarantee you a minimum payment of £5.00 (€5.83) per 60 minutes of your participation as required by Prolific, but if your earned points are beyond the minimum payment, you get the difference as your bonus. *For example, if you earn 3,000 points and the experiment takes 60 minutes, you earn £6.00: £5.00 for the minimum payment and £1.00 in addition as a bonus.*¹ The experiment will take a maximum of 100 minutes (1 hour 40 minutes), probably less. It may happen that you cannot be assigned to a group to participate in the study. In that case, you will receive the minimum amount as announced on Prolific.

Confidentiality of data processing

This study requires us to collect some of your personal data and demographics (age, gender, mother tongue, education, country, marital status, employment, household income). We need these data in order to be able to answer our research questions properly. Any personal data will all be anonymously stored separately from your Prolific ID so that they can never be traced back to you yourself. The computer on which your personal details is stored is secured to the highest standards, and only researchers involved will have access to this data. The data itself will also be protected by a security code. Your data will be stored for at least 10 years. This is in accordance with the guidelines provided by the VSNU Association of Universities in the Netherlands. Please refer to the website of the Authority for Personal Data: <https://autoriteitpersoonsgegevens.nl/nl/onderwerpen/avg-europese-privacywetgeving>, for more information about privacy.

Voluntary participation

Participation in this study is voluntary. You can end your participation in the study at any time, without any explanation and without any negative consequences. If you end your participation, we will use the data collected up to that point, unless you explicitly inform us otherwise.

¹Note that this example has been added after the first three sessions, due to many complaints about missing or incorrect bonus payments.

Questions, independent contact, and complaints officer

If you have any questions about the study, please contact the main researcher, Hendrik Nunner at h.nunner@uu.nl. If you have any comments about the study, please reach out to our independent contact Sanjana Singh at s.singh@uu.nl. If you have an official complaint about the study, you can email the complaints officer at klachtenfunctionaris-fetsocwet@uu.nl.

Consent

By checking the box and clicking the “Sign up” bottom below, I declare that I have read all information about the study “Progression of epidemics in a networking game” presented on this website and agree to participate in the study on on *DATE*, at

- *TIME* BST/IST/WEST (Ireland, Portugal, UK)
- *TIME* CEST (Austria, Belgium, Croatia, Czechia, Denmark, France, Germany, Hungary, Italy, Luxembourg, Malta, Netherlands, Poland, Slovakia, Slovenia, Spain, Sweden)
- *TIME* EEST (Bulgaria, Cyprus, Estonia, Finland, Greece, Latvia, Lithuania, Romania).

I know that I may at all times quit the study without any explanation or consequences.

With kind regards,

Hendrik Nunner (h.nunner@uu.nl),
Vincent Buskens (v.buskens@uu.nl),
Mirjam Kretzschmar (m.e.e.kretzschmar@umcutrecht.nl),
Rense Corten (r.corten@uu.nl)

Download the information and consent form co-signed by the main researcher here:

DOWNLOAD INFORMATION AND CONSENT FORM

I understand that I can only participate in the study at the specified date with the link that I will receive via Prolific.

SIGN UP

C.1.2 Registration confirmation

Thank you for signing up for the online academic study “Progression of epidemics in a networking game” on *DATE*, at

- *TIME* BST/IST/WEST (Ireland, Portugal, UK)
- *TIME* CEST (Austria, Belgium, Croatia, Czechia, Denmark, France, Germany, Hungary, Italy, Luxembourg, Malta, Netherlands, Poland, Slovakia, Slovenia, Spain, Sweden)
- *TIME* EEST (Bulgaria, Cyprus, Estonia, Finland, Greece, Latvia, Lithuania, Romania).

Please remember:

- You will receive a link to participate in the study about 10 minutes before the specified date and time via Prolific.
- You cannot join later than 10 minutes after the specified time.
- You will need a desktop computer or laptop (no smartphones or tablets).
- We advise you to use Google Chrome or Mozilla Firefox, as Internet Explorer may not support all features.

<i>BACK TO PROLIFIC</i>

C.1.3 Instructions

You are now taking part in a decision-making experiment. Please read the following instructions carefully. They contain everything you need to know to participate in the experiment. The data for this study is collected and controlled by Hendrik Nunner of the Utrecht University Institute (EUI). If you have any questions, please email Hendrik Nunner at h.nunner@uu.nl.

During the experiment, you can earn points. These points will be converted at the end of the experiment at the exchange rate of:

$$500 \text{ points} = \text{£}1.00 \text{ (€}1.17\text{)}$$

The number of points you earn depends on your own choices and the choices of other participants. Depending on your performance in the experiment, you will earn maximally £10.00 (€11.65). We will guarantee you a minimum payment of £5.00 (€5.83) per 60 minutes of your participation as required by Prolific, but if your earned points are beyond the minimum payment, you get the difference as your bonus. *For example, if you earn 3,000 points and the experiment takes 60 minutes, you earn £6.00: £5.00 for the minimum payment and £1.00 in addition as a bonus.*² The experiment will take a maximum of 100 minutes (1 hour 40 minutes), probably less. It may happen that you cannot be assigned to a group to participate in the study. In that case, you will receive the minimum amount as announced on Prolific. (Note that this is slightly adapted from the recruitment text, based on our experience and misunderstandings in an earlier session).

At the end of the experiment, you will be shown a message confirming your participation and a link redirecting you to Prolific when clicked on. Clicking this link is required to finalize your participation. Only when you completed all tasks and got redirected to Prolific, you will receive the money you earned during the experiment in your Prolific account.

Please note that you may at all times quit the study without any explanation or consequences.

PLEASE CLICK TO PROCEED

Part 1

In part 1 of the experiment, you will be confronted with a series of five situations. Each situation requires you to choose between two options. One option is to accept a sure payment of a particular number of points. The second option is to gamble with a 50:50 chance of getting 300 points or getting nothing. Realize that you will get paid according to your choice, and if you decide to gamble, the outcome indeed depends on a random flip of a coin which is executed by the computer.

The number of points you earn for part 1 is the average number of points related to the five decisions you made (score of each situation divided by five).

²Note that this example has been added after the first three sessions, due to many complaints about missing or incorrect bonus payments.

PLEASE CLICK TO PROCEED

Part 2

Overview

In part 2 of the experiment, you participate in two games. In each game, you are one person in a network of 60 participants. You are displayed as a square on your own screen. All other participants are shown as circles. We call your square and the circles of other participants “nodes”. Relations between participants are indicated as lines between the nodes. Each game consists of about 15 to 20 rounds. Before round 1, a starting network is generated on the screen and one participant in the network will be infected with a (hypothetical) disease. The node of this participant will be colored red. Nodes that have relations with an infected (red) node can also get infected in subsequent rounds.

Game rounds

Each game consists of about 15 to 20 rounds. Each round consists of 4 stages.

Stage 1. In this stage, nodes can get infected. In the first round, only one node is infected. In all other rounds, other nodes can get infected. The probability of getting infected depends on the number of red nodes with whom you are connected. If you are not connected with a red node, you cannot get infected. If you are connected to 1 red node, the probability of infection is 0.15 and increases for each additional infected node (see below).

Table C.1: Chance of infections.

# of infectious neighbors	0	1	2	3	4	5	6	7	8	9	10
chance of getting infected	0.00	0.15	0.28	0.39	0.48	0.56	0.62	0.68	0.73	0.77	0.80

If you get infected, your node will also turn red, and you will stay infected for exactly 4 rounds. After these four rounds, your node will turn green, you will get immune, and you cannot get the disease anymore in this game. All other immune nodes will also be colored green.

Stage 2. In stage 2, you can break relations or propose new relations in the network. There will be 12 nodes displayed on the right side of the screen, for which you can either choose to break the relation or to propose a new relation. The buttons to make the decisions are also colored (red if that node is infected with the disease, green if that node is immune and gray otherwise). Breaking relations you can decide yourself, but before you can establish a new relation, the other participant needs to agree on this relationship as well. You will have exactly 60 seconds to make your decisions in this stage in each round. After that, the game proceeds automatically.

Stage 3. In stage 3, you can indicate which relations that others have proposed to you, you indeed also would like to have. These are again displayed on the right side of the screen and colored as explained before. Other participants will simultaneously decide whether to accept your proposals.

Stage 4. In stage 4, your points for this round are calculated (see Earning points).

Note that stages 1 (disease transmission) and 4 (points computation) are done by the system and the outcomes will be presented as information on the screen (color of nodes, number of points). So in practice, you will move back and forth between stages 2 (breaking / proposing relations) and 3 (accepting proposed relations).

Earning points

How many points you earn depends on three things:

1. **Your relations:** If you have 0 relations, you earn 0 points. The first relation earns you quite some points, the second a bit less, and so on until you can earn maximally 100 points for 6 relations. If you have more than 6 relations, the number of points you earn will decrease again.

Table C.2: Points per number of relations.

# of relations	0	1	2	3	4	5	6	7	8	9	10	11	12	13	...
points per round	0	31	56	75	89	97	100	97	89	75	56	31	0	0	...

2. **Relations between nodes connected to you:** The two games you play will differ in how many points you earn based on the relations between your connections. The game you are playing will be indicated on the screen.

Game A You will be awarded 20 additional points if there are no relations between any of the nodes you are connected to. For each relation between these nodes, you will lose some of these 20 points.

Game B You will be awarded 20 additional points if you are connected to 5 other nodes that all have relations among each other as well, while you have one relation to another node not in this group. If your connections have fewer relations among each other, you lose some of these 20 points.

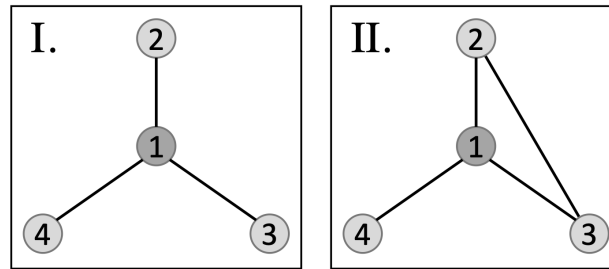


Figure C.1: **Illustration of clustering settings.** (I.) No relation between the connections of node 1. (II.) One relation between connections (2 and 3) of node 1, while one connection (4) is to a node not in this group.

3. **Having the disease:** 14 points will be deducted from the points you earn from your relations for each round that you are infected.

User interface

Figure 4.2 was shown accompanied by an instructional video to explain the elements of the user interface. The video is available upon request.

Final notes

We will first run a trial game with 3 rounds, in which you can try out the interface. You do not earn anything in the trial game, but you can get a feel for the game and experience how you earn points. After that, we start with the first paid game.

If you have questions or do not understand something, please email the main researcher Hendrik Nunner at h.nunner@uu.nl. Otherwise, click the continue button on the screen.

PLEASE CLICK TO PROCEED

C.1.4 Pop-up Game A

Disease transmission

The probability of getting infected depends on the number of infected (red) nodes with whom you are connected:

Table C.3: Chance of infections.

# of infectious neighbors	0	1	2	3	4	5	6	7	8	9	10
chance of getting infected	0.00	0.15	0.28	0.39	0.48	0.56	0.62	0.68	0.73	0.77	0.80

Infected nodes turn red (including your own node) and remain infected for 4 consecutive rounds. After 4 rounds, infected nodes turn green (including your own node), become immune, and can therefore not get infected again.

Earning points in Game A

How many points you earn depends on three things:

1. **Your relations:** The number of points you can earn depends on the number of relations you have. Six relations is optimal. Fewer or more relations will result in fewer points.

Table C.4: Points per number of relations.

# of relations	0	1	2	3	4	5	6	7	8	9	10	11	12	13	...
points per round	0	31	56	75	89	97	100	97	89	75	56	31	0	0	...

2. **Relations between nodes connected to you:** You will be awarded 20 additional points if there are no relations between any of the nodes you are connected to. For each relation between these nodes, you will lose some of these 20 points.

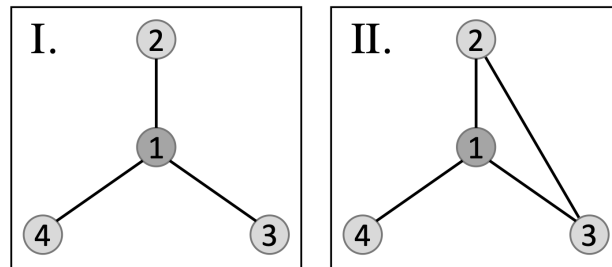


Figure C.2: **Illustration of clustering settings (Pop-up Game A).** (I.) No relation between the connections of node 1. (II.) One relation between connections (2 and 3) of node 1, while one connection (4) is to a node not in this group.

3. **Having the disease:** 14 points will be deducted from the points you earn from your relations for each round that you are infected.

C.1.5 Pop-up Game B

Disease transmission

The probability of getting infected depends on the number of infected (red) nodes with whom you are connected:

Table C.5: Chance of infections.

# of infectious neighbors	0	1	2	3	4	5	6	7	8	9	10
chance of getting infected	0.00	0.15	0.28	0.39	0.48	0.56	0.62	0.68	0.73	0.77	0.80

Infected nodes turn red (including your own node) and remain infected for 4 consecutive rounds. After 4 rounds, infected nodes turn green (including your own node), become immune, and can therefore not get infected again.

Earning points in Game B

How many points you earn depends on three things:

1. **Your relations:** The number of points you can earn depends on the number of relations you have. Six relations is optimal. Fewer or more relations will result in fewer points.

Table C.6: Points per number of relations.

# of relations	0	1	2	3	4	5	6	7	8	9	10	11	12	13	...
points per round	0	31	56	75	89	97	100	97	89	75	56	31	0	0	...

2. **Relations between nodes connected to you:** You will be awarded 20 additional points if you are connected to 5 other nodes that all have relations among each other as well, while you have one relation to another node not in this group. If your connections have fewer relations among each other, you lose some of these 20 points.

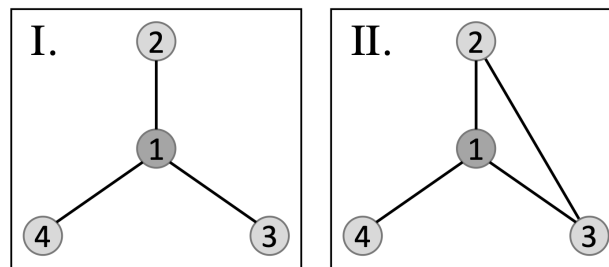


Figure C.3: **Illustration of clustering settings (Pop-up Game B).** (I.) No relation between the connections of node 1. (II.) One relation between connections (2 and 3) of node 1, while one connection (4) is to a node not in this group.

3. **Having the disease:** 14 points will be deducted from the points you earn from your relations for each round that you are infected.

C.1.6 Survey

- What is your age? *NUMBER FIELD*
- What is your gender? *DROP-DOWN FIELD*
 - 1 Female
 - 2 Male
 - 3 Other
- What is your mother tongue? *DROP-DOWN FIELD*
 - 1 English
 - 2 Other
- What is your highest completed level of education? *DROP-DOWN FIELD*
 - 1 No formal education
 - 2 Primary school (elementary education)
 - 3 Lower secondary (secondary completed does not allow entry to university: obligatory school)
 - 4 Upper secondary (programs that allows entry to university)
 - 5 Post secondary, non-tertiary (other upper secondary programs toward labor market or technical formation)
 - 6 Lower level tertiary, first stage (also technical schools at a tertiary level)
 - 7 Upper level tertiary (Master, Doctor)
 - 8 No answer
- What is your country of residence? *DROP-DOWN FIELD*
 - 93 Afghanistan
 - 355 Albania
 - 213 Algeria
 - ...
 - 263 Zimbabwe
- How concerned were/are your to attract COVID-19?
 - 1 Very
 - 2 A bit
 - 3 Hardly
 - 4 Not at all

- Did you get COVID-19?

0 No

1 Yes

C.1.7 Cannot be assigned

Dear participant,

thank you for participating in the academic study “Progression of epidemics in a networking game”! As explained in the instructions, participants are assigned to groups. This means that when not enough people show up to form a group of 60, not everybody gets to participate.

Unfortunately, you could not be assigned to a group.

The minimum payment as announced in the study will be transferred to you shortly. Please make sure to exit the study by clicking the “Back to Prolific” button below.

Kind regards,

Hendrik Nunner

<i>BACK TO PROLIFIC</i>

C.2 Additional definitions

C.2.1 Simulations to determine parameter settings for the experiment (*simulation 1*)

Determining parameter settings for the experiment was a four-step process. In the first step, we selected conditions ($LO:RA$, $LO:AS$, $HI:RA$, $HI:AS$) and fixed parameters for each setting (clustering: LO , HI ; social mixing: RA , AS) that we deemed feasible for the money available to conduct the experiment. Our goal was to use the largest networks possible with enough sessions to achieve sufficient statistical power. Our time limit was 90 minutes per session (resulting in at least £7.50 per participant), with two installments of the NIDT per session. Within this time limit, we estimated about 20 minutes for reading the instructions, about 30 minutes per NIDT, and 10 minutes buffer. For a single installment of the NIDT, we estimated 1 minute for stage 1 and 30 seconds for stage 2 of each round. Consequently, each installment of the NIDT had to stop after a maximum of 20 rounds. Preliminary simulations showed that parameter settings presented in Table C.7 and settings for risk perception parameters that were taken from a previous study by Vriens and Buskens (2020), which used the same staircase task to measure risk avoidance³, provided the best results. That is, the disease had sufficient time to spread, so that after 20 rounds most simulations had no more infected nodes in the network. Furthermore, simulations created epidemics of different size depending on settings for clustering and social mixing.

In the second step, we selected the baseline networks used for the two clustering settings. That is, we first generated 50 networks with a preferred proportion of 0.0 closed triads, and 50 networks with a preferred proportion of 0.667 closed triads. The remaining parameters were set as presented in Table C.7. To make sure that the initial networks were as similar as possible, except for clustering, we then selected one predefined network for each clustering setting (see Figure 5.2) so that they had the same largest available average degree (5.933) and the closest available value for closeness (LO : 0.975, HI : 0.955). Furthermore, we defined fixed orders for assigning participants to nodes based on their risk aversion score. That is, neighbors in the random social mixing setting have mostly different risk perception settings, while neighbors in the assortative mixing setting have mostly similar risk perception settings.

In the third step, we simulated the entire data collection, with 48 experimental sessions and two installments of the NIDT per experimental session. We also made sure that the agents between the two installments of the NIDT were not reinitialized with new risk perception parameters, but only assigned to the network positions according to the social mixing setting. The results of these simulations are denoted in the results as *simulation 1*.

³We rescaled the reported scores of $M = 19.7$ and $SD = 6.98$ to $M = 1.27$, $SD = 0.45$ to match the risk perception parameter in the simulations.

Table C.7: Parameter ranges used for network generation and determination of game parameters, as well as settings used for game parameters (*LO*, *HI*, *RA*, *AS*).

Parameter	Range	Setting	Use
<i>Utility/points</i>			
<i>I. Social network</i>			
Benefit per social tie [*]	$b_1 = 1.0$	1.0	all
Simple cost per tie [*]	$c_1 = 0.2$	0.2	all
Marginal cost per tie [*]	$c_2 \in \{0.1, 0.067, 0.05\}$	0.067	all
Benefit for triadic closure	$b_2 = 0.5$	0.5	all
Preferred proportion of closed triads [†]	$\alpha \in \{0.0, 0.667, 0.75\}$	0.0 0.667	<i>LO</i> , <i>RA</i> , <i>AS</i> <i>HI</i> , <i>RA</i> , <i>AS</i>
<i>II. Disease</i>			
Disease severity	$\sigma > 0$	0.34	all
Infectivity	$0 \leq \gamma \leq 1$	0.15	all
Recovery time in time steps	$\tau \in \mathbb{Z}^+$	4	all
<i>Homophily</i>			
Likelihood of ties similar in risk perception	$0 \leq \omega \leq 1$	0.0 0.8	<i>LO</i> , <i>HI</i> , <i>RA</i> <i>LO</i> , <i>HI</i> , <i>AS</i>
Network / offerings			
Network size	$N \in \{12, 24, 40, 60, 80\}$	60	all
Proportion of N offered per time step	$0 \leq \phi \leq 1$	0.2	all
Proportion of ϕ as distance 1 ties	$0 \leq \psi \leq 1$	0.5	all
Proportion of ϕ as distance 2 ties	$0 \leq \xi \leq 1$	0.3	all

Notes: ^{*}: The combination of $b_1 = 1.0$, $c_1 = 0.2$, and $c_2 \in \{0.1, 0.067, 0.05\}$ sets the preferred number of ties to 4, 6, and 8 respectively. [†]: The setting of $\alpha = 0.667$ with a preferred number of 6 ties ($c_2 = 0.067$) realizes a preference of 4 out of 6 neighbors sharing connections.

In the fourth step, we adjusted point rewards to create easy-to-understand rules. For this, we took the optimum of 6 ties as the basis and rather than awarding $b_1 \cdot t_i - (c_1 \cdot t_i + c_2 \cdot t_i^2) = 1.0 \cdot 6 - (0.2 \cdot 6 + 0.067 \cdot 6^2) = 2.388$ points, we used a factor of 41.55 to award 100 points. Rewards for clustering and costs for being infected were rescaled accordingly.

C.2.2 Additional simulations (*simulation 2*)

To get a better insight on how decision-making differed between the agents in our initially simulated experiments (*simulation 1*) and participants in the experiment, we performed an additional series of simulated experiments (*simulation 2*). We used the same setup as in the experiment, with as many parameters set according to data from our experiment as possible (see Section 4.2.3). Two settings, however, could not be immediately set so that we had to extend our model in two ways. First, we added a parameter that allows to set the probability to make a decision decreasing (rather than increasing or maintaining) point rewards ($\epsilon = 0.66$):

- Repeat until all decision opportunities have been processed:
 - If alter is infected or $\epsilon \leq U[0, 1]$:
 - Make decision to maximize point rewards using Equation 4.1
 - else:
 - Make decision to decrease point rewards using Equation 4.1

Second, we extended the third term of Equation 4.1 not only with a risk perception parameter ($0.0 < r_i < 2.0$) for susceptible agents as done in simulation 1 (see Equation 4.3), but also with an overestimation factor (η):

$$\left[\sigma \right] \rightarrow (\pi_i^{2-r} \cdot \sigma^r) \cdot \eta \quad (\text{C.1})$$

We set η to a value of 2.5 to achieve a similar strong avoidance reaction of susceptible agents towards infected agents, as observed in the experiment (see Section C.3.3). Note that we used an overestimation factor rather than inflating risk parameter (r). That is because inflating r would, on the one hand, require a value (1.8) that impedes variation between agents ($0.0 \leq r \leq 2.0$) without creating a long not observed tail towards lower values. On the other hand, artificially inflated risk scores do not correspond to the scores measured using the staircase task.

C.3 Additional results

C.3.1 Demographic data of the participants

Table C.8: Demographic data of the participants.

	%	N
<i>Gender</i>		
Female	41.2	1,185
Male	54.5	1,570
Other	1.0	30
No answer	3.3	94
<i>Age</i>		
18 to 29	72.0	2,072
30 to 39	16.7	482
40 to 49	6.1	177
50 to 59	2.7	79
60 and above	0.6	18
No answer	1.8	51
<i>Mother tongue</i>		
English	16.5	475
Other	80.8	2,325
No answer	2.7	79
<i>Education</i>		
No formal education	0.1	2
Primary school	1.0	28
Obligatory school	4.0	114
Secondary (allowing entry to university)	42.3	1,219
Tertiary (Bachelor, Maser, PhD)	49.5	1,424
No answer	3.2	92
<i>COVID-19 concern</i>		
Very	27.9	802
A bit	47.4	1,365
Hardly	13.8	398
Not at all	8.7	250
No answer	2.2	64
<i>COVID-19 positive (at some point)</i>		
Yes	13.8	397
No	84.4	2,429
No answer	1.8	53

Table C.9: Countries of residence of the participants.

	%	N
Portugal	22.5	647
Poland	21.2	609
UK	14.2	408
Italy	12.1	348
Spain	6.4	183
Greece	5.9	171
Hungary	2.6	76
Ireland	1.8	53
France	1.6	45
Estonia	1.4	41
Netherlands	1.3	38
Czech Republic	1.2	35
Slovenia	1.1	31
Germany	1.0	29
Belgium	0.9	25
Latvia	0.8	24
Sweden	0.8	23
Finland	0.7	21
Austria	0.2	5
Denmark	0.1	4
Luxembourg	0.1	3
USA/Canada/Caribbean	0.1	1
Thailand	0.1	1
New Zealand	0.1	1
No answer	2.0	57

C.3.2 Effect of settings and conditions on epidemics

Figure C.4 shows an extended version of Figure 4.4. As discussed in Chapter 4, we observe no significant differences of outcome measures (final size, duration, peak size) in the experiment. In simulation 1, however, the simulation performed to determine parameter settings for the experiment, we observe smaller and lower peaking epidemics for the baseline network consisting of multiple clusters (column 1). Furthermore, increase in social mixing shows a tendency for smaller, shorter, and lower peaking epidemics (column). In combination (column 3), low clustering and random mixing produce on average the largest, longest, and highest peaking epidemics, while high clustering and assortative mixing produce on average the smallest, shortest, and lowest peaking epidemics. After adjusting the parameter settings to data from the experiment (simulation 2), we observe that the epidemics remain on average larger, longer, and higher peaking as compared to the experiment, while the effect of conditions on epidemics disappears.

C.3.3 Avoidance of infected alters

Figure C.5 shows how quickly susceptible egos disconnected from infected alters depending on the number of infected neighbors of the ego ((a)) and how this affects the average degree of infected nodes over the course of an infection ((b)). We see that avoidance of infected nodes is much stronger in the experiment than in simulation 1. By increasing the perceived risk of getting infected and perceived severity of the disease by a factor of 2.5, we increase the avoidance behavior of agents to a degree similar to participants in the experiment (see Figure C.5 (a)). This alone, however, does not suffice to achieve a similar strong isolation of infected nodes as observed in the experiment (see Figure C.5 (b)).

C.3.4 Factors contributing to networking decisions

Table C.10 shows a comparison of factors contributing to networking decisions. Model 1 describes factors contributing to the attractiveness of a relation. We define a relation to be attractive as a binary variable composed of decisions to create a relation (proposals in stage 1 and accepted proposals in stage 2 of a round) and declined opportunities to dissolve a relation. Models 2, 3, and 4 describe factors contributing to whether a decision opportunity (proposing a relation to an alter in stage 1 of a round, accepting a proposal of an alter in stage 2 of a round, dissolving a relation) was taken.

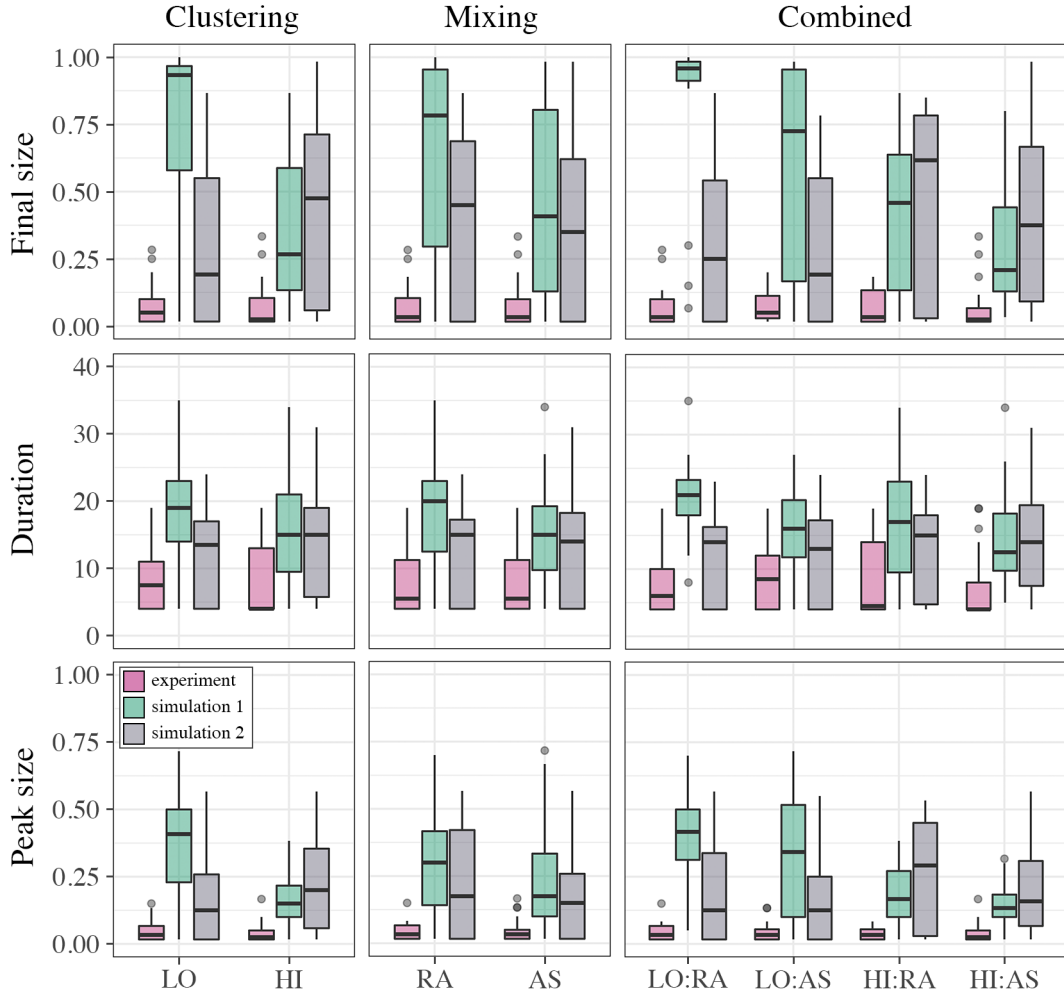


Figure C.4: **Effects of settings and conditions on epidemics.** Box-and-whisker plots show the median, interquartile range, minimum, maximum, outliers of final size (proportion of cumulatively infected nodes; row 1), duration (in rounds; row 2), and peak size (maximum proportion of nodes infected at the same time; row 3) by settings for clustering (column 1) and social mixing (column 2), as well as conditions (column 3).

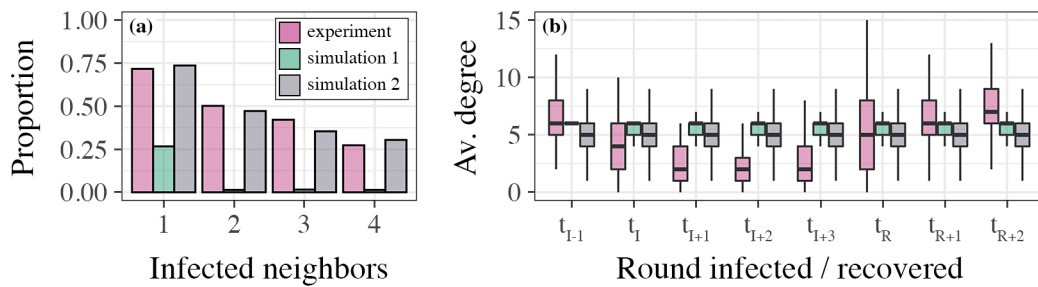


Figure C.5: **Avoidance of infected alters.** Plot (a) shows the proportion of performed relation dissolutions to dissolution opportunities of susceptible egos and infected alters depending on the number of infected neighbors. Plot (b) shows the average degree of nodes from one round before infection (t_{I-1}) until two rounds after recovery (t_{R+2}). Both plots are divided by data source (pink: experiment, green: simulation 1, purple: simulation 2).

Table C.10: Three-level random-intercept logistic regression of networking decisions.

	Model 1 (rel. attractive)	Model 2 (rel. proposed)	Model 3 (proposal accepted)	Model 4 (rel. dissolved)
Constant	5.866*** (0.074)	1.050*** (0.115)	3.013*** (0.160)	-5.927*** (0.133)
Disease states				
Ego ^S - Alter ^S without infected neighbor	-0.103* (0.059)	-0.522*** (0.088)	-0.217 (0.135)	0.025 (0.107)
Ego ^S - Alter ^S with infected neighbor	-1.070*** (0.061)	-1.350*** (0.091)	-0.593*** (0.138)	1.532*** (0.108)
Ego ^S - Alter ^I	-3.728*** (0.067)	-3.626*** (0.108)	-3.732*** (0.149)	4.523*** (0.118)
Ego ^S - Alter ^R	1.070*** (0.063)	1.018*** (0.092)	0.600*** (0.145)	-1.060*** (0.128)
Ego ^I - Alter ^S without infected neighbor [†]	0.200*** (0.062)	-0.035 (0.090)	0.199 (0.233)	
Ego ^I - Alter ^S with infected neighbor	-0.014 (0.079)	0.157 (0.113)	1.398*** (0.488)	0.447*** (0.130)
Ego ^I - Alter ^I	-1.347*** (0.094)	-0.307** (0.144)	-1.119*** (0.222)	2.367*** (0.139)
Ego ^I - Alter ^R	1.949*** (0.114)	2.282*** (0.148)	1.479*** (0.347)	-0.500* (0.270)
Ego ^R - Alter ^S without infected neighbor	-0.886*** (0.057)	-1.059*** (0.087)	-0.744*** (0.132)	0.779*** (0.107)
Ego ^R - Alter ^S with infected neighbor	-0.813*** (0.083)	-0.884*** (0.122)	-0.578*** (0.176)	1.078*** (0.163)
Ego ^R - Alter ^I	-1.560*** (0.084)	-1.411*** (0.150)	-1.136*** (0.171)	1.730*** (0.139)
Ego ^R - Alter ^R (reference)				
Risk aversion				
Risk aversion score (RAS)	0.185 (0.172)	0.602*** (0.226)	0.406 (0.403)	0.648* (0.340)
Ego ^S - Alter ^S without infected neighbor × RAS	-0.037 (0.161)	-0.445** (0.210)	-0.189 (0.402)	-0.636* (0.337)
Ego ^S - Alter ^S with infected neighbor × RAS	-0.130 (0.179)	-0.611** (0.240)	-0.220 (0.437)	-0.604* (0.355)
Ego ^S - Alter ^I × RAS	-0.833*** (0.235)	-1.429*** (0.392)	-1.664*** (0.510)	-0.083 (0.425)
Ego ^S - Alter ^R × RAS	0.182 (0.199)	-0.366 (0.252)	0.157 (0.505)	-0.889* (0.533)
Network properties				
Degree	-0.492*** (0.007)	-0.435*** (0.010)	-0.465*** (0.018)	0.613*** (0.017)
Degree ²	0.020*** (0.0004)	0.020*** (0.001)	0.022*** (0.001)	-0.022*** (0.001)
Degree of alter	-0.043*** (0.001)	-0.030*** (0.002)	-0.040*** (0.003)	0.027*** (0.003)
Clustering (<i>HI</i>)	0.026 (0.032)	-0.035 (0.043)	-0.213*** (0.056)	0.007 (0.037)
<i>HI</i> × <i>ECC</i> [†]	-10.258*** (0.106)	-10.570*** (0.142)	-14.200*** (0.277)	8.032*** (0.252)
<i>LO</i> × <i>ECC</i> [†]	6.655*** (0.137)	3.872*** (0.151)	2.491*** (0.287)	-27.789*** (0.461)
Decision opportunity types				
Number of offers by opportunity type	-0.098*** (0.003)	-0.029*** (0.005)	-0.342*** (0.009)	-0.045*** (0.007)
Creation opportunity in stage 1 of a round	-4.126*** (0.010)			
Creation opportunity in stage 2 of a round	-3.299*** (0.015)			
Log Likelihood	-286,474	-153,387	-43,380	-74,957
AIC	573,006	306,827	86,815	149,967
BIC	573,339	307,119	87,065	150,240
Observations	711,159	360,493	79,360	271,306
Number of groups (rounds)	52,650	52,650	39,403	52,441
Number of groups (nodes)	5,758	5,758	5,736	5,758
Number of groups (games)	96	96	96	96
Variance (rounds)	0.28	0.47	0.31	0.31
Variance (nodes)	0.87	1.65	1.25	0.98
Variance (games)	0.01	0.01	0.04	0.01

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$, SEs in parentheses. [†]*ECC*: Expected Change in Clustering. Models are three-level random-intercept logistic regression of whether (Model 1) a relation is attractive (0: declined opportunities to create a relation in stage 1 and 2 of a round, and accepted opportunities to break a relation, 1: accepted opportunities to create a relation in stage 1 and 2 of a round, and declined opportunities to break a relation), (Model 2) an opportunity to create a relation in stage 1 was accepted, (Model 3) an opportunity to create a relation in stage 2 was accepted, and (Model 4) an opportunity to dissolve a relation was accepted.

C.3.5 Progression of properties over time (extended)

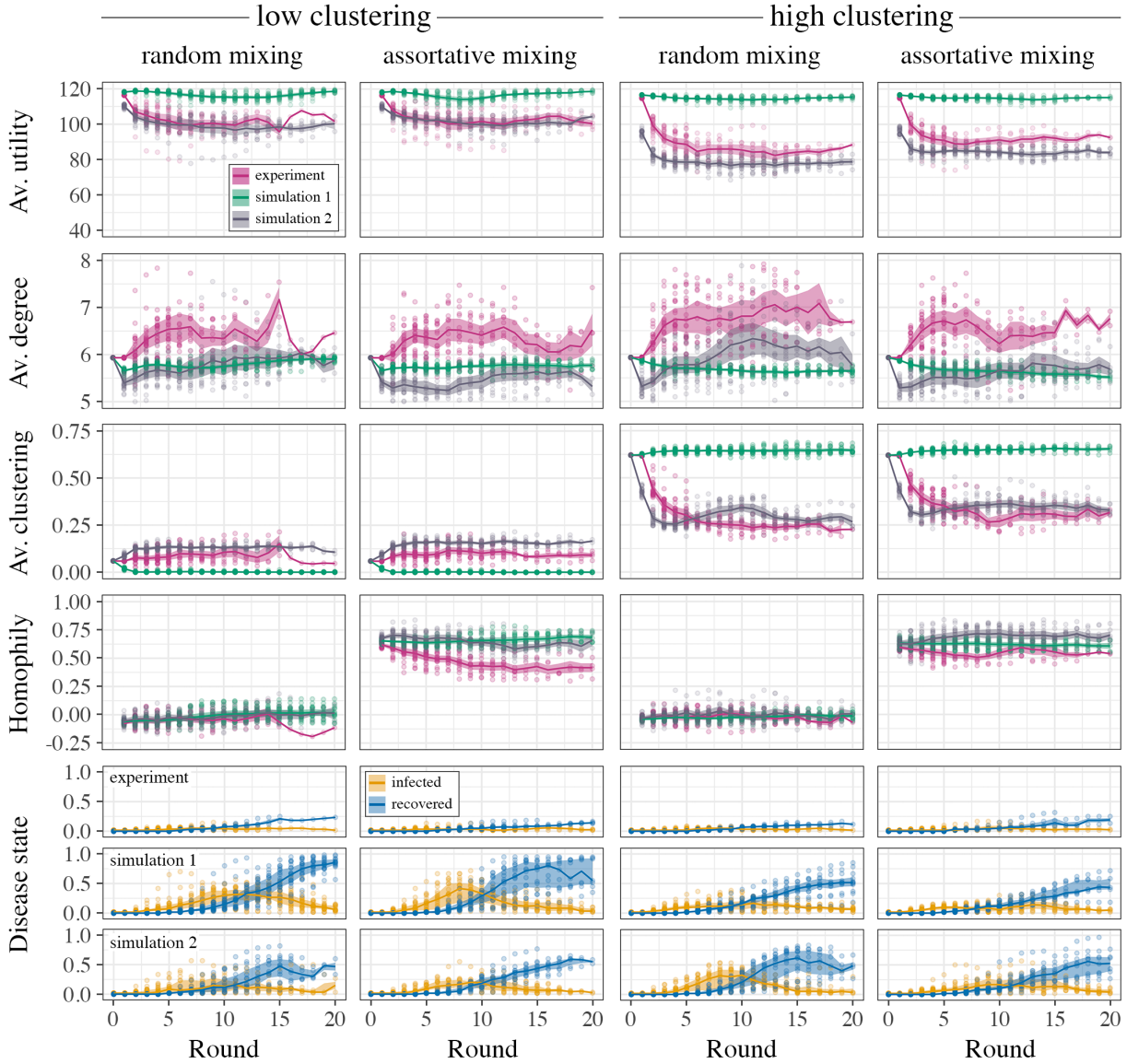


Figure C.6: **Progression of properties over time (extended)**. Plots are divided by different properties (rows) and settings for clustering and social mixing (columns). Average utility (row 1), average degree (row 2), average clustering (row 3), and degree of homophily (row 4) depict the results for the experiment (pink), simulation 1, which has been performed before the experiment (green), and experiment 2 with parameter settings taken from the experiment (purple). Progression of disease states has been divided into individual rows per data source (experiment: row 4, simulation 1: row 5, simulation 2: row 6). Each plot shows the progression of the corresponding property over time (max. 20 rounds). Dots within the plots show the mean of all nodes per single NIDT. Lines depict the mean and ribbons the standard deviation for NIDT means. Note that this is an extended version of Figure 4.5.

C.3.6 Effect of parameter variations on network structure and epidemic size

Fig. C.7 shows the relationships between parameters that have been set in simulation 2 based on the data collected in the experiment and affected outcome measures (i.e., average degree, average clustering, final size of epidemics). The data is based on 8,000 additional simulations with random variations of marginal costs for the number of relations (c_2), the probability to make a random rather than utility maximizing decision (ϵ), and the factor overestimating the perceived risk of getting infected and the perceived severity of an infection (η). We observe that higher marginal costs, and lower probabilities of random decisions have negative effects on the number of relations. Furthermore, completely random decisions ($\epsilon \approx 1.0$) result in an average clustering of around 0.2 in both clustering settings, while decisions based fully only utility maximization ($\epsilon \approx 0.0$) result in higher average clustering in the *HI* setting ($M = 0.61$, $SD = 0.02$) than in the *LO* setting ($M = 0.16$, $SD = 0.04$). Finally, higher values for risk overestimates reduce the final size. Additionally, diseases do not spread at all when agents perceive the risk of getting infected and the severity of the disease about three times the unaltered perceived value ($\eta = 3.0$).

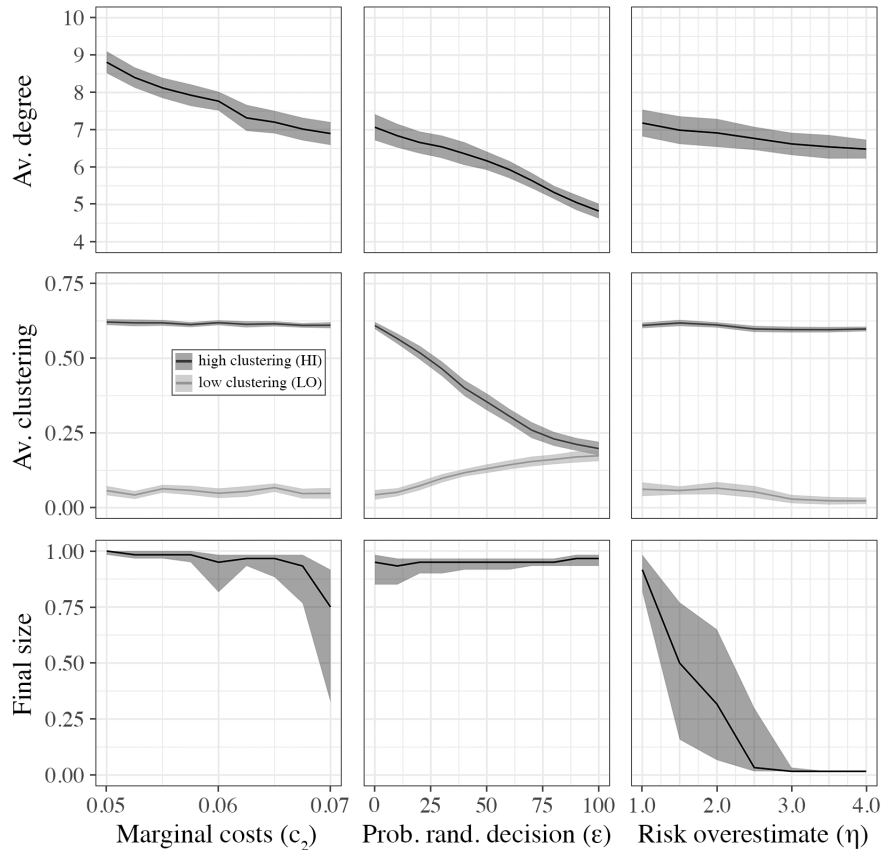
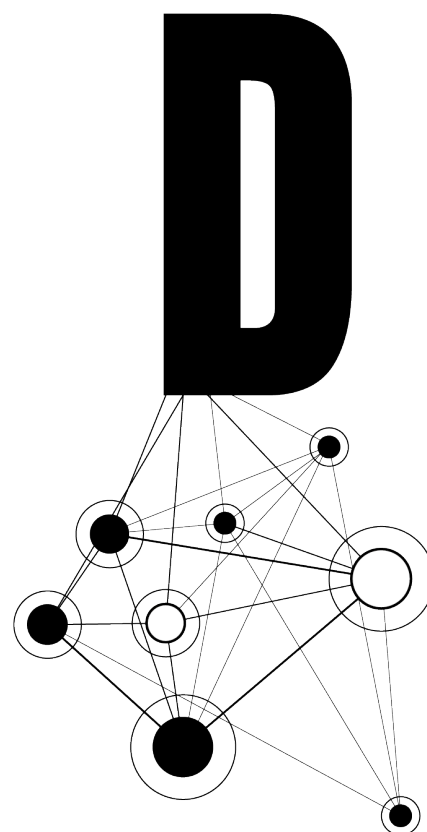


Figure C.7: **Parameter variations and their effects in simulations.** Plots are divided by different outcome measures (rows) and simulation parameters (columns). Average clustering is divided into the two clustering settings (*LO*, *HI*). Lines depict mean, ribbons standard deviations of the corresponding outcome measure.



Appendix D

Supplements to Chapter 5

D.1 Supplementary results

Table D.1 reports detailed descriptive statistics on the properties of input networks, initially infected nodes/*index cases*, and course of epidemics split by normal vs. lockdown networks, as well as course of epidemics split by counter measures.

Tables D.2 – D.5 report how well size and degree of occupational groups in the generated networks match the empirical data.

Table D.6 reports the average final sizes of epidemics across conditions (columns) and control variables (rows). These data are the basis for the differences of final size between the baseline and test conditions presented in Table 5.2.

Table D.7 reports the average peak sizes of epidemics across conditions (columns) and control variables (rows). Table D.8 reports the average peak sizes of epidemics of the baseline condition (column 2) and differences by test condition in percent points (columns 3 – 5), similar to differences of final size reported in Table 5.2.

Table D.9 reports the average duration of epidemics across conditions (columns) and control variables (rows). Table D.10 reports the average duration of epidemics of the baseline condition (column 2) and differences by test condition in percent points (columns 3 – 5), similar to differences of final size reported in Table 5.2.

Table D.1: Descriptive statistics of input networks for the simulations ($N = 133,200$).

	Mean	SD	Median	Min	Max	Skew
Network						
<i>I. Normal</i>						
Av. degree (\mathcal{D}_G)	4.92	0.21	4.92	4.46	5.3	-0.18
Clustering (\mathcal{C}_G)	0.38	0.1	0.4	0.21	0.56	-0.18
Av. path length (\mathcal{L}_G)	7.72	0.82	7.7	6.38	9.97	0.32
Assortativity, occupation (\mathcal{A}_G^p)	0.12	0.05	0.09	0.06	0.2	0.29
<i>II. Lockdown</i>						
Av. degree (\mathcal{D}_G)	2.1	0.02	2.1	2.06	2.16	0.52
Clustering (\mathcal{C}_G)	0.12	0.03	0.11	0.06	0.18	0.16
Av. path length (\mathcal{L}_G)	13.83	1.55	14.01	11.28	18.31	0.11
Assortativity, occupation (\mathcal{A}_G^p)	0.07	0.02	0.06	0.04	0.11	0.39
Index case						
<i>I. Normal</i>						
Degree	4.53	3.57	4	1	85	3.96
Clustering	0.39	0.26	0.33	0	1	0.3
Closeness	0.96	0.12	0.98	0	1	-7.58
Assortativity (occupation)	0.09	0.18	0	0	1	2.4
<i>II. Lockdown</i>						
Degree	1.98	1.76	2	0	26	1.73
Clustering	0.12	0.26	0	0	1	2.41
Closeness	0.44	0.32	0.64	0	0.74	-0.6
Assortativity (occupation)	0.06	0.18	0	0	1	3.69
Epidemic						
<i>I. Normal</i>						
Final size	50.8	35.86	62.11	0.01	98.24	-0.42
Duration	72.24	46.4	72	10	484	1.11
Epidemic peak size	2,908.22	2,466.26	2,886	1	8,395	0.22
<i>II. Lockdown</i>						
Final size	7.14	15.34	0.03	0.01	61.95	1.96
Duration	46.3	60.5	16	10	511	1.85
Epidemic peak size	193.57	460.08	3	1	2,716	2.64
Counter measures						
<i>I. Vaccinations</i>						
Vaccinated (proportion)	0.25	0.16	0.2	0	0.5	0.16
Immunized (proportion)	0.19	0.13	0.18	0	0.46	0.37
<i>II. Baseline</i>						
Final size	58.81	39.63	59.64	0.01	98.24	-0.51
Duration	66.73	41.91	63	10	208	0.49
Peak size	3,787.89	3,188.37	2,454.5	1	8,395	0.05
<i>III. Random vaccine distribution</i>						
Final size	35.72	35.11	34.58	0.01	94.85	0.27
Duration	62.85	49.21	63	10	394	0.93
Peak size	2,077.23	2,425.01	899	1	7,978	0.78
<i>IV. Vaccine distribution prioritizing occupational groups by largest degree</i>						
Final size	20.56	32.52	0.06	0.01	94.35	1.18
Duration	55.27	61.39	22	10	511	1.64
Peak size	900.29	1,682.96	5	1	7,130	1.88

Table D.2: Comparison of occupational groups in empirical and generated networks (1).

	Mean	SD	Median	Min	Max	Skew
<i>Architecture and Engineering Occupations</i>						
Share of the labor market (empirical)	0.01					
Share of the labor market (generated)	0.01	0	0.01	0.01	0.02	0.09
Degree, normal (empirical)	3.47	8.99	0	0	150	8.35
Degree, normal (generated)	3.6	0.9	4	0	6	-1
Degree, lockdown (empirical)	1.88	6.31	0	0	100	8.73
Degree, lockdown (generated)	1.93	1.21	2	0	5	0.26
<i>Arts, Design, Entertainment, Sports, and Media Occupations</i>						
Share of the labor market (empirical)	0.01					
Share of the labor market (generated)	0.01	0	0.01	0.01	0.02	-0.19
Degree, normal (empirical)	3.23	11.03	0	0	200	10.5
Degree, normal (generated)	3.32	0.82	3	0	5	-0.85
Degree, lockdown (empirical)	2.51	21.18	0	0	500	20.67
Degree, lockdown (generated)	2.58	1.04	3	0	5	-0.39
<i>Building and Grounds Cleaning and Maintenance Occupations</i>						
Share of the labor market (empirical)	0.03					
Share of the labor market (generated)	0.03	0	0.03	0.02	0.03	0.22
Degree, normal (empirical)	3.93	15.38	0	0	100	5.46
Degree, normal (generated)	3.94	0.98	4	0	6	-1.04
Degree, lockdown (empirical)	1.26	4.95	0	0	60	9.41
Degree, lockdown (generated)	1.3	1.12	1	0	6	0.72
<i>Business and Financial Operations Occupations</i>						
Share of the labor market (empirical)	0.04					
Share of the labor market (generated)	0.04	0	0.04	0.04	0.05	-0.09
Degree, normal (empirical)	3.66	15.98	0	0	400	15.75
Degree, normal (generated)	3.72	0.93	4	0	6	-1.03
Degree, lockdown (empirical)	1.52	5.49	0	0	100	11.51
Degree, lockdown (generated)	1.56	1.17	1	0	6	0.52
<i>Community and Social Service Occupations</i>						
Share of the labor market (empirical)	0.01					
Share of the labor market (generated)	0.01	0	0.01	0.01	0.02	0.24
Degree, normal (empirical)	2.84	8.85	0	0	100	7.43
Degree, normal (generated)	2.91	0.66	3	0	5	-1.46
Degree, lockdown (empirical)	1.08	2.61	0	0	20	4.38
Degree, lockdown (generated)	1.11	0.95	1	0	4	0.51
<i>Computer and Mathematical Occupations</i>						
Share of the labor market (empirical)	0.02					
Share of the labor market (generated)	0.02	0	0.02	0.02	0.03	-0.24
Degree, normal (empirical)	2.85	7.77	0	0	75	5.34
Degree, normal (generated)	2.92	0.69	3	0	5	-1.26
Degree, lockdown (empirical)	1.21	3.89	0	0	54	7.2
Degree, lockdown (generated)	1.24	0.99	1	0	5	0.44

Table D.3: Comparison of occupational groups in empirical and generated networks (2).

	Mean	SD	Median	Min	Max	Skew
<i>Construction and Extraction Occupations</i>						
Share of the labor market (empirical)	0.03					
Share of the labor market (generated)	0.03	0	0.03	0.03	0.04	0.12
Degree, normal (empirical)	3.6	18.53	0	0	400	15.28
Degree, normal (generated)	3.67	0.91	4	0	6	-1.07
Degree, lockdown (empirical)	1.99	11.58	0	0	300	19.52
Degree, lockdown (generated)	2.05	1.25	2	0	6	0.18
<i>Educational Instruction and Library Occupations</i>						
Share of the labor market (empirical)	0.04					
Share of the labor market (generated)	0.04	0	0.04	0.04	0.05	-0.47
Degree, normal (empirical)	12.76	62.44	0	0	1,100	12.45
Degree, normal (generated)	11.86	5.07	13	0	33	-0.31
Degree, lockdown (empirical)	2.61	18.87	0	0	500	20.5
Degree, lockdown (generated)	3.08	2.18	3	0	16	0.83
<i>Farming, Fishing, and Forestry Occupations</i>						
Share of the labor market (empirical)	0.01					
Share of the labor market (generated)	0	0	0	0	0.01	-0.36
Degree, normal (empirical)	2.15	4.98	0	0	36	3.75
Degree, normal (generated)	2.23	0.79	2	0	4	-0.45
Degree, lockdown (empirical)	1.76	4.98	0	0	50	5.72
Degree, lockdown (generated)	1.7	0.81	2	0	4	-0.07
<i>Food Preparation and Serving Related Occupations</i>						
Share of the labor market (empirical)	0.06					
Share of the labor market (generated)	0.06	0	0.06	0.06	0.07	-0.01
Degree, normal (empirical)	5.32	22.29	0	0	300	9.43
Degree, normal (generated)	5.32	1.43	6	0	9	-1.26
Degree, lockdown (empirical)	1.57	5.86	0	0	80	8.41
Degree, lockdown (generated)	1.62	1.32	1	0	8	0.78
<i>Healthcare Practitioners and Technical Occupations</i>						
Share of the labor market (empirical)	0.04					
Share of the labor market (generated)	0.04	0	0.04	0.04	0.05	-0.38
Degree, normal (empirical)	20.17	256.51	0	0	5,000	19.07
Degree, normal (generated)	15.5	9.02	14	0	97	0.93
Degree, lockdown (empirical)	5.19	23.75	0	0	440	11.49
Degree, lockdown (generated)	5.34	3.49	5	0	40	1.05
<i>Healthcare Support Occupations</i>						
Share of the labor market (empirical)	0.03					
Share of the labor market (generated)	0.03	0	0.03	0.03	0.04	0.03
Degree, normal (empirical)	5.44	19.44	0	0	200	6.62
Degree, normal (generated)	5.46	1.43	6	0	9	-1.34
Degree, lockdown (empirical)	3.7	12.65	0	0	100	5.61
Degree, lockdown (generated)	3.81	1.57	4	0	9	-0.11

Table D.4: Comparison of occupational groups in empirical and generated networks (3).

	Mean	SD	Median	Min	Max	Skew
<i>Installation, Maintenance, and Repair Occupations</i>						
Share of the labor market (empirical)	0.03					
Share of the labor market (generated)	0.03	0	0.03	0.02	0.03	0.17
Degree, normal (empirical)	3.76	12.36	0	0	200	8.04
Degree, normal (generated)	3.82	0.95	4	0	6	-1.15
Degree, lockdown (empirical)	2.83	11.24	0	0	200	10.1
Degree, lockdown (generated)	2.91	1.18	3	0	6	-0.3
<i>Legal Occupations</i>						
Share of the labor market (empirical)	0.01					
Share of the labor market (generated)	0.01	0	0.01	0	0.01	0.2
Degree, normal (empirical)	8.28	22.06	0	0	110	3.38
Degree, normal (generated)	8.13	2.22	9	0	14	-1.61
Degree, lockdown (empirical)	0.92	2.98	0	0	28	6.73
Degree, lockdown (generated)	1.84	1.67	1	0	10	1.1
<i>Life, Physical, and Social Science Occupations</i>						
Share of the labor market (empirical)	0.01					
Share of the labor market (generated)	0.01	0	0.01	0	0.01	0.21
Degree, normal (empirical)	4.63	14.67	0	0	200	6.4
Degree, normal (generated)	4.67	1.15	5	0	7	-1.06
Degree, lockdown (empirical)	3.55	18.64	0	0	300	13.06
Degree, lockdown (generated)	3.65	1.34	4	0	7	-0.27
<i>Management Occupations</i>						
Share of the labor market (empirical)	0.05					
Share of the labor market (generated)	0.05	0	0.05	0.04	0.05	-0.37
Degree, normal (empirical)	5.84	23.54	0	0	500	10.79
Degree, normal (generated)	5.83	1.54	6	0	10	-1.4
Degree, lockdown (empirical)	1.32	4.79	0	0	100	9.7
Degree, lockdown (generated)	1.56	1.34	1	0	8	0.89
<i>Office and Administrative Support Occupations</i>						
Share of the labor market (empirical)	0.09					
Share of the labor market (generated)	0.1	0	0.09	0.09	0.1	-0.29
Degree, normal (empirical)	4.42	16.51	0	0	300	9.91
Degree, normal (generated)	4.46	1.15	5	0	7	-1.15
Degree, lockdown (empirical)	1.94	7.66	0	0	100	9.04
Degree, lockdown (generated)	2	1.37	2	0	7	0.52
<i>Personal Care and Service Occupations</i>						
Share of the labor market (empirical)	0.02					
Share of the labor market (generated)	0.02	0	0.02	0.02	0.03	-0.06
Degree, normal (empirical)	12.82	79.3	0	0	1,000	10.31
Degree, normal (generated)	11.84	5.05	13	0	29	-0.28
Degree, lockdown (empirical)	4.58	41.84	0	0	891	17.9
Degree, lockdown (generated)	4.71	2.89	4	0	24	0.72

Table D.5: Comparison of occupational groups in empirical and generated networks (4).

	Mean	SD	Median	Min	Max	Skew
<i>Production Occupations</i>						
Share of the labor market (empirical)	0.04					
Share of the labor market (generated)	0.04	0	0.04	0.04	0.05	0.2
Degree, normal (empirical)	2.87	9.51	0	0	100	6.54
Degree, normal (generated)	2.91	0.67	3	0	5	-1.37
Degree, lockdown (empirical)	2.19	8.86	0	0	200	12.15
Degree, lockdown (generated)	2.25	0.93	2	0	5	-0.44
<i>Protective Service Occupations</i>						
Share of the labor market (empirical)	0.02					
Share of the labor market (generated)	0.02	0	0.02	0.01	0.02	0.3
Degree, normal (empirical)	1.12	3.3	0	0	20	3.83
Degree, normal (generated)	0.99	0.1	1	0	1	-10.33
Degree, lockdown (empirical)	1.05	2.37	0	0	10	2.66
Degree, lockdown (generated)	0.99	0.1	1	0	1	-10.33
<i>Sales and Related Occupations</i>						
Share of the labor market (empirical)	0.07					
Share of the labor market (generated)	0.07	0	0.07	0.06	0.08	-0.11
Degree, normal (empirical)	5.59	16.91	0	0	210	6.75
Degree, normal (generated)	5.58	1.45	6	0	9	-1.4
Degree, lockdown (empirical)	3.3	25.59	0	0	500	16.87
Degree, lockdown (generated)	3.4	1.71	3	0	9	0.07
<i>Transportation and Material Moving Occupations</i>						
Share of the labor market (empirical)	0.06					
Share of the labor market (generated)	0.06	0	0.06	0.06	0.06	-0.07
Degree, normal (empirical)	5.31	21.69	0	0	400	11.17
Degree, normal (generated)	5.3	1.43	6	0	9	-1.22
Degree, lockdown (empirical)	2.69	9.08	0	0	100	6.91
Degree, lockdown (generated)	2.77	1.65	3	0	9	0.36
<i>Unemployed</i>						
Share of the labor market (empirical)	0.05					
Share of the labor market (generated)	0.05	0	0.05	0.04	0.05	-0.09
Degree, normal (empirical)	2.34	12.45	NA	NA	NA	NA
Degree, normal (generated)	2.61	0.7	3	0	4	-1.35
Degree, lockdown (empirical)	0.96	5.08	NA	NA	NA	NA
Degree, lockdown (generated)	0.99	0.88	1	0	4	0.54
<i>Retired</i>						
Share of the labor market (empirical)	0.21					
Share of the labor market (generated)	0.21	0	0.21	0.2	0.22	0.11
Degree, normal (empirical)	2.13	12.56	NA	NA	NA	NA
Degree, normal (generated)	2.47	0.69	3	0	4	-0.94
Degree, lockdown (empirical)	0.87	5.12	NA	NA	NA	NA
Degree, lockdown (generated)	0.9	0.84	1	0	4	0.58

Table D.6: Final sizes of the simulated epidemics by conditions and controls.

	Baseline			Random			Targeted		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
<i>No lockdown</i>									
Overall	90.97	19.07	95.48	60.78	29.98	69.16	38.58	37.24	38.87
Vaccine availability 5%				85.24	21.21	91.18	79.7	25.94	89.35
Vaccine availability 10%				79.7	22.38	86.46	65.72	30.39	80.01
Vaccine availability 20%				68.47	23.48	76.07	43.26	31.74	57.91
Vaccine availability 30%				55.7	24.95	64.63	23.66	27.3	0.48
Vaccine availability 40%				43.85	24.21	52.41	12.31	20.07	0.07
Vaccine availability 50%				31.74	22.6	38.75	6.84	14.09	0.04
Vaccine effectivity 60%				67.22	27.11	74.44	50.65	34.58	60.76
Vaccine effectivity 75%				60.8	29.52	68.5	37.18	37.01	30.03
Vaccine effectivity 90%				54.33	31.73	62.27	27.92	36.52	0.13
Low av. degree (4.46-4.74)	84.91	24.05	90.16	53	31.96	62.65	28.18	34.29	0.26
Medium av. degree (4.74-5.02)	92.25	17.18	96.53	61.87	29.56	69.72	37.39	37.16	36.11
High av. degree (5.02-5.3)	93.16	16.82	95.69	64.14	28.4	72.71	45.73	37.4	54.46
Low clustering (0.21-0.33)	95.11	13.04	97.32	66.2	27.4	74.34	47.23	37.16	56.29
Medium clustering (0.33-0.44)	91.98	18.68	95.46	62.32	29.43	71.94	41.19	37.42	45.34
High clustering (0.44-0.56)	83.59	23.97	90	50.74	31.81	60.24	22.28	31.48	0.14
Low av. path length (6.38-7.58)	94.74	14.76	97.3	66.27	27.45	74.66	46.01	37.36	55.19
Medium av. path length (7.58-8.77)	89.79	19.99	95.17	58.95	30.5	67.21	37.22	36.95	35.92
High av. path length (8.77-9.97)	83.81	24.07	90.76	50.73	31.94	60.23	21.08	30.93	0.12
Low homophily (0.06-0.11)	89.16	20.99	95.04	58.38	30.86	67.65	33.11	36.19	5.38
Medium homophily (0.11-0.16)	95.16	15.3	97.6	67.99	26.22	75.82	49.32	36.94	59.42
High homophily (0.16-0.2)	92.9	16.21	95.49	63.04	28.89	72.41	44.73	37.44	52.78
<i>Lockdown</i>									
Overall	26.64	26.59	39.66	10.66	17.64	0.04	2.53	8.86	0.02
Vaccine availability 5%				22.69	23.92	0.2	10.63	16.81	0.06
Vaccine availability 10%				19.13	21.4	0.11	3.77	9.48	0.04
Vaccine availability 20%				11.73	16.14	0.05	0.62	3	0.02
Vaccine availability 30%				6.5	11.29	0.04	0.09	0.54	0.02
Vaccine availability 40%				2.87	6.91	0.03	0.05	0.18	0.02
Vaccine availability 50%				1.06	3.71	0.02	0.03	0.09	0.01
Vaccine effectivity 60%				12.67	18.79	0.05	3.87	11.01	0.03
Vaccine effectivity 75%				10.43	17.43	0.04	2.25	8.27	0.02
Vaccine effectivity 90%				8.9	16.4	0.03	1.48	6.55	0.02
Low av. degree (2.06-2.1)	25.6	26.11	0.35	9.79	16.93	0.04	2.2	8.2	0.02
Medium av. degree (2.1-2.13)	27.59	26.93	41.42	11.29	18.06	0.04	2.73	9.17	0.02
High av. degree (2.13-2.16)	26.05	27.45	0.18	12.24	19.16	0.05	3.56	11.1	0.02
Low clustering (0.06-0.1)	35.33	28.58	56.81	14.86	20.76	0.05	4.63	12.42	0.02
Medium clustering (0.1-0.14)	25.43	25.73	0.28	10.07	16.76	0.04	1.95	7.2	0.02
High clustering (0.14-0.18)	17.6	21.5	0.09	6.3	12.85	0.04	0.79	3.83	0.02
Low av. path length (11.28-13.62)	34.05	28.55	54.69	14.48	20.46	0.05	4.42	12.07	0.02
Medium av. path length (13.62-15.97)	23.28	24.69	0.18	8.83	15.66	0.04	1.52	6.19	0.02
High av. path length (15.97-18.31)	15.28	20.32	0.08	5.48	11.89	0.04	0.59	3.04	0.02
Low homophily (0.04-0.06)	23.62	25.61	0.15	9.59	16.73	0.04	2.15	8.01	0.02
Medium homophily (0.06-0.09)	33.57	28.76	56.94	14.68	20.59	0.05	4.75	12.54	0.02
High homophily (0.09-0.11)	29.06	26.6	48.48	10.87	17.59	0.04	2.27	8.25	0.02

Note: Final size is measured in percent of all nodes that got infected at any time per simulation run.

Table D.7: Peak sizes of the simulated epidemics by conditions and controls.

	Baseline			Random			Targeted		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
<i>No lockdown</i>									
Overall	66.91	16.1	69.84	39.14	22.77	42.08	17.99	20.69	9.15
Vaccine availability 5%				60.76	17.04	64.18	46.95	18.16	51.8
Vaccine availability 10%				54.93	17.25	58.84	28.98	16.94	32.99
Vaccine availability 20%				43.7	16.82	47.65	15.22	13.86	14.76
Vaccine availability 30%				32.72	16.36	36.8	7.7	10.39	0.21
Vaccine availability 40%				23.6	14.68	26.68	3.86	6.98	0.05
Vaccine availability 50%				15.53	12.55	16.16	2.11	4.67	0.03
Vaccine effectivity 60%				43.86	21.03	46.61	24.6	20.92	22.09
Vaccine effectivity 75%				38.81	22.42	40.91	16.76	20.23	6.41
Vaccine effectivity 90%				34.13	23.51	34.87	12.03	18.47	0.09
Low av. degree (4.46-4.74)	56.22	17.59	58.47	30.03	21.03	31.95	11.63	16.71	0.17
Medium av. degree (4.74-5.02)	68.25	14.18	72.04	39.69	22.39	43.05	17.34	20.41	7.75
High av. degree (5.02-5.3)	71.71	14.1	71.18	43.91	22.33	47.26	22.38	21.8	18.52
Low clustering (0.21-0.33)	76.64	11.06	78.54	47.31	22.73	51.31	24.7	23.33	19.77
Medium clustering (0.33-0.44)	66.8	13.94	69.19	38.93	21.38	42.23	18.01	19.55	10.75
High clustering (0.44-0.56)	53.87	16.08	59.23	27.5	19.77	29	8.14	13.19	0.1
Low av. path length (6.38-7.58)	75.82	12.24	77.84	46.71	22.6	50.62	23.56	23	18.67
Medium av. path length (7.58-8.77)	63.19	14.78	67.53	35.62	21.28	38.45	15.81	18.57	7.57
High av. path length (8.77-9.97)	53.19	16.02	59.48	26.79	19.57	28.12	7.36	12.41	0.09
Low homophily (0.06-0.11)	63.25	17.04	65.12	36.15	22.48	38.97	15.03	19.23	3.24
Medium homophily (0.11-0.16)	79.42	12.74	81.43	50.66	22.51	55.48	26.29	23.8	21.78
High homophily (0.16-0.2)	70.12	13.15	70.36	41.68	21.97	44.89	20.92	21.13	16.09
<i>Lockdown</i>									
Overall	8.84	9.35	8.68	3.09	5.54	0.04	0.53	1.93	0.02
Vaccine availability 5%				7.06	7.93	0.13	2.15	3.77	0.05
Vaccine availability 10%				5.58	6.7	0.08	0.66	1.74	0.03
Vaccine availability 20%				2.99	4.45	0.04	0.13	0.46	0.02
Vaccine availability 30%				1.47	2.73	0.03	0.04	0.09	0.02
Vaccine availability 40%				0.59	1.45	0.03	0.03	0.05	0.02
Vaccine availability 50%				0.22	0.7	0.02	0.02	0.04	0.01
Vaccine effectivity 60%				3.65	5.93	0.04	0.83	2.5	0.02
Vaccine effectivity 75%				2.98	5.41	0.04	0.45	1.7	0.02
Vaccine effectivity 90%				2.52	5.01	0.03	0.28	1.23	0.02
Low av. degree (2.06-2.1)	8.17	8.82	0.22	2.73	5.12	0.04	0.43	1.66	0.02
Medium av. degree (2.1-2.13)	9.38	9.63	9.65	3.21	5.63	0.04	0.56	1.98	0.02
High av. degree (2.13-2.16)	9.4	10.69	0.15	3.76	6.6	0.04	0.84	2.79	0.02
Low clustering (0.06-0.1)	13.47	11.09	19.59	4.82	7.22	0.04	1.03	2.89	0.02
Medium clustering (0.1-0.14)	7.8	7.97	0.16	2.67	4.63	0.04	0.34	1.22	0.02
High clustering (0.14-0.18)	4.63	5.72	0.07	1.43	2.99	0.03	0.15	0.59	0.02
Low av. path length (11.28-13.62)	13	11.05	18.66	4.58	7	0.04	0.96	2.75	0.02
Medium av. path length (13.62-15.97)	6.78	7.25	0.11	2.26	4.14	0.04	0.26	0.99	0.02
High av. path length (15.97-18.31)	3.6	4.82	0.06	1.12	2.48	0.03	0.12	0.44	0.02
Low homophily (0.04-0.06)	7.64	8.82	0.09	2.72	5.12	0.04	0.47	1.78	0.02
Medium homophily (0.06-0.09)	13.4	11.69	22.14	4.87	7.41	0.04	1.06	2.97	0.02
High homophily (0.09-0.11)	9.06	8.5	13.21	2.95	5.04	0.04	0.39	1.43	0.02

Note: Peak size is measured as the maximum percentage of simultaneously infected nodes per simulation run.

Table D.8: Mean peak size of baseline condition (in percent of infected nodes) and difference by test condition in percent points.

	Baseline	Random to Baseline	Targeted to Baseline	Targeted to Random
<i>No lockdown</i>				
Overall	66.91	−27.77	−48.92	−21.15
Vaccine availability 5%		−6.16	−19.97	−13.81
Vaccine availability 10%		−11.99	−37.94	−25.95
Vaccine availability 20%		−23.21	−51.69	−28.48
Vaccine availability 30%		−34.19	−59.21	−25.02
Vaccine availability 40%		−43.32	−63.05	−19.73
Vaccine availability 50%		−51.38	−64.81	−13.43
Vaccine effectivity 60%		−23.05	−42.31	−19.26
Vaccine effectivity 75%		−28.10	−50.15	−22.05
Vaccine effectivity 90%		−32.78	−54.89	−22.11
Low clustering	76.64	−29.33	−51.94	−22.61
Medium clustering	66.80	−27.87	−48.79	−20.92
High clustering	53.87	−26.37	−45.72	−19.36
Low homophily	63.25	−27.10	−48.22	−21.12
Medium homophily	79.42	−28.76	−53.13	−24.37
High homophily	70.12	−28.44	−49.20	−20.76
<i>Lockdown</i>				
Overall	8.84	−5.75	−8.31	−2.56
Vaccine availability 5%		−1.78	−6.70	−4.92
Vaccine availability 10%		−3.26	−8.18	−4.92
Vaccine availability 20%		−5.85	−8.72	−2.87
Vaccine availability 30%		−7.38	−8.81	−1.43
Vaccine availability 40%		−8.25	−8.81	−0.56
Vaccine availability 50%		−8.62	−8.82	−0.20
Vaccine effectivity 60%		−5.20	−8.02	−2.82
Vaccine effectivity 75%		−5.87	−8.39	−2.53
Vaccine effectivity 90%		−6.33	−8.57	−2.24
Low clustering	13.47	−8.65	−12.44	−3.79
Medium clustering	7.80	−5.13	−7.46	−2.32
High clustering	4.63	−3.20	−4.48	−1.29
Low homophily	7.64	−4.92	−7.17	−2.25
Medium homophily	13.40	−8.53	−12.34	−3.81
High homophily	9.06	−6.11	−8.67	−2.56

Table D.9: Duration of the simulated epidemics by conditions and controls in time steps.

	Baseline			Random			Targeted		
	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median
<i>No lockdown</i>									
Overall	61.47	13.44	62	68.27	27.21	72	76.81	60.26	76
Vaccine availability 5%				63.18	15.68	64	75.07	24.21	77
Vaccine availability 10%				64.91	17.49	67	95.64	46.55	97
Vaccine availability 20%				68.58	21.53	72	101.85	68.39	106
Vaccine availability 30%				70.19	27.08	76	83.43	73.09	59
Vaccine availability 40%				71.81	32.63	80	59.01	61.46	25
Vaccine availability 50%				70.93	39.61	82	45.85	54.19	19
Vaccine effectivity 60%				68.02	23.47	71	82.98	46.59	89
Vaccine effectivity 75%				68.61	26.87	72	79.28	61.14	77
Vaccine effectivity 90%				68.16	30.8	72	68.16	69.76	33
Low av. degree (4.46-4.74)	64.62	16.03	67	70.03	32.45	76	69.77	62.9	41
Medium av. degree (4.74-5.02)	61.74	12.72	62	69.33	26.94	72	76.44	61.19	75
High av. degree (5.02-5.3)	59.4	12.1	60	66.19	23.86	69	81.2	57.32	81
Low clustering (0.21-0.33)	55.57	8.6	56	62.58	20.99	65	76.95	53.31	77
Medium clustering (0.33-0.44)	62.89	12.69	64	71.04	26.63	75	83.57	62.62	84
High clustering (0.44-0.56)	67.07	16.69	69	71.22	33.72	78	64.94	63.14	30
Low av. path length (6.38-7.58)	55.96	9.3	56	63.86	21.58	66	77.59	54.81	77
Medium av. path length (7.58-8.77)	64.03	13.71	65	70.74	28.49	75	79.65	62.92	80
High av. path length (8.77-9.97)	69.05	16.34	71.5	72.85	34.86	80	64.53	64.68	28
Low homophily (0.06-0.11)	62.94	14.47	64	69.92	29.35	74	73.56	62.2	67
Medium homophily (0.11-0.16)	53.36	9.07	53	59.95	18.47	62	75.7	49.6	76
High homophily (0.16-0.2)	60.75	11.86	61	67.3	24.82	70	81.79	58.73	82
<i>Lockdown</i>									
Overall	72	57.25	91.5	57.43	63.59	19	33.73	54.58	14
Vaccine availability 5%				72.29	61.08	39	73.85	85.93	22
Vaccine availability 10%				71.82	63.82	30	51.92	73.55	19
Vaccine availability 20%				66.53	69.17	22	28.01	40.84	15
Vaccine availability 30%				57.56	68.47	19	18.08	16.88	12
Vaccine availability 40%				44.62	60.36	16	16.18	12.34	11
Vaccine availability 50%				31.79	44.85	14	14.36	8.64	10
Vaccine effectivity 60%				62.7	65.31	21	40.76	60.53	16
Vaccine effectivity 75%				57.22	63.3	19	32.35	52.51	14
Vaccine effectivity 90%				52.38	61.69	18	28.09	49.33	12
Low av. degree (2.06-2.1)	72.86	58.6	55.5	55.55	62.89	19	32.73	54.07	14
Medium av. degree (2.1-2.13)	71.79	56.27	94	58.98	64.22	20	34.47	55.16	14
High av. degree (2.13-2.16)	65.7	55.01	36.5	58.46	62.83	20	35.3	52.78	14
Low clustering (0.06-0.1)	74.33	50.61	99	61.86	62.38	21	40.62	62.4	14
Medium clustering (0.1-0.14)	73.1	58.59	46.5	58.13	64.6	19	33.07	54.87	14
High clustering (0.14-0.18)	67.43	62.65	26	50.85	63.03	18	26.13	40.87	13
Low av. path length (11.28-13.62)	72.41	50.27	99	61.56	62.53	21	39.7	60.86	14
Medium av. path length (13.62-15.97)	72.7	60.21	38	55.81	64.05	19	30.94	51.77	14
High av. path length (15.97-18.31)	66.05	66.29	22	49.66	64.09	18	25.11	37.66	13
Low homophily (0.04-0.06)	68.68	58.08	34	55.88	64	19	31.4	49.59	13
Medium homophily (0.06-0.09)	69.83	48.59	96	60.11	60.03	21	40.57	61.97	14
High homophily (0.09-0.11)	78.86	58.67	109	59.06	64.25	20	34.98	59.23	14

Note: Duration is measured as the number of time steps until all once infected nodes have recovered.

Table D.10: Mean duration of baseline condition and difference by test condition in percent points.

	Baseline	Random to Baseline	Targeted to Baseline	Targeted to Random
<i>No lockdown</i>				
Overall	61.47	6.80	15.34	8.54
Vaccine availability 5%		1.71	13.60	11.89
Vaccine availability 10%		3.44	34.17	30.73
Vaccine availability 20%		7.11	40.38	33.27
Vaccine availability 30%		8.72	21.96	13.24
Vaccine availability 40%		10.34	−2.46	−12.80
Vaccine availability 50%		9.46	−15.62	−25.08
Vaccine effectivity 60%		6.55	21.51	14.96
Vaccine effectivity 75%		7.14	17.81	10.67
Vaccine effectivity 90%		6.69	6.69	0.00
Low clustering	55.57	7.01	21.38	14.36
Medium clustering	62.89	8.15	20.68	12.53
High clustering	67.07	4.16	−2.13	−6.29
Low homophily	62.94	6.99	10.62	3.63
Medium homophily	53.36	6.59	22.34	15.75
High homophily	60.75	6.55	21.04	14.49
<i>Lockdown</i>				
Overall	72.00	−14.56	−38.26	−23.70
Vaccine availability 5%		0.29	1.85	1.57
Vaccine availability 10%		−0.18	−20.07	−19.89
Vaccine availability 20%		−5.47	−43.99	−38.52
Vaccine availability 30%		−14.43	−53.92	−39.48
Vaccine availability 40%		−27.38	−55.81	−28.44
Vaccine availability 50%		−40.21	−57.64	−17.43
Vaccine effectivity 60%		−9.30	−31.24	−21.94
Vaccine effectivity 75%		−14.78	−39.65	−24.87
Vaccine effectivity 90%		−19.61	−43.90	−24.29
Low clustering	74.33	−12.47	−33.72	−21.24
Medium clustering	73.10	−14.97	−40.03	−25.06
High clustering	67.43	−16.57	−41.30	−24.72
Low homophily	68.68	−12.79	−37.28	−24.48
Medium homophily	69.83	−9.72	−29.26	−19.54
High homophily	78.86	−19.80	−43.88	−24.08

D.2 Supplementary methods

Note that the following elaborations are partly taken from another manuscript (Nunner et al., 2022a) that presents the model used for network generation and simulation of epidemics. It is a specific model case of the *Networking during Infectious Disease Model (NIDM)* (Nunner et al., 2021), designed to simulate epidemics in small-world networks with control over degree per node and network clustering. The information provided is complete in terms of components used for the current study.

D.2.1 Network generation

Network generation is based on utility maximization for network positions. That is, utility for an agent i is a trade-off between benefits (B_i) and maintenance costs (C_i) a social tie creates:

$$U_i(\mathbf{G}) = B_i(\mathbf{G}) - C_i(\mathbf{G}). \quad (\text{D.1})$$

We consider networks to be undirected, unweighted, and non-reflexive graphs:

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1N} \\ g_{21} & g_{22} & \cdots & g_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ g_{N1} & g_{N2} & \cdots & g_{NN} \end{bmatrix}, \quad (\text{D.2})$$

with $g_{ij} \in \{0, 1\}$, $g_{ii} = 0$, and $g_{ij} = g_{ji} = 1$ if a tie exists. Social benefit is defined as the sum of weighted benefits for direct ties and weighted benefits for the proportion of closed triads of agent i :

$$B_i(\mathbf{G}) = b_1 \cdot t_i + b_2 \cdot \left(1 - 2 \cdot \frac{|x_i - \alpha|}{\max(\alpha, 1 - \alpha)} \right), \quad (\text{D.3})$$

with x_i denoting the actual proportion of closed triads i belongs to in i 's ego-network, α the preferred proportion of closed triads of agent i , and t_i the number of ties agent i possesses:

$$t_i = \sum_j g_{ij}. \quad (\text{D.4})$$

Social maintenance costs are assumed to be quadratic in the number of ties t_i to model increasing marginal costs of additional ties:

$$C_i(\mathbf{G}) = c_1 \cdot t_i + c_2 \cdot t_i^2. \quad (\text{D.5})$$

Network generation for reported contact numbers prior to lockdown was realized using the NIDM network formation procedure. That is, time is modeled as discrete time steps. Within each time step, we simulate the formation and dissolution of ties:

- Create an empty set of unprocessed agents A^1 and repeat until all agents have been processed:
 - Randomly select an unprocessed agent i , either:
 - from A if $A \neq \{\}$ and not all agents from A have been processed yet, or
 - from the entire population N if $A = \{\}$.
 - Repeat until agent i has evaluated $\phi \cdot (N - 1)$ ties other agents j :
 - Create a new set J of agents to be evaluated through one of the three following options:
 1. With probability ψ : J becomes all agents at distance 1 that have not yet been evaluated by i in the current time step.
 2. With probability ξ : J becomes all agents at distance 2 that have not yet been evaluated by i in the current time step.
 3. With probability $1 - \psi - \xi$: J becomes all agents from the entire population that have not yet been evaluated by i in the current time step.
 - Select a single agent j from J either:
 - Belonging to the same occupational group (with probability ω), or
 - a randomly selected agent (with probability $1 - \omega$).
 - If j is directly tied to i :
 - Dissolve tie ij , if i 's utility excluding ij exceeds the utility including ij .
 - else:
 - Form tie ij , if both i 's and j 's utilities including ij exceed the utilities excluding ij .
 - Add j to A if j is directly tied to i .

Generation of lockdown networks was realized using a network pruning algorithm:

- For all networks N :
 - Repeat until N is “in lockdown” (av. degree of all occupational groups \leq reported av. lockdown degree + 3%):
 - For all edges e connecting nodes n_1 and n_2 in N (in random order):
 - If n_1 and n_2 are not yet “in lockdown” (degree $>$ reported av. lockdown degree of occupational group + 3%):
 - Remove e with a probability depending on the reported percentage decrease in degree between prior and during lockdown (t^-): $\frac{t_{n_1}^- + t_{n_2}^-}{2}$

¹The use of A (in combination with parameters ψ and ξ that are introduced later) and thus the consecutive processing of an agent's neighbors allows producing high levels of clustering and homophilous mixing in networks too large for a purely randomized approach.

D.2.2 Disease spread

Disease states (susceptible, infected, recovered, or vaccinated) are defined for each node in the network:

$$\mathbf{d} \in \{S, I, R, V\}^N. \quad (\text{D.6})$$

The probability for a node i to get infected per time step

$$\pi_i(\mathbf{G}, \mathbf{d}) = 1 - (1 - \gamma)^{t_{iI}}, \quad (\text{D.7})$$

depends on the probability to get infected per single contact (γ) and the number of infected neighbors:

$$t_{iI} = \sum_{j, d_j=I} g_{ij}. \quad (\text{D.8})$$

As before, time is modeled as discrete time steps. Within each time step, we simulate disease transmission events:

- Repeat until all nodes have been processed:
 - Randomly select an unprocessed node i :
 - If i is infected, compute whether node recovers: passed time steps since infection $\geq \tau$.
 - If i is susceptible, compute whether i gets infected from infected neighbors (see Equation D.7).

D.2.3 Parameters and submodels

Table D.11 presents all parameters including the range of possible settings and the initial settings used for network generation and the simulation of epidemics. We systematically varied the settings for α (to control clustering) and ω (to control occupational group homophily) and selected for each combination the 10 best fits, resulting in 90 normal (prior to lockdown) networks. Afterwards, we used the pruning algorithm described earlier to generate 90 lockdown networks.

Based on these 180 networks, we simulated 20 epidemics for each of the conditions (Baseline, Random, Targeted). For the two vaccination campaign conditions, we systematically varied the vaccine parameters (availability, effectivity) to end up with $1 \cdot 20 \cdot 180 + 2 \cdot 20 \cdot 180 \cdot 6 \cdot 3 = 133,200$ simulated epidemics.

Network size was selected to ensure large enough networks to make meaningful inferences, while limiting computational demands to allow a large number of network variations and simulated epidemics. Other fixed parameters were selected because of one of two reasons. First, they are backed by empirical data. Recent US labor market numbers were taken to set occupational group size (U.S. Bureau of Labor Statistics, 2019). Clustering in contact networks was found to be around 0.46 with clustering getting lower for older

persons (Danon et al., 2013). We, therefore, varied α between 0.3, 0.4, and 0.5. Average degree per contact in our model can be realized in the model by keeping the b_1 and c_1 constant, and setting c_2 to define the optimum number of ties: $\frac{b_1 - c_1}{2 \cdot t_{\text{target}}}$. This is done for each node individually, dependent on reported contact numbers per occupational group. Other fixed parameters were selected because pilot runs showed that they produce informative and interesting variations of epidemics regarding final size, epidemic peak size, and epidemic peak time.

Table D.11: Parameters, ranges, and initial settings.

Parameter	Range	Initial setting
<i>Network Generation</i>		
<i>I. Utility</i>		
Benefit per social tie	$b_1 \in \mathbb{R}_0^+$	$b_1 = 1.0$
Benefit for triadic closure	$b_2 \in \mathbb{R}_0^+$	$b_2 = 0.5$
Preferred proportion of closed triads	$0 \leq \alpha \leq 1$	$\alpha \in \{0.3, 0.4, 0.5\}$
Simple cost per tie	$c_1 \in \mathbb{R}_0^+$	$c_1 = 0.2$
Marginal cost per tie	$c_2 \in \mathbb{R}_0^+$	$c_2 = [0.02, 0.36]^*$
<i>II. Network</i>		
Number of agents	$N \in \mathbb{Z}_0^+$	$N = 10,000$
<i>III. Simulation</i>		
Proportion of nodes offered to evaluate per time step	$0 \leq \phi \leq 1$	$\phi = 0.001$
Proportion of ϕ as distance 1 ties	$0 \leq \psi \leq 1$	$\psi = 0.3$
Proportion of ϕ as distance 2 ties	$0 \leq \xi \leq 1$	$\xi = 0.5$
Likelihood of offered alter from the same occ. group	$0 \leq \omega \leq 1$	$\omega \in \{0.0, 0.4, 0.8\}$
<i>Simulation of Epidemics</i>		
<i>I. Disease</i>		
Probability of disease transmission per time step	$0 \leq \gamma \leq 1$	$\gamma = 0.15$
Recovery time in time steps	$\tau \in \mathbb{Z}^+$	$\tau = 10$
<i>II. Vaccination</i>		
Vaccine availability	$0 \leq \theta \leq 1$	$\theta \in \{0.05, 0.1, 0.2, 0.3, 0.4, 0.5\}$
Vaccine effectivity	$0 \leq \eta \leq 1$	$\eta \in \{0.6, 0.75, 0.9\}$

Note: *Settings for c_2 depended on occupational group membership and reported contact numbers. Healthcare Practitioners and Technical Occupations, for example, had the highest reported contact numbers (20.17), thus receiving the lowest marginal costs per tie: $c_2 = \frac{1.0 - 0.2}{2 \cdot 20.17} = 0.02$. Protective Service Occupations, on the other hand, had the lowest reported contact numbers (1.12), thus receiving the highest marginal costs per tie: $c_2 = \frac{1.0 - 0.2}{2 \cdot 1.12} = 0.36$.

Nederlandse samenvatting

Achtergrond

Sinds het begin van de COVID-19 pandemie begin 2020 tot de voltooiing van dit proefschrift in juli 2022 heeft de Wereldgezondheidsorganisatie (WHO) bijna 560 miljoen gevallen en meer dan 6 miljoen sterfgevallen van COVID-19 geregistreerd (World Health Organization, 2022b). Hoewel er een aantal effectieve vaccins zijn ontwikkeld en er wereldwijd grootschalige vaccinatiecampagnes zijn gevoerd, hebben maar 58 van de 194 lidstaten van de WHO de doelstelling van 70 procent vaccinatiegraad gehaald (World Health Organization, 2022a). Inzicht in de dynamiek van besmettelijke ziektes in haar sociale context blijft noodzakelijk om de negatieve medische en sociale gevolgen van dergelijke epidemieën te verzachten. Wiskundige modellen van ziekteverspreiding zijn een van de belangrijkste instrumenten om de verspreiding van een dergelijke dynamiek te begrijpen en kunnen zo de beleidsvorming voor de volksgezondheid ondersteunen.

Klassieke deterministische modellen in de epidemiologie zijn bijvoorbeeld belangrijke hulpmiddelen voor het voorspellen van het verloop en de prevalentie van ziekten in grote gemeenschappen. Dit is onder meer belangrijk voor het schatten van ziekenhuisopnames en het bepalen van hoeveel vaccinaties nodig zijn om het aantal gevallen voor een komende COVID-19 golf onder controle te houden. Deze modellen maken vaak gebruik van gewone differentiaalvergelijkingen die veronderstellen dat de populatie van potentiële gastheren perfect gemengd is. We weten echter dat mensen zich niet homogeen vermengen, maar sociale banden vormen op basis van persoonlijke voorkeuren en de mogelijkheden die hun omgeving biedt (Feld, 1981; McPherson et al., 2001). Netwerkmodellen van ziekteverspreiding, het type model dat in dit proefschrift wordt gebruikt, versoepelen de aanname van perfect gemengde populaties en stellen ons in staat de effecten van sociale structuur en interactiepatronen op de verspreiding van infectieziekten te begrijpen. Dit is belangrijk voor ziekten die zich voornamelijk via interpersoonlijke contacten verspreiden, zoals COVID-19 (Li et al., 2020; Shen et al., 2020). Eerdere simulatiestudies hebben bijvoorbeeld herhaaldelijk gesuggereerd dat de aanwezigheid van clusters binnen een sociaal netwerk de verspreiding van ziekten kan afremmen (Badham & Stocker, 2010; Keeling, 1999; Miller, 2009). Clusters beschrijven dicht verbonden gebieden binnen een netwerk waarin slechts enkele relaties bruggen vormen naar andere gebieden. Voor een besmettelijke ziekte vormen deze bruggen een knelpunt voor de verspreiding van het ene cluster naar het andere.

Maar zoals elk wetenschappelijk model maken netwerkmodellen vereenvoudigende veronderstellingen die de interpretatie van hun theoretische resultaten in de echte wereld bemoeilijken. Een typische veronderstelling is dat menselijk gedrag (contactfrequentie, gezondheidsgerelateerd gedrag) constant is, wat leidt tot statische netwerkstructuren (Bansal et al., 2010). Het verwaarlozen van netwerkveranderingen kan een goede benadering zijn, vooral wanneer dergelijke veranderingen zich voordoen op een veel ruimere tijdschaal dan de verspreiding van infecties (Pastor-Satorras et al., 2015). Voorspellingen over de dynamiek van ziekteverspreiding zouden echter wezenlijk anders kunnen zijn als rekening wordt gehouden met het co-evolutionaire proces tussen sociale netwerken en de verspreiding van infecties (Bansal et al., 2010; Marceau et al., 2010). Een benadering die met dergelijke processen

rekening houdt, is het ontwerpen van op het individu gebaseerde modellen voor de overdracht van infectieziekten (Bedson et al., 2021; Verelst et al., 2016; Willem et al., 2017). Uit simulaties op basis van willekeurige netwerkveranderingen op individueel niveau blijkt bijvoorbeeld dat rekening houden met netwerkdynamiek kan leiden tot kleinere en kortere epidemieën (Gross et al., 2006; Leung et al., 2018; Risau-Gusman & Zanette, 2009).

Gezondheidsgerelateerd gedrag en de dynamiek van sociale netwerken

Wij zien netwerkveranderingen niet als willekeurige veranderingen, maar als bewuste beslissingen die verband houden met de reacties van mensen op en hun bezorgdheid over ziekte. Om te illustreren hoe deze reacties op gezondheid de structuur van onze sociale netwerken systematisch kunnen veranderen, stelt u zich voor dat het begin 2020 is en de eerste gevallen van COVID-19 in uw land zijn gemeld. U bent bang om een ziekte op te lopen die nog niet goed wordt begrepen en u weet niet hoe u zich het best kunt beschermen. Nog voordat officiële afstandsmaatregelen zijn aangekondigd, besluit u een skireis naar Zuid-Tirol af te zeggen en vermijdt u sociale contacten die niet absoluut noodzakelijk zijn. Enkele maanden later, nadat veel mensen aan COVID-19-infecties zijn overleden, is de eerste golf van de pandemie afgezwakt door een landelijke lockdown. Het is op dat moment dat u en een vriend een bezoek plannen omdat geen van u beiden symptomen van COVID-19-infectie ondervindt en niemand in uw huishouden de afgelopen tijd iemand anders heeft gezien.

Dit voorbeeld beschrijft een netwerkmechanisme dat de basis vormt voor afstandsmaatregelen. Dat wil zeggen, een verandering in de netwerkstructuur (hier: vermindering van het aantal sociale contacten) beïnvloedt de verspreiding van de ziekte (hier: vermindering van het aantal nieuwe gevallen). Uit het voorbeeld blijkt verder dat er verschillende redenen zijn voor verandering in de netwerkstructuur. Ten eerste kunnen mensen sterven en al hun sociale relaties definitief verdwijnen. Ten tweede kunnen autoriteiten tijdelijke afstandsmaatregelen opleggen. Ten derde kunnen individuen uit eigen beweging besluiten hun interpersoonlijke contacten tijdelijk te verminderen. In dit proefschrift zijn wij geïnteresseerd in netwerkveranderingen op basis van deze laatste reden, die een wisselwerking vormt tussen (i.) de verspreiding van infectieziekten in sociale netwerken, (ii.) de aanpassing van het netwerkgedrag van mensen aan de verspreiding van ziekten, en (iii.) de invloed van structurele netwerkveranderingen op de verspreiding van infecties. Daarom is de hoofdvraag van dit proefschrift: Hoe bepaalt gezondheidsgerelateerd gedrag de co-evolutie van sociale netwerken en infectieziekten?

Een model voor de co-evolutie van sociale netwerken en infectieziekten

Om deze vraag te beantwoorden hebben we het *Networking during Infectious Diseases Model* (NIDM) gecreëerd. Het NIDM is een agentgebaseerd simulatiemodel dat theorie uit drie verschillende gebieden integreert. Vanuit de sociologie weten we dat sociale relaties bijdragen aan iemands tevredenheid (Ormel et al., 1999). Bovendien kunnen sociale relaties strategisch worden gebruikt om normen af te dwingen (Coleman, 1994) of de informatiestroom tussen mensen te controleren (Burt, 1992). Tegelijkertijd zijn er grenzen aan het aantal relaties dat iemand actief kan onderhouden, omdat ze cognitief veeleisend zijn (Dunbar, 1993; Lindenfors et al., 2021) en tijd, moeite en mogelijk geld kosten (Jackson, 2008). Uit de epidemiologie gebruiken we het concept van *compartmental models* (Hens et al., 2012). Deze modellen verdelen een populatie in verschillende compartimenten, afhankelijk van de ziekte-toestand die individuen ervaren, bijvoorbeeld vatbaar, geïnfecteerd en hersteld. Voor sociale netwerken betekent dit dat een vatbaar individu een infectie kan oplopen van iemand met wie hij een sociale relatie deelt. Eenmaal geïnfecteerd duurt het, afhankelijk van de ziekte, enige tijd voordat het individu is hersteld. Uit de gezondheidspsychologie weten we dat individuen de neiging hebben om openbare plaatsen en besmettelijke anderen te vermijden in tijden van verhoogd infectierisico (Ahmed et al., 2018; Bish & Michie, 2010; Jones & Salathé, 2009; Tracy et al., 2009). Bovendien weten we dat personen die infectie als waarschijnlijker en ernstiger beschouwen, eerder geneigd zijn tot dit vermijdingsgedrag (Bish & Michie, 2010; Ferguson, 2007; Goodwin et al., 2009; Jones & Salathé, 2009; Leppin & Aro, 2009).

Op basis van deze kennis beschrijft het NIDM netwerkbeslissingen als een afweging tussen de voordelen, kosten, en mogelijke schade van infecties door een sociale relatie. Dat wil zeggen, het NIDM beloont gunstige netwerkposities (bijvoorbeeld het aantal sociale relaties, deel uitmaken van een cluster) en brengt kosten met zich mee voor het oplopen van een infectie. De mate waarin agenten besmettelijke anderen vermijden hangt af van individuele risicopercepties over vatbaarheid voor en ernst van een besmettelijke ziekte. Verder nemen we aan dat agenten de verwachte beloning op korte termijn maximaliseren. Dat betekent dat agenten alleen relaties aangaan of onderhouden als het directe voordeel van die relatie opweegt tegen de waargenomen potentiële kosten.

Overzicht en discussie van de resultaten

Homofilie van gezondheidsgerelateerd gedrag beperkt de verspreiding van infectieziekten

Om een volledig beeld te krijgen van hoe de co-evolutie van sociale netwerken en infectieziekten wordt gevormd door het NIDM, zijn we begonnen met twee studies waarbij agentgebaseerde simulaties zijn gebruikt. Deze simulaties suggereren dat hogere niveaus

van waargenomen gezondheidsrisico's sterkere ziektevermijdende reacties uitlokken, wat leidt tot een lager aantal sociale relaties en dus minder agenten die infecties oplopen. Bovendien bleek dat de dynamiek zeer gevoelig is voor veranderingen in gedrag en dus netwerkstructuur, omdat kleine variaties in ziektevermijding het verschil kunnen maken of een ziekte-uitbraak een epidemie wordt. Een andere bevinding is dat dicht verbonden gebieden van agenten die een hoog infectierisico waarnemen, elkaar kunnen beschermen tegen infectie. Ter illustratie: denk aan een familie van drie generaties (grootouders, ouders, kinderen) die in hetzelfde huis wonen. Iedereen heeft een nauwe band met alle anderen. Bovendien ervaart iedereen de kans op en het risico van besmetting met COVID-19 als vergelijkbaar. In geval van een nieuwe golf van COVID-19 besluit de familie zich te isoleren door binnenshuis te blijven en hun sociale contacten te minimaliseren. Op grotere schaal isoleert dergelijk gedrag niet alleen individuen, maar hele netwerkgebieden en beschermt hen zo tegen besmetting. Dit gecombineerde effect van clustering in sociale netwerken en homofilie op het gezondheidsgerelateerd gedrag (de neiging om relaties aan te knopen met anderen die dezelfde risicoperceptie over het infectierisico delen) op de verspreiding van de ziekte leidt tot de laagste waargenomen infectiepercentages in al onze simulaties.

We menen dat dit inzicht een interessante aanvulling is op de baanbrekende studie van Watts and Strogatz (1998), die stelt dat bruggen tussen clusters sluiproutes zijn waarlangs besmettelijke ziekten snel elk deel van een netwerk kunnen bereiken. Hoewel we het eens zijn met de algemene notie van Watts and Strogatz (1998), zien we ook dat deze sluiproutes zwakke punten zijn voor de verspreiding van ziekten en daarom kunnen worden gebruikt om afstandsmaatregelen te verbeteren. Dat wil zeggen, aan de ene kant leidt vermindering van het contact met verre relaties (ruimtelijke afstand, emotionele afstand) tot minder verstoring van het netwerk. Aan de andere kant vermindert een beperking van het contact met verre relaties de kans op een snelle verspreiding van de ziekte door het netwerk. Wij denken bovendien dat dit effect kan worden benut voor epidemische interventie, zoals gerichte voorlichtingscampagnes voor groepen die weinig gezondheidsrisico's ervaren. De literatuur geeft verschillende voorbeelden van hoe risicoperceptie samenvalt met persoonlijke kenmerken die homofilie op groepsniveau bevorderen, zoals geslacht, leeftijd, opleidingsniveau, enzovoort. (Bish & Michie, 2010).

Vermijding van infectieziekten kan ten koste gaan van de sociale cohesie

Uit de simulatieresultaten konden we hypothesen afleiden over de gedragsaannames en gevolgen binnen het NIDM. Specifiek testen we of besmettelijke anderen vermeden worden door vatbare ego's (**H4.1**), of een hoger waargenomen infectierisico een sterkere vermijding van besmettelijke anderen veroorzaakt (**H4.2**), of een hoger waargenomen infectierisico de individuele kans om besmet te raken verlaagt (**H4.3**), of epidemieën kleiner zijn in netwerken met een hogere graad van clustering (**H4.4**), en of epidemieën kleiner zijn in netwerken met een hogere graad van homofilie in gezondheidsgerelateerd gedrag (**H4.5**). Voor dit doel ontwikkelden we de *Networking during Infectious Diseases Task* (NIDT),

een netwerkspel dat punten uitkeert op basis van dezelfde beloningen gedefinieerd in het NIDM. De punten werden omgezet in geldbedragen en aan het eind van het experiment aan de deelnemers uitbetaald.

Onze resultaten ondersteunen de gedragsaannames van het NIDM. Deelnemers vermijden actief geïnfecteerde anderen (**H4.1**) en meer risicomijdende deelnemers vertonen sterkere vermijdingsreacties (**H4.2**). Het vermijdingsgedrag was echter zo sterk dat de ziekte zich nauwelijks binnen de netwerken verspreidde. Daardoor konden we de overige hypothesen (**H4.3-H4.5**) niet bevestigen noch verwerpen. Aanvullende simulaties suggereerden dat sterkere ziektevermijding alleen de experimentele resultaten niet kan verklaren. Deelnemers gebruikten veeleer meer verfijnde strategieën om infectie en besmettelijke anderen te vermijden dan verwacht. Dat wil zeggen, besmettelijke deelnemers werden niet alleen vermeden door vatbare deelnemers, maar door alle deelnemers, ongeacht hun ziekte-toestand. Bovendien werden niet alleen besmettelijke deelnemers vermeden, maar ook vatbare deelnemers met besmettelijke burens. Samenvattend kunnen we stellen dat deelnemers het vermijden van ziekte prioriteit geven boven het handhaven van de netwerkstructuur wat ten koste ging van hun opbrengsten uit de netwerkstructuur.

We identificeren drie redenen voor het verschil tussen de modelvoorspellingen en de experimentele resultaten. Ten eerste bieden infecties een opvallend signaal (rode kleur voor geïnfecteerde deelnemers) dat onmiddellijke gedragsreacties uitlokt (vermijding). In vergelijking daarmee vergt clustering meer inspanning om bij te houden (bijhouden van het aantal gemeenschappelijke burens in iemands netwerk). Ten tweede, volgens *Prospect Theory* (Kahneman & Tversky, 2013; Tversky & Kahneman, 1992) hebben mensen de neiging om beschikbare voordelen af te wijzen om lage verlieskansen te vermijden (Denes-Raj & Epstein, 1994). Deze neiging kan hebben geleid tot de sterke ziektevermijdende reactie ten koste van netwerkvoordelen. Ten derde kunnen de media, culturele groepen en interpersoonlijke netwerken het risico van de COVID-19 pandemie, die aan de gang was op het ogenblik dat we het experiment uitvoerden, hebben versterkt (Kasperson et al., 1988) en zo reacties hebben uitgelokt die verder gingen dan de onmiddellijke beperking van de schade.

Ondanks de moeilijkheid om de resultaten van ons experiment te generaliseren naar de echte wereld, geloven we dat deze prioritering van ziektevermijding boven het behoud van sociale cohesie kan worden beschouwd als een waarschuwingssignaal. Dat wil zeggen, infectie kan worden beschouwd als een opvallend signaal (bijvoorbeeld zichtbare symptomen, zelfrapportages) met duidelijke gevolgen (het ervaren van symptomen, absentisme) die onmiddellijke gedragsreacties uitlokken (het vermijden van besmettelijke anderen). Verlies van sociale cohesie is echter betrekkelijk moeilijk vast te stellen omdat het een gevolg is van ziektevermijding dat pas na verloop van tijd optreedt. Verder kan een persoon volgens de *social production function* (SPF) theorie (Ormel et al., 1999) een laag niveau van sociaal welzijn niet compenseren met een hoog niveau van lichamelijk welzijn, omdat beide universele, niet-substitueerbare doelen onafhankelijk van elkaar moeten worden nagestreefd (Ormel et al., 1999). Studies tijdens de COVID-19 pandemie hebben aangetoond dat sociaal isolement een hele reeks persoonlijke problemen kan veroorzaken, zoals depressie, angst,

slaapgebrek, drugsmisbruik, enzovoort. (Banerjee & Rai, 2020; Heape, 2021; Kim & Jung, 2021; Pietrabissa & Simpson, 2020; Sepúlveda-Loyola et al., 2020). De prioritering van ziektevermijding kan dus isolement versterken door opgelegde afstandsmaatregelen en zo op de lange termijn een langzaam en moeilijk te herkennen verval van sociale cohesie en algemeen welzijn in de hand werken.

Prioriteit geven aan contactberoepen verhoogt effectiviteit van vaccinatiecampagnes

In het laatste deel van dit proefschrift wisselen we het perspectief van hoe individueel netwerkgedrag de verspreiding van besmettelijke ziekten beïnvloedt naar hoe we netwerkeigenschappen kunnen gebruiken voor epidemische interventie. Specifiek gaan we in op een twintig jaar oud idee dat het prioriteren van personen met een veel contacten de effectiviteit van vaccinatiecampagnes kan verbeteren. Een open vraag was echter hoe personen met veel contacten te identificeren, omdat contactinformatie niet alleen moeilijk te achterhalen, maar ook privé is. In plaats van personen met veel contacten te identificeren, stellen we voor de aansluiting bij een beroepsgroep te gebruiken als benadering voor verbondenheid in een sociaal netwerk.

Om te testen of beroepsgroepen als een geschikte benadering kunnen fungeren, hebben we eerst SWIDM-netwerken toegepast op contactgegevens die zijn verzameld tijdens de vroege stadia van de COVID-19-pandemie in China, Zuid-Korea, Japan, Italië, Groot-Britannië en de Verenigde Staten (Belot et al., 2020). Vervolgens hebben we epidemieën gesimuleerd die variëren in manier van verspreiden van vaccinaties (basisscenario/geen vaccinatie, willekeurige distributie, prioritering van beroepen met veel contacten), beschikbaarheid van vaccins (5-50% van de bevolking) en waarschijnlijkheid van immunisatie (60-90%). Onze resultaten wijzen erop dat, in vergelijking met het basisscenario, het aantal besmettingen in vergelijkbare mate kan worden verminderd wanneer de vaccins willekeurig worden verdeeld of wanneer prioriteit wordt gegeven aan beroepsgroepen met veel contacten met slechts de helft van het aantal vaccins. Bovendien kan de prioritering van beroepsgroepen de pieken verminderen en vertragen, zodat er minder gelijktijdige infecties optreden. Onze resultaten suggereren dus dat de beroepsgroep een gemakkelijk beschikbare en operationele benadering is om de doeltreffendheid van vaccinatiecampagnes te verbeteren.

Beperkingen en toekomstig werk

Met het NIDM hebben we een zeer flexibel model en een agentgebaseerde simulatie gepresenteerd die zinvolle en interessante inzichten oplevert. Uit ons experiment is echter gebleken dat delen van ons model een betere empirische onderbouwing vereisen met uitbreidingen voor het meenemen van de verfijnde strategieën gebruikt door deelnemers om

besmetting te vermijden. Hierbij valt te denken aan het meenemen van ziekte-toestanden van burens van burens voor netwerkbeslissingen en het verminderen van de aantrekkelijkheid van besmette anderen ongeacht de ziekte-toestand van ego.

Bovendien gebruikten we in dit proefschrift modellen die risicoperceptie beschouwden als een constante voorkeur. Onderzoek in de beginperiode van de COVID-19 pandemie heeft echter aangetoond dat risicopercepties niet alleen verschillen tussen individuen en landen (Dryhurst et al., 2020), maar ook veranderen over de tijd (Wise et al., 2020). Een element dat de risicoperceptie en het gezondheidsgerelateerd gedrag beïnvloedt, is de blootstelling aan media. Informatie over de wereldwijde prevalentie van een ziekte, zoals gerapporteerd door de media, zou dus de risicoperceptie van de agenten door het hele netwerk kunnen veranderen. Informatie over de lokale prevalentie, zoals gerapporteerd door persoonlijke contacten, zou de parameter juist in sterk besmette gebieden van het netwerk kunnen veranderen. Een dergelijke benadering zou het pad kunnen effenen voor een beter begrip van de invloed van wereldwijde informatiecampagnes en sociale beïnvloeding op het vermijden van ziekten.

Bovendien wordt gedragsaanpassing in het NIDM gemodelleerd vanuit het perspectief van vatbare individuen. Dat wil zeggen, afhankelijk van heterogene risicopercepties kunnen individuen sociale contacten aangaan of vermijden die infecties dragen. Hoewel dit in overeenstemming is met een groot aantal empirische studies (bijvoorbeeld Bish & Michie, 2010; Leppin & Aro, 2009) en een veelgebruikt mechanisme in op het individugebaseerde modellen voor gezondheidsgerelateerd gedrag, zoals het *Health Belief Model (HBM)*, zou het NIDM kunnen worden uitgebreid met een component die het voor besmette individuen minder aantrekkelijk maakt om sociale contacten aan te gaan. Afhankelijk van de ziekte kan een besmet persoon symptomen ervaren die hem dwingen thuis te blijven. Bovendien gaat besmetting gepaard met een ethische/normatieve verantwoordelijkheid om andere mensen niet in gevaar te brengen.

Met betrekking tot beleidsvorming kan de dynamiek van het NIDM belangrijke mechanismen blootleggen die moeten worden overwogen voor het ontwerp van interventies. Het laatste deel van dit proefschrift geeft een voorbeeld van hoe NIDM-simulaties en empirische gegevens expressief kunnen zijn voor de vergelijking van vaccinatiecampagnes binnen beroepsnetwerken. Daarnaast laat het zien dat het ontbreken van sociale netwerkdynamiek op zich geen probleem is voor netwerkmodellen van ziekteverspreiding. Het hangt eerder van de onderzoeksvragen af of verwacht wordt dat de netwerkdynamiek de infectiedynamiek significant zal veranderen. Statische netwerkmodellen blijven een plausibele abstractie wanneer bijvoorbeeld tijdschalen of soorten netwerken worden bestudeerd die minder vatbaar zijn voor substantiële relatieveranderingen.

Conclusie

Concluderend biedt dit proefschrift een nieuwe benadering om de co-evolutie van sociale netwerken en infectieziekten te beschrijven. Met het NIDM stellen we een flexibel modelkader voor dat individuele netwerkbeslissingen beschrijft als een afweging tussen de voorde-

len, kosten en mogelijke schade van een relatie door besmetting. Specifieke modelgevallen kunnen inzichten aan het licht brengen die kunnen worden gebruikt om epidemische interventies te ontwerpen of te verbeteren en toekomstig onderzoek te sturen. Deze bijdragen zijn van groot belang sinds de COVID-19-pandemie heeft aangetoond hoe kwetsbaar onze steeds meer geglobaliseerde wereld is voor uitbraken van infectieziekten.

References

- Abdulkareem, S. A., Augustijn, E.-W., Filatova, T., Musial, K., & Mustafa, Y. T. (2020). Risk perception and behavioral change during epidemics: Comparing models of individual and collective learning. *PLOS ONE*, *15*(1), e0226483.
- Abel, T., & McQueen, D. (2020). The COVID-19 pandemic calls for spatial distancing and social closeness: Not for social distancing! *International Journal of Public Health*, *65*, 231.
- Adam, D. C., Wu, P., Wong, J. Y., Lau, E. H. Y., Tsang, T. K., Cauchemez, S., Leung, G. M., & Cowling, B. J. (2020). Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nature Medicine*, *26*(11), 1714–1719.
- Ahmed, F., Zviedrite, N., & Uzicanin, A. (2018). Effectiveness of workplace social distancing measures in reducing influenza transmission: A systematic review. *BMC Public Health*, *18*(1), 518.
- Altieri, M. A., & Nicholls, C. I. (2020). Agroecology and the emergence of a post COVID-19 agriculture. *Agriculture and Human Values*, *37*(3), 525–526.
- Altig, D., Baker, S., Barrero, J. M., Bloom, N., Bunn, P., Chen, S., Davis, S. J., Leather, J., Meyer, B., Mihaylov, E., Mizen, P., Parker, N., Renault, T., Smietanka, P., & Thwaites, G. (2020). Economic uncertainty before and during the COVID-19 pandemic. *Journal of Public Economics*, *191*, 104274.
- Amit, S., Regev-Yochay, G., Afek, A., Kreiss, Y., & Leshem, E. (2021). Early rate reductions of SARS-CoV-2 infection and COVID-19 in BNT162b2 vaccine recipients. *The Lancet*, *397*(10277), 875–877.
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, *9*(4), 383–405.
- Arenas, A., Cota, W., Gómez-Gardenes, J., Gómez, S., Granell, C., Matamalas, J. T., Soriano, D., & Steinegger, B. (2020). A mathematical model for the spatiotemporal epidemic spreading of COVID-19. *MedRxiv*.
- Badham, J., & Stocker, R. (2010). The impact of network clustering and assortativity on epidemic behaviour. *Theoretical Population Biology*, *77*(1), 71–75.
- Balcan, D., Colizza, V., Gonçalves, B., Hu, H., Ramasco, J. J., & Vespignani, A. (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*, *106*(51), 21484–21489.
- Banerjee, D., & Rai, M. (2020). Social isolation in COVID-19: The impact of loneliness. *International Journal of Social Psychiatry*, *66*(6), 525–527.
- Bansal, S., Read, J., Pourbohloul, B., & Meyers, L. A. (2010). The dynamic nature of contact networks in infectious disease epidemiology. *Journal of Biological Dynamics*, *4*, 478–489.
- Barber, J. G., & Crisp, B. R. (1995). Social support and prevention of relapse following treatment for alcohol abuse. *Research on Social Work Practice*, *5*(3), 283–296.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

- Bedson, J., Skrip, L. A., Pedi, D., Abramowitz, S., Carter, S., Jalloh, M. F., Funk, S., Gobat, N., Giles-Vernick, T., Chowell, G., de Almeida, J. R., Elessawi, R., Scarpino, S. V., Hammond, R. A., Briand, S., Epstein, J. M., Hébert-Dufresne, L., & Althouse, B. M. (2021). A review and agenda for integrated disease models including social and behavioural factors. *Nature Human Behaviour*, 5(7), 834–846.
- Belot, M., Choi, S., Jamison, J. C., Papageorge, N. W., Tripodi, E., & van den Broek-Altenburg, E. (2020). Six-country survey on COVID-19. Retrieved May 28, 2020, from <https://osf.io/aubkc/>
- Berkman, L. F. (1984). Assessing the physical health effects of social networks and social support. *Annual Review of Public Health*, 5(1), 413–432.
- Berkman, L. F., & Leonard Syme, S. (1979). Social networks, host resistance, and mortality: A nine-year follow-up study of Alameda County residents. *American Journal of Epidemiology*, 109(2), 186–204.
- Bessi, A., Petroni, F., Vicario, M. D., Zollo, F., Anagnostopoulos, A., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). Homophily and polarization in the age of misinformation. *The European Physical Journal Special Topics*, 225(10), 2047–2059.
- Bi, Q., Wu, Y., Mei, S., Ye, C., Zou, X., Zhang, Z., Liu, X., Wei, L., Truelove, S. A., Zhang, T., Gao, W., Cheng, C., Tang, X., Wu, X., Wu, Y., Sun, B., Huang, S., Sun, Y., Zhang, J., ... Feng, T. (2020). Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A retrospective cohort study. *The Lancet Infectious Diseases*, 20(8), 911–919.
- Bish, A., & Michie, S. (2010). Demographic and attitudinal determinants of protective behaviours during a pandemic: A review. *British Journal of Health Psychology*, 15(4), 797–824.
- Block, P., Hoffman, M., Raabe, I. J., Dowd, J. B., Rahal, C., Kashyap, R., & Mills, M. C. (2020). Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nature Human Behaviour*, 4(6), 588–596.
- Borio, C. (2020). The COVID-19 economic crisis: Dangerously unique. *Business Economics*, 55(4), 181–190.
- Boroditsky, L. (2011). How language shapes thought. *Scientific American*, 304(2), 62–65.
- Brauer, F. (2008). Compartmental models in epidemiology. *Lecture Notes in Mathematics*, 1945, 19–79.
- Breban, R. (2011). Health newscasts for increasing influenza vaccination coverage: An inductive reasoning game approach. *PLOS ONE*, 6(12), e28300.
- Brethouwer, J. T., van de Rijt, A., Lindelauf, R., & Fokkink, R. (2021). “Stay nearby or get checked”: A COVID-19 control strategy. *Infectious Disease Modelling*, 6, 36–45.
- Brotherhood, L., Kircher, P., Santos, C., & Tertilt, M. (2020). An economic model of the COVID-19 epidemic: The importance of testing and age-specific policies. *CESifo Working Paper*.
- Buechel, B., & Buskens, V. (2013). The dynamics of closeness and betweenness. *Journal of Mathematical Sociology*, 37(3), 159–191.

- Bults, M., Beaujean, D. J. M. A., Richardus, J. H., & Voeten, H. A. C. M. (2015). Perceptions and behavioral responses of the general public during the 2009 Influenza A (H1N1) pandemic: A systematic review. *Disaster Medicine and Public Health Preparedness*, 9(2), 207–219.
- Burk, W. J., Steglich, C. E. G., & Snijders, T. A. B. (2007). Beyond dyadic interdependence: Actor-oriented models for co-evolving social networks and individual behaviors. *International Journal of Behavioral Development*, 31(4), 397–404.
- Burt, R. S. (1992). Structural holes: The social structure of competition.
- Burt, R. S. (2001). The social capital of structural holes. *New Directions in Economic Sociology*. SAGE.
- Buskens, V., Corten, R., & Przepiorka, W. (2022). Human behavior experiments in computational social science, (in preparation).
- Buskens, V., & Snijders, C. (2016). Effects of network characteristics on reaching the payoff-dominant equilibrium in coordination Games: A simulation study. *Dynamic Games and Applications*, 6(4), 477–494.
- Buskens, V., & Yamaguchi, K. (1999). A new model for information diffusion in heterogeneous social networks. *Sociological Methodology*, 29(1), 281–325.
- Camerer, C. F. (2003). Behavioural studies of strategic thinking in games. *Trends in Cognitive Sciences*, 7(5), 225–231.
- Catania, J. A., Kegeles, S. M., & Coates, T. J. (1990). Towards an understanding of risk behavior: An AIDS risk reduction model (ARRM). *Health Education Quarterly*, 17(1), 53–72.
- Centola, D., & Macy, M. (2007). Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3), 702–734.
- Chang, S. L., Piraveenan, M., Pattison, P., & Prokopenko, M. (2020). Game theoretic modelling of infectious disease dynamics and intervention methods: A review. *Journal of Biological Dynamics*, 14(1), 57–89.
- Cho, S. Y., Kang, J.-M., Ha, Y. E., Park, G. E., Lee, J. Y., Ko, J.-H., Lee, J. Y., Kim, J. M., Kang, C.-I., Jo, I. J., Ryu, J. G., Choi, J. R., Kim, S., Huh, H. J., Ki, C.-S., Kang, E.-S., Peck, K. R., Dhong, H.-J., Song, J.-H., ... Kim, Y.-J. (2016). MERS-CoV outbreak following a single patient exposure in an emergency room in South Korea: An epidemiological outbreak study. *The Lancet*, 388(10048), 994–1001.
- Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England Journal of Medicine*, 357(4), 370–379.
- Cohen, R., Havlin, S., & Ben-Avraham, D. (2003). Efficient immunization strategies for computer networks and populations. *Physical Review Letters*, 91(24), 247901.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *American Journal of Sociology*, 94, S95–S120.
- Coleman, J. S. (1994). *Foundations of Social Theory*. Harvard University Press.
- Colizza, V., Barrat, A., Barthélemy, M., Valleron, A.-J., & Vespignani, A. (2007). Modeling the worldwide spread of pandemic influenza: Baseline case and containment interventions. *PLOS Medicine*, 4(1), e13.

- Cook, K. S., Emerson, R. M., Gillmore, M. R., & Yamagashi, T. (1983). The distribution of power in exchange networks. *American Journal of Sociology*, 89(2), 275–305.
- Corten, R., & Buskens, V. (2010). Co-evolution of conventions and networks: An experimental study. *Social Networks*, 32(1), 4–15.
- Courtemanche, C., Garuccio, J., Le, A., Pinkston, J., & Yelowitz, A. (2020). Strong social distancing measures in the United States reduced the COVID-19 growth rate. *Health Affairs*, 39(7), 1237–1246.
- d’Andrea, V., Gallotti, R., Castaldo, N., & Domenico, M. D. (2022). Individual risk perception and empirical social structures shape the dynamics of infectious disease outbreaks. *PLOS Computational Biology*, 18, e1009760.
- Danon, L., Ford, A. P., House, T., Jewell, C. P., Keeling, M. J., Roberts, G. O., Ross, J. V., & Vernon, M. C. (2011). Networks and the epidemiology of infectious disease. *Interdisciplinary Perspectives on Infectious Diseases*, 2011, 28.
- Danon, L., Read, J. M., House, T. A., Vernon, M. C., & Keeling, M. J. (2013). Social encounter networks: Characterizing Great Britain. *Proceedings of the Royal Society B: Biological Sciences*, 280(1765), 20131037.
- Davis, R., Campbell, R., Hildon, Z., Hobbs, L., & Michie, S. (2015). Theories of behaviour and behaviour change across the social and behavioural sciences: A scoping review. *Health Psychology Review*, 9(3), 323–344.
- Della Rossa, F., Salzano, D., Di Meglio, A., De Lellis, F., Coraggio, M., Calabrese, C., Guarino, A., Cardona-Rivera, R., De Lellis, P., Liuzza, D., Lo Iudice, F., Russo, G., & di Bernardo, M. (2020). A network model of Italy shows that intermittent regional strategies can alleviate the COVID-19 epidemic. *Nature Communications*, 11(1), 1–9.
- Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, 66(5), 819–829.
- Dezső, Z., & Barabási, A.-L. (2002). Halting viruses in scale-free networks. *Physical Review E*, 65(5), 4.
- Dönges, P., Wagner, J., Contreras, S., Iftekhhar, E., Bauer, S., Mohr, S. B., Dehning, J., Valdez, A. C., Kretzschmar, M., Mäs, M., Nagel, K., & Priesemann, V. (2021). Interplay between risk perception, behaviour, and COVID-19 spread. *Frontiers in Physics*, 0, 68.
- Doreian, P., & Stokman, F. N. (1997). *Evolution of Social Networks*. Routledge.
- Dryhurst, S., Schneider, C. R., Kerr, J., Freeman, A. L. J., Recchia, G., van der Bles, A. M., Spiegelhalter, D., & van der Linden, S. (2020). Risk perceptions of COVID-19 around the world. *Journal of Risk Research*, 23(7-8), 994–1006.
- Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences*, 16(4), 681–694.
- Durham, D. P., & Casman, E. A. (2011). Incorporating individual health-protective decisions into disease transmission models: A mathematical framework. *Journal of The Royal Society Interface*, 9(68), 562–570.

- Elbay, R. Y., Kurtulmuş, A., Arpacioğlu, S., & Karadere, E. (2020). Depression, anxiety, stress levels of physicians and associated factors in COVID-19 pandemics. *Psychiatry Research*, 290, 113130.
- Elixir Core Team. (2021, January 4). Elixir. <https://elixir-lang.org/>
- Elmer, T., Mepham, K., & Stadtfeld, C. (2020). Students under lockdown: Comparisons of students' social networks and mental health before and during the COVID-19 crisis in Switzerland. *PLOS ONE*, 15(7), e0236337.
- Endo, A., Abbott, S., Kucharski, A. J., Funk, S., et al. (2020). Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Research*, 5.
- Falk, A., Becker, A., Dohmen, T. J., Huffman, D., & Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *IZA Discussion Papers*, 9674.
- Feld, S. L. (1981). The focused organization of social ties. *American Journal of Sociology*, 86, 1015–1035.
- Feld, S. L. (1991). Why your friends have more friends than you do. *American Journal of Sociology*, 96(6), 1464–1477.
- Ferguson, N. (2007). Capturing human behaviour. *Nature*, 446(7137), 733.
- Ferguson, N. M., Cummings, D. A. T., Fraser, C., Cajka, J. C., Cooley, P. C., & Burke, D. S. (2006). Strategies for mitigating an influenza pandemic. *Nature*, 442(7101), 448–452.
- Firth, J. A., Hellewell, J., Klepac, P., Kissler, S., Jit, M., Atkins, K. E., Clifford, S., Villabona-Arenas, C. J., Meakin, S. R., Diamond, C., Bosse, N. I., Munday, J. D., Prem, K., Foss, A. M., Nightingale, E. S., van Zandvoort, K., Davies, N. G., Gibbs, H. P., Medley, G., ... Spurgin, L. G. (2020). Using a real-world network to model localized COVID-19 control strategies. *Nature Medicine*, 26(10), 1616–1622.
- Flap, H., & Völker, B. (2004). *Creation and Returns of Social Capital*. Routledge.
- Flaxman, S., Mishra, S., Gandy, A., Unwin, H. J. T., Mellan, T. A., Coupland, H., Whitaker, C., Zhu, H., Berah, T., Eaton, J. W., Monod, M., Perez-Guzman, P. N., Schmit, N., Cilloni, L., Ainslie, K. E. C., Baguelin, M., Boonyasiri, A., Boyd, O., Cattarino, L., ... Bhatt, S. (2020). Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature*, 584(7820), 257–261.
- Fong, M. W., Gao, H., Wong, J. Y., Xiao, J., Shiu, E. Y. C., Ryu, S., & Cowling, B. J. (2020). Nonpharmaceutical measures for pandemic influenza in nonhealth-care settings-social distancing measures. *Emerging Infectious Diseases*, 26(5), 976–984.
- Funk, S., Gilad, E., Watkins, C., & Jansen, V. A. A. (2009). The spread of awareness and its impact on epidemic outbreaks. *Proceedings of the National Academy of Sciences*, 106(16), 6872–6877.
- Funk, S., Salathé, M., & Jansen, V. A. A. (2010). Modelling the influence of human behaviour on the spread of infectious diseases: A review. *Journal of the Royal Society Interface*, 7(50), 1247–1256.

- Galvani, A. P., & May, R. M. (2005). Dimensions of superspreading. *Nature*, 438(7066), 293–295.
- Gigerenzer, G., & Todd, P. M. (1999). *Simple Heuristics That Make Us Smart*. Oxford University Press.
- Glass, R. J., Glass, L. M., Beyeler, W. E., & Min, H. J. (2006). Targeted social distancing design for pandemic influenza. *Emerging Infectious Diseases*, 12(11), 1671–1681.
- Goldberg, E. L., Natta, P. V., & Comstock, G. W. (1985). Depressive symptoms, social networks and social support of elderly women. *American Journal of Epidemiology*, 121(3), 448–456.
- Goldberg, Y., Mandel, M., Woodbridge, Y., Fluss, R., Novikov, I., Yaari, R., Ziv, A., Freedman, L., & Huppert, A. (2021). Protection of previous SARS-CoV-2 infection is similar to that of BNT162b2 vaccine protection: A three-month nationwide experience from Israel. *MedRxiv*.
- Goodwin, R., Haque, S., Neto, F., & Myers, L. B. (2009). Initial psychological responses to influenza A, H1N1 ("Swine flu"). *BMC Infectious Diseases*, 9(1), 166.
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.
- Green, E. C., & Murphy, E. (2014). Health belief model. *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society* (pp. 766–769). Wiley.
- Grimm, V., Berger, U., Deangelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, 221, 2760–2768.
- Gross, T., D’Lima, C. J., & Blasius, B. (2006). Epidemic dynamics on an adaptive network. *Physical Review Letters*, 96, 208701.
- Guzzetta, G., Riccardo, F., Marziano, V., Poletti, P., Trentini, F., Bella, A., Andrianou, X., Del Manso, M., Fabiani, M., Bellino, S., Boros, S., Urdiales, A. M., Vescio, M. F., Piccioli, A., Brusaferrero, S., Rezza, G., Pezzotti, P., Ajelli, M., & Merler, S. (2020). Impact of a nationwide lockdown on SARS-CoV-2 transmissibility, Italy. *Emerging Infectious Diseases*, 27(1), 267.
- Halevy, N. (2016). Strategic thinking. *Advances in Experimental Social Psychology* (pp. 1–66). Elsevier.
- Hall, V. J., Foulkes, S., Saei, A., Andrews, N., Oguti, B., Charlett, A., Wellington, E., Stowe, J., Gillson, N., Atti, A., Islam, J., Karagiannis, I., Munro, K., Khawam, J., The SIREN Study Group, Chand, M., Brown, C., Ramsay, M., Bernal, J. L., & Hopkins, S. (2021). Effectiveness of BNT162b2 mRNA vaccine against infection and COVID-19 vaccine coverage in healthcare workers in England. *Multicentre Prospective Cohort Study (the SIREN Study)*, 10.
- Hamner, L. (2020). High SARS-CoV-2 attack rate following exposure at a choir practice — Skagit County, Washington, March 2020. *MMWR. Morbidity and mortality weekly report*, 69.
- Heape, A. (2021). Loneliness and social isolation in older adults: The effects of a pandemic. *Perspectives of the ASHA Special Interest Groups*, 6(6), 1729–1736.

- Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, 44(2), 174–199.
- Heesterbeek, H., Anderson, R. M., Andreasen, V., Bansal, S., DeAngelis, D., Dye, C., Eames, K. T. D., Edmunds, W. J., Frost, S. D. W., Funk, S., Hollingsworth, T. D., House, T., Isham, V., Klepac, P., Lessler, J., Lloyd-Smith, J. O., Metcalf, C. J. E., Mollison, D., Pellis, L., ... Isaac Newton Institute IDD Collaboration. (2015). Modeling infectious disease dynamics in the complex landscape of global health. *Science*, 347(6227).
- Hens, N., Shkedy, Z., Aerts, M., Faes, C., Van Damme, P., & Beutels, P. (2012). The SIR model. *Modeling Infectious Disease Parameters based on Serological and Social Contact Data* (pp. 25–58). Springer Science & Business Media.
- Herrmann, H. A., & Schwartz, J. M. (2020). Why COVID-19 models should incorporate the network of social interactions. *Physical Biology*, 17(6), 065008.
- Hoffmann, M., Schroeder, S., Kleine-Weber, H., Müller, M. A., Drosten, C., & Pöhlmann, S. (2020). Nafamostat mesylate blocks activation of SARS-CoV-2: New treatment option for COVID-19. *Antimicrobial Agents and Chemotherapy*, 64(6).
- House, J. S., Robbins, C., & Metzner, H. L. (1982). The association of social relationships and activities with mortality: Prospective evidence from the Tecumseh community health study. *American Journal of Epidemiology*, 116(1), 123–140.
- Hyafil, A., & Morina, D. (2020). Analysis of the impact of lockdown on the reproduction number of the SARS-Cov-2 in Spain. *Gaceta Sanitaria*.
- Institute of Medicine and National Research Council. (2011). *The Science of Adolescent Risk-Taking: Workshop Report*. Committee on the Science of Adolescence. The National Academies Press.
- Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press.
- Jackson, M. O., & Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1), 44–74.
- James, A., Pitchford, J. W., & Plank, M. J. (2007). An event-based model of superspreading in epidemics. *Proceedings of the Royal Society B: Biological Sciences*, 274(1610), 741–747.
- Jones, J. H., & Salathé, M. (2009). Early assessment of anxiety and behavioral response to novel swine-origin influenza A (H1N1). *PLOS ONE*, 4(12), e8032.
- Kahneman, B. Y. D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. *Handbook of the Fundamentals of Financial Decision Making: Part I*, 99–127.
- Karaivanov, A. (2020). A social network model of COVID-19. *PLOS ONE*, 15(10), e0240878.
- Kasl, S. V., & Cobb, S. (1966). Health behavior, illness behavior and sick role behavior: I. Health and illness behavior. *Archives of Environmental Health: An International Journal*, 12(2), 246–266.
- Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., Kasperson, J. X., & Ratick, S. (1988). The social amplification of risk: A conceptual framework. *Risk Analysis*, 8(2), 177–187.

- Kaur, S., Bherwani, H., Gulia, S., Vijay, R., & Kumar, R. (2020). Understanding COVID-19 transmission, health impacts and mitigation: Timely social distancing is the key. *Environment, Development and Sustainability*, 1–17.
- Kay, J. (2020). COVID-19 superspreader events in 28 countries: Critical patterns and lessons. *Quillette*, 23.
- Keeling, M. J. (1999). The effects of local spatial structure on epidemiological invasions. *Proceedings of the Royal Society B: Biological Sciences*, 266, 859–867.
- Kim, D. A., Hwong, A. R., Stafford, D., Hughes, D. A., O'Malley, A. J., Fowler, J. H., & Christakis, N. A. (2015). Social network targeting to maximise population behaviour change: A cluster randomised controlled trial. *The Lancet*, 386(9989), 145–153.
- Kim, H. H. S., & Jung, J. H. (2021). Social isolation and psychological distress during the COVID-19 pandemic: A cross-national analysis. *Gerontologist*, 61(1), 103–113.
- Kitchovitch, S., & Li, P. (2010). Risk perception and disease spread on social networks. *Procedia Computer Science*, 1(1), 2345–2354.
- Klov Dahl, A. S. (1985). Social networks and the spread of infectious diseases: The AIDS example. *Social Science & Medicine*, 21(11), 1203–1216.
- Koku, E., & Felsher, M. (2020). The effect of social networks and social constructions on HIV risk perceptions. *AIDS and Behavior*, 24, 206–221.
- Kretzschmar, M. E., Rozhnova, G., Bootsma, M. C. J., van Boven, M., van de Wijgert, J. H. H. M., & Bonten, M. J. M. (2020). Impact of delays on effectiveness of contact tracing strategies for COVID-19: A modelling study. *The Lancet Public Health*, 5(8), e452–e459.
- Kumari, R., Kumar, S., Poonia, R. C., Singh, V., Raja, L., Bhatnagar, V., & Agarwal, P. (2021). Analysis and predictions of spread, recovery, and death caused by COVID-19 in India. *Big Data Mining and Analytics*, 4(2), 65–75.
- Laguarta, J., Hueto, F., & Subirana, B. (2020). COVID-19 artificial intelligence diagnosis using only cough recordings. *IEEE Open Journal of Engineering in Medicine and Biology*, 1, 275–281.
- Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Schubert, J., Bania, J., & Khosrawipour, T. (2020). The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *Journal of Travel Medicine*, 27(3), taaa037.
- Lau, J. T. F., Yang, X., Tsui, H. Y., & Kim, J. H. (2005). Impacts of SARS on health-seeking behaviors in general population in Hong Kong. *Preventive Medicine*, 41, 454–462.
- Leifeld, P. (2013). texreg: Conversion of statistical model output in R to LaTeX and HTML tables. *Journal of Statistical Software*, 55(8), 1–24.
- Leppin, A., & Aro, A. R. (2009). Risk perceptions related to SARS and avian influenza: Theoretical foundations of current empirical research. *International Journal of Behavioral Medicine*, 16(1), 7–29.
- Leshem, E., & Lopman, B. A. (2021). Population immunity and vaccine protection against infection. *The Lancet*.

- Leung, K. Y., Ball, F., Sirl, D., & Britton, T. (2018). Individual preventive social distancing during an epidemic may have negative population-level outcomes. *Journal of The Royal Society Interface*, 15(145), 20180296.
- Leung, K., Jit, M., Lau, E. H., & Wu, J. T. (2017). Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Scientific Reports*, 7(1), 4–8.
- Levine-Tiefenbrun, M., Yelin, I., Katz, R., Herzel, E., Golan, Z., Schreiber, L., Wolf, T., Nadler, V., Ben-Tov, A., Kuint, J., Gazit, S., Patalon, T., Chodick, G., & Kishony, R. (2021). Initial report of decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine. *Nature Medicine*, 27(5), 790–792.
- Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K. S., Lau, E. H., Wong, J. Y., Xing, X., Xiang, N., Wu, Y., Li, C., Chen, Q., Li, D., Liu, T., Zhao, J., Liu, M., ... Feng, Z. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia [PMID: 31995857]. *New England Journal of Medicine*, 382(13), 1199–1207.
- Lin, J., Yan, K., Zhang, J., Cai, T., & Zheng, J. (2020). A super-spreader of COVID-19 in Ningbo city in China. *Journal of Infection and Public Health*, 13(7), 935–937.
- Lin, N., Ye, X., & Ensel, W. M. (1999). Social support and depressed mood : A structural analysis. *Journal of Health and Social Behavior*, 40(4), 344–359.
- Lindenfors, P., Wartel, A., & Lind, J. (2021). ‘Dunbar’s number’ deconstructed. *Biology Letters*, 17(5), 20210158.
- Little, S. J., Pond, S. L. K., Anderson, C. M., Young, J. A., Wertheim, J. O., Mehta, S. R., May, S., & Smith, D. M. (2014). Using HIV networks to inform real time prevention interventions. *PLOS ONE*, 9(6), e98443.
- Liu, F., Li, X., & Zhu, G. (2020). Using the contact network model and Metropolis-Hastings sampling to reconstruct the COVID-19 spread on the “Diamond Princess”. *Science Bulletin*, 65(15), 1297–1305.
- Lloyd-Smith, J. O., Schreiber, S. J., Kopp, P. E., & Getz, W. M. (2005). Superspreading and the effect of individual variation on disease emergence. *Nature*, 438(7066), 355–359.
- Lo Iacono, S., Przepiorka, W., Buskens, V., Corten, R., & van de Rijt, A. (2021). COVID-19 vulnerability and perceived norm violations predict loss of social trust: A pre-post study. *Social Science & Medicine*, 291, 114513.
- Long, S. J. (1997). *Regression Models for Categorical and Limited Dependent Variables*. SAGE.
- Longini Jr, I. M., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummings, D. A. T., & Halloran, M. E. (2005). Containing pandemic influenza at the source. *Science*, 309(5737), 1083–1087.
- Lunn, P. D., & Ní Choisdealbha, Á. (2018). The case for laboratory experiments in behavioural public policy. *Behavioural Public Policy*, 2(1), 22–40.

- Maher, P. J., MacCarron, P., & Quayle, M. (2020). Mapping public health responses with attitude networks: The emergence of opinion-based groups in the UK's early COVID-19 response phase. *British Journal of Social Psychology*, 59(3), 641–652.
- Mallapaty, S. (2021). Can COVID vaccines stop transmission? Scientists race to find answers. *Nature*.
- Manzo, G., & van de Rijt, A. (2020). Halting SARS-CoV-2 by targeting high-contact individuals. *arXiv preprint arXiv:2005.08907*.
- Mao, L., & Yang, Y. (2012). Coupling infectious diseases, human preventive behavior, and networks: A conceptual framework for epidemic modeling. *Social Science & Medicine*, 74, 167–175.
- Marceau, V., Noël, P. A., Hébert-Dufresne, L., Allard, A., & Dubé, L. J. (2010). Adaptive networks: Coevolution of disease and topology. *Physical Review E*, 82(3), 036116.
- Matrajt, L., & Leung, T. (2020). Evaluating the effectiveness of social distancing interventions to delay or flatten the epidemic curve of coronavirus disease. *Emerging Infectious Diseases*, 26(8), 1740–1748.
- May, R. M., & Anderson, R. M. (1987). Transmission dynamics of HIV infection. *Nature*, 326(6109), 137–142.
- McCord, C. (2020, January 26). Phoenix framework. <https://phoenixframework.org/>
- McKibbin, W., & Fernando, R. (2020). The economic impact of COVID-19. *Economics in the Time of COVID-19*, 45(10.1162).
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444.
- Milgram, S. (1967). The small-world problem. *Psychology Today*, 1(1), 61–67.
- Miller, D., Martin, M. A., Harel, N., Tirosh, O., Kustin, T., Meir, M., Sorek, N., Gefen-Halevi, S., Amit, S., Vorontsov, O., Shaag, A., Wolf, D., Peretz, A., Shemer-Avni, Y., Roif-Kaminsky, D., Kopelman, N. M., Huppert, A., Koelle, K., & Stern, A. (2020). Full genome viral sequences inform patterns of sars-cov-2 spread into and within israel. *Nature Communications*, 11(1), 5518.
- Miller, J. C. (2009). Percolation and epidemics in random clustered networks. *Physical Review E*, 80, 020901.
- Morsella, E., Bargh, J., & Gollwitzer, P. (Eds.). (2009). *Oxford Handbook of Human Action*. Oxford University Press.
- Mossong, J., Hens, N., Jit, M., Beutels, P., Auranen, K., Mikolajczyk, R., Massari, M., Salmaso, S., Tomba, G. S., Wallinga, J., Heijne, J., Sadkowska-Todys, M., Rosinska, M., & Edmunds, W. J. (2008). Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Medicine*, 5, 0381–0391.
- Newman, M. E. (2002). Assortative mixing in networks. *Physical Review Letters*, 89(20), 208701.
- Nishi, A., Dewey, G., Endo, A., Neman, S., Iwamoto, S. K., Ni, M. Y., Tsugawa, Y., Iosifidis, G., Smith, J. D., & Young, S. D. (2020). Network interventions for managing the COVID-19 pandemic and sustaining economy. *Proceedings of the National Academy of Sciences*, 117(48), 30285–30294.

- Nunner, H. (2020). Networking during infectious diseases model (NIDM) simulator [version: v4.1.0. Commit: f17e0b0. doi: 10.5281/zenodo.4290115].
- Nunner, H., Buskens, V., Corten, R., Kaandorp, C., & Kretzschmar, M. (2023). *Disease avoidance may come at the cost of social cohesion: Insights from a large-scale social networking experiment* [unpublished manuscript].
- Nunner, H., Buskens, V., & Kretzschmar, M. (2021). A model for the co-evolution of dynamic social networks and infectious disease dynamics. *Computational Social Networks*, 8(1), 19.
- Nunner, H., Buskens, V., Teslya, A., & Kretzschmar, M. (2022a). Health behavior homophily can mitigate the spread of infectious diseases in small-world networks. *Social Science & Medicine*, 312, 115350.
- Nunner, H., van de Rijt, A., & Buskens, V. (2022b). Prioritizing high-contact occupations raises effectiveness of vaccination campaigns. *Scientific Reports*, 12(1), 1–13.
- Nyabadza, F., Chiyaka, C., Mukandavire, Z., & Hove-Musekwa, S. D. (2010). Analysis of an HIV/AIDS model with public-health information campaigns and individual withdrawal. *Journal of Biological Systems*, 18, 357–375.
- Ormel, J., Lindenberg, S., Stevererink, N., & Verbrugge, L. M. (1999). Subjective well-being and social production functions. *Social Indicators Research*, 46(1), 61–90.
- Palla, G., Barabási, A.-L., & Vicsek, T. (2007). Quantifying social group evolution. *Nature*, 446(7136), 664–667.
- Pang, J., Huang, Y., Xie, Z., Li, J., & Cai, Z. (2021). Collaborative city digital twin for the COVID-19 pandemic: A federated learning solution. *Tsinghua Science and Technology*, 26(5), 759–771.
- Paré, P. E., Beck, C. L., & Başar, T. (2020). Modeling, estimation, and analysis of epidemics over networks: An overview. *Annual Reviews in Control*, 50, 345–360.
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P., & Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of Modern Physics*, 87(3), 925.
- Pastor-Satorras, R., & Vespignani, A. (2002). Immunization of complex networks. *Physical Review E*, 65(3), 036104.
- Peirlinck, M., Linka, K., Sahli Costabal, F., & Kuhl, E. (2020). Outbreak dynamics of COVID-19 in China and the United States. *Biomechanics and Modeling in Mechanobiology*, 19(6), 2179–2193.
- Perez, L., & Dragicevic, S. (2009). An agent-based approach for modeling dynamics of contagious disease spread. *International Journal of Health Geographics*, 8(50), 1–17.
- Pietrabissa, G., & Simpson, S. G. (2020). Psychological consequences of social isolation during COVID-19 outbreak. *Frontiers in Psychology*, 11, 2201.
- Pigné, Y., Dutot, A., Guinand, F., & Olivier, D. (2008). GraphStream: A Tool for bridging the gap between complex systems and dynamic graphs. *CoRR*, abs/0803.2.
- Poletti, P., Ajelli, M., & Merler, S. (2012). Risk perception and effectiveness of uncoordinated behavioral responses in an emerging epidemic. *Mathematical Biosciences*, 238(2), 80–89.

- R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Reluga, T. C. (2010). Game theory of social distancing in response to an epidemic. *PLOS Computational Biology*, 6(5), e1000793.
- Risau-Gusman, S., & Zanette, D. H. (2009). Contact switching as a control strategy for epidemic outbreaks. *Journal of Theoretical Biology*, 257, 52–60.
- Rocha, L. E., Liljeros, F., & Holme, P. (2011). Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLOS Computational Biology*, 7(3), e1001109.
- Santamaría, L., & Hortal, J. (2020). COVID-19 effective reproduction number dropped during Spain’s nationwide dropdown, then spiked at lower-incidence regions. *Science of the Total Environment*, 751, 142257.
- Scata, M., Attanasio, B., Aiosa, G. V., & Corte, A. L. (2020). The dynamical interplay of collective attention, awareness, and epidemics spreading in the multiplex social networks during COVID-19. *IEEE Access*, 8, 189203–189223.
- Sepúlveda-Loyola, W., Rodríguez-Sánchez, I., Pérez-Rodríguez, P., Ganz, F., Torralba, R., Oliveira, D. V., & Rodríguez-Mañas, L. (2020). Impact of social isolation due to COVID-19 on health in older people: Mental and physical effects and recommendations. *The Journal of Nutrition, Health & Aging*, 24(9), 938–947.
- Settersten, R. A., Bernardi, L., Härkönen, J., Antonucci, T. C., Dykstra, P. A., Heckhausen, J., Kuh, D., Mayer, K. U., Moen, P., Mortimer, J. T., Mulder, C. H., Smeeding, T. M., van der Lippe, T., Hagestad, G. O., Kohli, M., Levy, R., Schoon, I., & Thomson, E. (2020). Understanding the effects of COVID-19 through a life course lens. *Advances in Life Course Research*, 45, 100360.
- Shen, Y., Xu, W., Li, C., Handel, A., Martinez, L., Ling, F., Ebell, M., Fu, X., Pan, J., Ren, J., Gu, W., & Chen, E. (2020). A cluster of novel coronavirus disease 2019 infections indicating person-to-person transmission among casual contacts from social gatherings: An outbreak case-contact investigation [ofaa231]. *Open Forum Infectious Diseases*, 7(6).
- Siche, R., & Siche, R. (2020). What is the impact of COVID-19 disease on agriculture? *Scientia Agropecuaria*, 11(1), 3–6.
- Silva, C. J., Cantin, G., Cruz, C., Fonseca-Pinto, R., Passadouro, R., Soares dos Santos, E., & Torres, D. F. M. (2022). Complex network model for COVID-19: Human behavior, pseudo-periodic solutions and multiple epidemic waves. *Journal of Mathematical Analysis and Applications*, 514(2), 125171.
- Simmel, G. (1950). *The Sociology of Georg Simmel*. Simon; Schuster.
- Singh, K. K., & Singh, A. (2021). Diagnosis of COVID-19 from chest X-ray images using wavelets-based depthwise convolution network. *Big Data Mining and Analytics*, 4(2), 84–93.
- Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science*, 15(6), 322–325.

- Slovic, P., & Weber, E. U. (2002). Perception of risk posed by extreme events. *Risk Management Strategies in an Uncertain World*, 1–21.
- Stein, R. A. (2011). Super-spreaders in infectious diseases. *International Journal of Infectious Diseases*, 15(8), e510–e513.
- Sun, J., He, W. T., Wang, L., Lai, A., Ji, X., Zhai, X., Li, G., Suchard, M. A., Tian, J., Zhou, J., Veit, M., & Su, S. (2020). COVID-19: Epidemiology, evolution, and cross-disciplinary perspectives. *Trends in Molecular Medicine*, 26(5), 483–495.
- Sun, L., Axhausen, K. W., Lee, D.-H., & Cebrian, M. (2014). Efficient detection of contagious outbreaks in massive metropolitan encounter networks. *Scientific Reports*, 4(1), 1–6.
- Tagini, S., Brugnera, A., Ferrucci, R., Mazzocco, K., Compare, A., Silani, V., Pravettoni, G., & Poletti, B. (2021). It won't happen to me! Psychosocial factors influencing risk perception for respiratory infectious diseases: A scoping review. *Applied Psychology: Health and Well-Being*, 13(4), 835–852.
- Tande, A. J., Pollock, B. D., Shah, N. D., Farrugia, G., Virk, A., Swift, M., Breeher, L., Binnicker, M., & Berbari, E. F. (2021). Impact of the COVID-19 vaccine on asymptomatic infection among patients undergoing pre-procedural COVID-19 molecular screening. *Clinical Infectious Diseases*.
- Teovanović, P., Lukić, P., Zupan, Z., Lazić, A., Ninković, M., & Žeželj, I. (2021). Irrational beliefs differentially predict adherence to guidelines and pseudoscientific practices during the COVID-19 pandemic. *Applied Cognitive Psychology*, 35(2), 486–496.
- Teslya, A., Nunner, H., Buskens, V., & Kretzschmar, M. E. (2022a). The effect of competition between health opinions on epidemic dynamics. *PNAS Nexus*, (in press).
- Teslya, A., Pham, T. M., Godijk, N. G., Kretzschmar, M. E., Bootsma, M. C. J., & Rozhnova, G. (2020). Impact of self-imposed prevention measures and short-term government-imposed social distancing on mitigating and delaying a COVID-19 epidemic: A modelling study. *PLOS Medicine*, 17(7), e1003166.
- Teslya, A., Rozhnova, G., Pham, T. M., van Wees, D. A., Nunner, H., Godijk, N. G., Bootsma, M., & Kretzschmar, M. E. (2022b). The importance of sustained compliance with physical distancing during COVID-19 vaccination rollout. *Communications Medicine*, 2(1), 1–17.
- Thompson, M. G., Burgess, J. L., Naleway, A. L., Tyner, H. L., Yoon, S. K., Meece, J., Olsho, L. E. W., Caban-Martinez, A. J., Fowlkes, A., Lutrick, K., et al. (2021). Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers—eight US locations, December 2020 – March 2021. *Morbidity and Mortality Weekly Report*, 70(13), 495.
- Thu, T. P. B., Ngoc, P. N. H., Hai, N. M., & Tuan, L. A. (2020). Effect of the social distancing measures on the spread of COVID-19 in 10 highly infected countries. *Science of the Total Environment*, 742, 140430.

- Tracy, C. S., Rea, E., & Upshur, R. E. G. (2009). Public perceptions of quarantine: Community-based telephone survey following an infectious disease outbreak. *BMC Public Health*, 9(1), 470.
- Tran Kiem, C., Massonnaud, C. R., Levy-Bruhl, D., Poletto, C., Colizza, V., Bosetti, P., Fontanet, A., Gabet, A., Olié, V., Zanetti, L., Boëlle, P.-Y., Crépey, P., & Cauchemez, S. (2021). A modelling study investigating short and medium-term challenges for COVID-19 vaccination: From prioritisation to the relaxation of measures. *EClinicalMedicine*, 38, 101001.
- Trepte, S., Reinecke, L., & Juechems, K. (2012). The social side of gaming: How playing online computer games creates online and offline social support. *Computers in Human Behavior*, 28(3), 832–839.
- Tunc, I., Shkarayev, M. S., & Shaw, L. B. (2013). Epidemics in adaptive social networks with temporary link deactivation. *Journal of Statistical Physics*, 151(1-2), 355–366.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.
- U.S. Bureau of Labor Statistics. (2018). 2018 standard occupational classification system. Retrieved May 3, 2021, from https://www.bls.gov/soc/2018/major_groups.htm
- U.S. Bureau of Labor Statistics. (2019). Employment by major occupational group, 2019 and projected 2029. Retrieved May 3, 2021, from <https://www.bls.gov/emp/tables/emp-by-major-occupational-group.htm>
- Ustun, G. (2021). Determining depression and related factors in a society affected by COVID-19 pandemic. *The International Journal of Social Psychiatry*, 67(1), 54.
- Vaccines and immunizations - Interim list of categories of essential workers mapped to standardized industry codes and titles. (2021). Retrieved March 29, 2021, from <https://www.cdc.gov/vaccines/covid-19/categories-essential-workers.html>
- Valdez, L. D., Macri, P. A., & Braunstein, L. A. (2012a). Intermittent social distancing strategy for epidemic control. *Physical Review E*, 85(3), 036108.
- Valdez, L. D., Macri, P. A., & Braunstein, L. A. (2012b). Temporal percolation of the susceptible network in an epidemic spreading. *PLOS ONE*, 7(9), e44188.
- Valente, T. W. (2012). Network interventions. *Science*, 337(6090), 49–53.
- Verelst, F., Willem, L., & Beutels, P. (2016). Behavioural change models for infectious disease transmission: A systematic review (2010-2015). *Journal of the Royal Society Interface*, 13, 20160820.
- VoPham, T., Weaver, M. D., Hart, J. E., Ton, M., White, E., & Newcomb, P. A. (2020). Effect of social distancing on COVID-19 incidence and mortality in the US. *MedRxiv*.
- Vriens, E., & Buskens, V. (2020). *Managing risk heterogeneity in risk-sharing groups: A multi-method study on risk aversion and solidarity* [Accessed: 2022-05-19].
- Wang, Z., Andrews, M. A., Wu, Z.-X., Wang, L., & Bauch, C. T. (2015). Coupled disease–behavior dynamics on complex networks: A review. *Physics of Life Reviews*, 15, 1–29.
- Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.

- Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, 105(2), 493–527.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440.
- Weeden, K. A., & Cornwell, B. (2020). The small-world network of college classes: Implications for epidemic spread on a university campus. *Sociological Science*, 7, 222–241.
- Weesie, J., & Flap, H. (1990). *Social Networks Through Time*. Isor.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.
- Willem, L., Verelst, F., Bilcke, J., Hens, N., & Beutels, P. (2017). Lessons from a decade of individual-based models for infectious disease transmission: A systematic review (2006-2015). *BMC Infectious Diseases*, 17, 612.
- Wise, T., Zbozinek, T. D., Michelini, G., Hagan, C. C., & Mobbs, D. (2020). Changes in risk perception and self-reported protective behaviour during the first week of the COVID-19 pandemic in the United States: COVID-19 risk perception and behavior. *Royal Society Open Science*, 7(9).
- Woike, J. K., Hafenbrädl, S., Kanngiesser, P., & Hertwig, R. (2022). The transmission game: Testing behavioral interventions in a pandemic-like simulation. *Science Advances*, 8(8), 1–12.
- Wong, G., Liu, W., Liu, Y., Zhou, B., Bi, Y., & Gao, G. F. (2015). MERS, SARS, and Ebola: The role of super-spreaders in infectious disease. *Cell Host & Microbe*, 18(4), 398–401.
- Woolhouse, M. E. J., Dye, C., Etard, J.-F., Smith, T., Charlwood, J. D., Garnett, G. P., Hagan, P., Hii, J. L. K., Ndhlovu, P. D., Quinnell, R. J., Watts, C. H., Chandiwana, S. K., & Anderson, R. M. (1997). Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proceedings of the National Academy of Sciences*, 94(1), 338–342.
- World Health Organization. (2021). Coronavirus disease (COVID-19): How is it transmitted? Retrieved December 5, 2021, from <https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-how-is-it-transmitted>
- World Health Organization. (2022a). Vaccine equity. Retrieved July 19, 2022, from <https://www.who.int/campaigns/vaccine-equity>
- World Health Organization. (2022b). WHO coronavirus (COVID-19) dashboard. Retrieved July 19, 2022, from <https://covid19.who.int/>
- Yamaguchi, K. (1990). Homophily and social distance in the choice of multiple friends: An analysis based on conditionally symmetric log-bilinear association model. *Journal of the American Statistical Association*, 85(410), 356–366.
- Zhan, X. X., Liu, C., Zhou, G., Zhang, Z. K., Sun, G. Q., Zhu, J. J. H., & Jin, Z. (2018). Coupling dynamics of epidemic spreading and information diffusion on complex networks. *Applied Mathematics and Computation*, 332, 437–448.

Acknowledgments

Every so often it feels like the times are changing fast. For me, and arguably for many others, this has never been more true than during the past few years. When I started working on my dissertation in 2017, I had to explain the importance of studying how social networks and infectious diseases co-evolve. About two and a half years later, the COVID-19 pandemic began and questions about the relevance of my work stopped. But to be fair, I was not some kind of visionary who was ahead of times. I was just a Masters student finishing his thesis, who discovered his passion for science and found an interesting sounding project with some of the best mentors, collaborators, and supporters one could wish for.

First and foremost, I am extremely grateful to Vincent and Mirjam, without whom I could not have undertaken this journey. Besides the interesting sounding idea of the project, it was really a personal meeting with Vincent before I applied that made me realize that this is what I want to dedicate the next few years of my professional life to. Your mix of being knowledgeable, empathic, and a bit nerdy in just the right ways was what convinced me that I will get the support needed for the challenge of becoming a scientist. I already miss the meetings in which we just kept talking about ideas, results, how to proceed with our work, or just personal things. Thank you, also for being patient explaining some statistical concepts that didn't always come easy to me, and especially for your empathy in times of personal hardship. In any case, there is one particular memory I will never forget, and that is the moment when what, we thought, was a tranquil dinner in St. Petersburg turned into a full-blown party with you, Eva, and me doing traditional dances.

Mirjam, I am extremely grateful for all your support over the last years. On a professional level, you always had good advice that inspired me to take my work a step further by limiting myself. With your help, I have learned to focus better on what's interesting and relevant, to work more effectively, and to put my work into words better. On a personal level, I will always fondly remember the social gatherings at your house when the colleagues from the Modeling Group came to your potluck dinners. Thank you, also for advice that went beyond the boundaries of our project and for introducing me to your extended network. I very much look forward to continuing to see you on a regular basis.

Special thanks go to Chris and Wojtek who recognized my scientific talent before anyone else, including me. Thanks, Chris, for the opportunity to gather practical experience during a research internship and for supporting me especially throughout the early stages of becoming a scientist. Thank you, Wojtek, for playing a significant part in my transition to the social sciences. I am very proud of the work we three did together and grateful for all the discussions and experiences along the way.

Special thanks also go to my collaborators Sasha, Arnout, Rense, and Casper and to Eva Vriens, Eva Jaspers, Marcus, and Kasper for your support. Sasha, I want to thank you for being such a great colleague and friend who accompanied me for the most part of my journey. The discussions we had, professionally and even more so personally, will always have a special place in my memory. Thank you, Arnout and Rense, for securing additional funding and considering me a suitable candidate to perform the associated work. I know

without this opportunity, my thesis could not have been the same. Casper, I am truly grateful for all the hard work you put into realizing the software for the experiment and for sharing the responsibility of conducting it. It was such a nice change to work with a fellow programmer again for a while. Thank you, Eva (Vriens) for your advice setting up the experiment and for all the great memories in St. Petersburg, Florence, and of course Utrecht. Thank you, Eva (Jaspers) and Marcus, for your guidance, especially during my early teaching experiences. Thank you, Kasper for your help with the Dutch summary and thus saving me from embarrassment.

Without friends, old and new, this journey would have been long, lonesome, and surely less fun. Many have been supportive on a practical level and have already been mentioned. Others have been supportive on a more personal level. Thank you, Sanjana, for always having a listening ear, whether it was stress at work, lack of sleep, or other hardships along the way. You were sorely missed at the last pre-Christmas baking session. Many thanks to Jan and Bene for keeping in touch so regularly over all these years. Without our online sessions, I might have had more sleep, but definitely poorer mental well-being. I also want to thank, Tara and Rita, the best roommates one could wish for. It was much fun to talk to you guys about work, the weekend, holidays, future plans, and everything else. Unfortunately, our time together ended so suddenly. Thanks also to all the other ICS PhD candidates and colleagues from the sociology department that are too many to mention personally. You gave the extra something, a feeling of belonging, whether it was travelling together to the ICS events, having lunch together, or celebrating Sinterklaas.

Many thanks to everyone who commented on my work and thus played a role in shaping it. Specifically, I want to mention the previously unmentioned members of the Collaborative Relations Seminar Group, Werner and Philipp, the members of the Infectious Disease Modelling Group, Ganna, Kim, Martin, Noortje, Mui, Emil, Tahir, and Michiel, the members of the SocNetID project consortium, Mart, Nicole, and Jon, the early and former members of the SocNetID project group, Ema and Caitlin, as well as everyone from the Centre for Complex Systems Studies at Utrecht University.

Many thanks also go to the Social Networks Lab at ETH Zürich for having me for a short but very memorable research internship. Thank you, Christoph and Alvaro, for discussing my research in so much detail and for introducing me to the concept of Dynamic Network Actor Models. Thank you, Ulrik, Alejandro, Kieran, Jennifer, Seyed, Meher, and Denise, for the inspiration, the company, and all the great coffee.

I would like to extend my sincere thanks to my new colleagues from the University of Lübeck, especially André, Leonard, Lilian, Maged, and Toni for the warm welcome. André, I am truly grateful for the understanding and time you provide me for polishing my thesis. I am very much looking forward to what comes next.

Zum Schluss möchte ich noch meiner Familie danken, die mir zu jeder Zeit und in allen Lebenslagen zur Seite steht. Hierbei geht ein ganz besonderer Dank an meine lieben Eltern, die mir überhaupt erst eine Ausbildung ermöglicht haben, welche mit dem Anfertigen einer Doktorarbeit ihren Höhepunkt erreicht. Vielen Dank, für all die Freiheiten und dass ihr mich auf jegliche Art und Weise unterstützt habt, meine Ideen und Ziele zu verfolgen, auch

wenn sie für euch vielleicht nicht immer unmittelbar nachvollziehbar waren. Ein besonderer Dank geht auch an meine Schwester Sandra, die immer an meinen Erfolg geglaubt hat (vielleicht manchmal mehr als ich) und mir emotional immer zur Seite steht. Vielen Dank auch an meinen Sohn Nikolas, der es schafft wie kein anderer mich von Stress und Sorgen rund um die Wissenschaften abzulenken. Danke auch, dass du in mancher Zeit auf mich verzichten musstest und es mir niemals vorgehalten hast. Mein allergrößter Dank aber geht an meine Frau Katerina. Danke für die tägliche Unterstützung, dein Verständnis und dafür, dass wir es immer wieder gemeinsam schaffen, die Herausforderungen wissenschaftlichen Arbeitens und familiären Lebens zu meistern; sind die Hürden auch noch so groß. Ohne dich hätte ich diese Reise nicht begonnen, ganz zu schweigen, beenden können. Für dieses Zuhause, das unabhängig von Ort und Zeit ist, werde ich euch auf ewig dankbar sein.

About the author

Hendrik Nunner was born in Henstedt-Ulzburg, Germany, on January 29, 1983. In 2010, he obtained his bachelor's degree in Media Informatics at Fachhochschule Wedel, Germany. After several years of professional experience as a software developer, he completed the research master Artificial Intelligence in 2017 at Utrecht University. In the same year, he started working as a Ph.D. Candidate at the Interuniversity Centre for Social Science Theory and Methodology (ICS), the department of Sociology, and the research group Cooperative Relations at Utrecht University. For the time of his employment as Ph.D. Candidate he was a member of the Infectious Disease Modelling group at UMC Utrecht and a junior member of the Centre for Complex Systems Studies (CCSS) at Utrecht University. He wrote his dissertation under the supervision of prof. dr. ir. Vincent Buskens (Department of Sociology, Utrecht University) and prof. dr. Mirjam Kretzschmar (Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht).

Hendrik Nunner currently works as a postdoctoral fellow at the Institute for Multimedia and Interactive Systems in Lübeck, Germany. His research revolves around the interplay of human behavior and infectious disease dynamics, with a focus on disease spread on social networks and the effects of opinions and (dis-)information on health behavior. He is proficient in a variety of theoretical and methodological tools, including agent-based models and simulations, network analysis, mathematical models of disease spread, lab and online experiments, game theory, and cognitive modeling.

Peer-reviewed publications

- Janssen, C. P., Everaert, E., Hendriksen, H. M. A., and Mensing, G. L., Tigchelaar, L. J., & Nunner, H. (2019). The influence of rewards on (sub-)optimal interleaving. *PLOS ONE*, 14(3), e0214027.
- Nunner, H., Buskens, V., & Kretzschmar, M. (2021). A model for the co-evolution of dynamic social networks and infectious disease dynamics. *Computational Social Networks*, 8(1), 19.
- Nunner, H., Buskens, V., Teslya, A., & Kretzschmar, M. (2022). Health behavior homophily can mitigate the spread of infectious diseases in small-world networks. *Social Science & Medicine*, 312, 115350.
- Nunner, H., Przepiorka, W., & Janssen, C. (2022). The role of reinforcement learning in the emergence of conventions: Simulation experiments with the repeated volunteer's dilemma. *Journal of Artificial Societies and Social Simulation*, 25(1), 7.
- Nunner, H., van de Rijt, A., & Buskens, V. (2022). Prioritizing high-contact occupations raises effectiveness of vaccination campaigns. *Scientific Reports*, 12(1), 1–13.
- Teslya, A., Nunner, H., Buskens, V., & Kretzschmar, M. E. (2022). The effect of competition between health opinions on epidemic dynamics. *PNAS Nexus*, (in press).
- Teslya, A., Rozhnova, G., Pham, T. M., van Wees, D. A., Nunner, H., Godijk, N. G., Bootsma, M., & Kretzschmar, M. (2022). The importance of sustained compliance with physical distancing during COVID-19 vaccination rollout. *Communications Medicine*, 2(1), 1–17.

Working papers

- Nunner, H., Buskens, V., Corten, R., Kaandorp, C., & Kretzschmar, M. (2023). Disease avoidance may come at the cost of social cohesion: Insights from a large-scale social networking experiment.

Data collections

- Nunner, H., Buskens, V., Corten, R., Kaandorp, C., & Kretzschmar, M. (2021). Effects of risk perception on epidemics in a small-world network game. Data available upon request.

Media appearances

- Omroep West (2022, January 14). Coronavirus: ADO wil vanaf 28 januari weer publiek in de stadions. <https://www.omroepwest.nl/nieuws/4513203/coronavirus-ado-wil-vanaf-28-januari-weer-publiek-in-de-stadions>

ICS dissertation series

The ICS series presents dissertations of the Interuniversity Center for Social Science Theory and Methodology. Each of these studies aims at integrating explicit theory formation with state-of-the-art empirical research or at the development of advanced methods for empirical research. The ICS was founded in 1986 as a cooperative effort of the universities of Groningen and Utrecht. Since 1992, the ICS has expanded to the University of Nijmegen and, since 2017, to the University of Amsterdam. Most of the projects are financed by the participating universities or by the Dutch Research Council. The international composition of the ICS graduate students is mirrored in the increasing international orientation of the projects and thus of the ICS series itself.

1. C. van Liere (1990), *Lastige leerlingen. Een empirisch onderzoek naar sociale oorzaken van probleemgedrag op basisscholen*. Amsterdam: Thesis Publishers
2. Marco H.D. van Leeuwen (1990), *Bijstand in Amsterdam, ca. 1800–1850. Armenzorg als beheersings- en overlevingsstrategie*. ICS dissertation, Utrecht
3. I. Maas (1990), *Deelname aan podiumkunsten via de podia, de media en actieve beoefening. Substitutie of leereffecten?* Amsterdam: Thesis Publishers
4. M.I. Broese van Groenou (1991), *Gescheiden netwerken. De relaties met vrienden en verwanten na echtscheiding*. Amsterdam: Thesis Publishers
5. Jan M.M. van den Bos (1991), *Dutch EC policy making. A model-guided approach to coordination and negotiation*. Amsterdam: Thesis Publishers
6. Karin Sanders (1991), *Vrouwelijke pioniers. Vrouwen en mannen met een ‘mannelijke’ hogere beroepsopleiding aan het begin van hun loopbaan*. Amsterdam: Thesis Publishers
7. Sjerp de Vries (1991), *Egoism, altruism, and social justice. Theory and experiments on cooperation in social dilemmas*. Amsterdam: Thesis Publishers
8. Ronald S. Batenburg (1991), *Automatisering in bedrijf*. Amsterdam: Thesis Publishers
9. Rudi Wielers (1991), *Selectie en allocatie op de arbeidsmarkt. Een uitwerking voor de informele en geïnstitutionaliseerde kinderopvang*. Amsterdam: Thesis Publishers
10. Gert P. Westert (1991), *Verschillen in ziekenhuisgebruik*. ICS dissertation, Groningen
11. Hanneke Hermesen (1992), *Votes and policy preferences. Equilibria in party systems*. Amsterdam: Thesis Publishers
12. Cora J.M. Maas (1992), *Probleemleerlingen in het basisonderwijs*. Amsterdam: Thesis Publishers
13. Ed A.W. Boxman (1992), *Contacten en carrière. Een empirisch-theoretisch onderzoek naar de relatie tussen sociale netwerken en arbeidsmarktposities*. Amsterdam: Thesis Publishers
14. Conny G.J. Taes (1992), *Kijken naar banen. Een onderzoek naar de inschatting van arbeidsmarktkansen bij schoolverlaters uit het middelbaar beroepsonderwijs*. Amsterdam: Thesis Publishers

15. Peter van Roozendaal (1992), *Cabinets in multi-party democracies. The effect of dominant and central parties on cabinet composition and durability*. Amsterdam: Thesis Publishers
16. Marcel van Dam (1992), *Regio zonder regie. Verschillen in en effectiviteit van gemeentelijk arbeidsmarktbeleid*. Amsterdam: Thesis Publishers
17. Tanja van der Lippe (1993), *Arbeidsverdeling tussen mannen en vrouwen*. Amsterdam: Thesis Publishers
18. Marc A. Jacobs (1993), *Software: Kopen of kopiëren? Een sociaal-wetenschappelijk onderzoek onder PC-gebruikers*. Amsterdam: Thesis Publishers
19. Peter van der Meer (1993), *Verdringing op de Nederlandse arbeidsmarkt. Sector- en sekseverschillen*. Amsterdam: Thesis Publishers
20. Gerbert Kraaykamp (1993), *Over lezen gesproken. Een studie naar sociale differentiatie in leesgedrag*. Amsterdam: Thesis Publishers
21. Evelien Zeggelink (1993), *Strangers into friends. The evolution of friendship networks using an individual oriented modeling approach*. Amsterdam: Thesis Publishers
22. Jaco Berveling (1994), *Het stempel op de besluitvorming. Macht, invloed en besluitvorming op twee Amsterdamse beleidsterreinen*. Amsterdam: Thesis Publishers
23. Wim Bernasco (1994), *Coupled careers. The effects of spouse's resources on success at work*. Amsterdam: Thesis Publishers
24. Liset van Dijk (1994), *Choices in child care. The distribution of child care among mothers, fathers and non-parental care providers*. Amsterdam: Thesis Publishers
25. Jos de Haan (1994), *Research groups in Dutch sociology*. Amsterdam: Thesis Publishers
26. K. Boahene (1995), *Innovation adoption as a socio-economic process. The case of the Ghanaian cocoa industry*. Amsterdam: Thesis Publishers
27. Paul E.M. Ligthart (1995), *Solidarity in economic transactions. An experimental study of framing effects in bargaining and contracting*. Amsterdam: Thesis Publishers
28. Roger Th. A.J. Leenders (1995), *Structure and influence. Statistical models for the dynamics of actor attributes, network structure, and their interdependence*. Amsterdam: Thesis Publishers
29. Beate Völker (1995), *Should auld acquaintance be forgot...? Institutions of communism, the transition to capitalism and personal networks: The case of East Germany*. Amsterdam: Thesis Publishers
30. A. Cancrinus-Matthijssse (1995), *Tussen hulpverlening en ondernemerschap. Beroepsuitoefening en taakopvattingen van openbare apothekers in een aantal West-Europese landen*. Amsterdam: Thesis Publishers
31. Nardi Steverink (1996), *Zo lang mogelijk zelfstandig. Naar een verklaring van verschillen in oriëntatie ten aanzien van opname in een verzorgingstehuis onder fysiek kwetsbare ouderen*. Amsterdam: Thesis Publishers

32. Ellen Lindeman (1996), *Participatie in vrijwilligerswerk*. Amsterdam: Thesis Publishers
33. Chris Snijders (1996), *Trust and commitments*. Amsterdam: Thesis Publishers
34. Koos Postma (1996), *Changing prejudice in Hungary. A study on the collapse of state socialism and its impact on prejudice against gypsies and Jews*. Amsterdam: Thesis Publishers
35. Jooske T. van Busschbach (1996), *Uit het oog, uit het hart? Stabiliteit en verandering in persoonlijke relaties*. Amsterdam: Thesis Publishers
36. René Torenvlied (1996), *Besluiten in uitvoering. Theorieën over beleidsuitvoering modelmatig getoetst op sociale vernieuwing in drie gemeenten*. Amsterdam: Thesis Publishers
37. Andreas Flache (1996), *The Double edge of networks. An analysis of the effect of informal networks on cooperation in social dilemmas*. Amsterdam: Thesis Publishers
38. Kees van Veen (1997), *Inside an internal labor market: Formal rules, flexibility and career lines in a Dutch manufacturing company*. Amsterdam: Thesis Publishers
39. Lucienne van Eijk (1997), *Activity and well-being in the elderly*. Amsterdam: Thesis Publishers
40. Róbert Gál. (1997), *Unreliability. Contract discipline and contract governance under economic transition*. Amsterdam: Thesis Publishers
41. Anne-Geerte van de Goor (1997), *Effects of regulation on disability duration*. ICS dissertation, Utrecht
42. Boris Blumberg (1997), *Das Management von Technologiekooperationen. Partnersuche und Verhandlungen mit dem Partner aus Empirisch-Theoretischer Perspektive*. ICS dissertation, Utrecht
43. Marijke von Bergh (1997), *Loopbanen van oudere werknemers*. Amsterdam: Thesis Publishers
44. Anna Petra Nieboer (1997), *Life-events and well-being: A prospective study on changes in well-being of elderly people due to a serious illness event or death of the spouse*. Amsterdam: Thesis Publishers
45. Jacques Niehof (1997), *Resources and social reproduction: The effects of cultural and material resources on educational and occupational careers in industrial nations at the end of the twentieth century*. ICS dissertation, Nijmegen
46. Ariana Need (1997), *The kindred vote. Individual and family effects of social class and religion on electoral change in the Netherlands, 1956–1994*. ICS dissertation, Nijmegen
47. Jim Allen (1997), *Sector composition and the effect of education on wages: An international comparison*. Amsterdam: Thesis Publishers
48. Jack B.F. Hutten (1998), *Workload and provision of care in general practice. An empirical study of the relation between workload of Dutch general practitioners and the content and quality of their care*. ICS dissertation, Utrecht

49. Per B. Kropp (1998), *Berufserfolg im Transformationsprozeß, Eine theoretisch-empirische Studie über die Gewinner und Verlierer der Wende in Ostdeutschland*. ICS dissertation, Utrecht
50. Maarten H.J. Wolbers (1998), *Diploma-inflatie en verdringing op de arbeidsmarkt. Een studie naar ontwikkelingen in de opbrengsten van diploma's in Nederland*. ICS dissertation, Nijmegen
51. Wilma Smeenk (1998), *Opportunity and marriage. The impact of individual resources and marriage market structure on first marriage timing and partner choice in the Netherlands*. ICS dissertation, Nijmegen
52. Marinus Spreen (1999), *Sampling personal network structures: Statistical inference in ego-graphs*. ICS dissertation, Groningen
53. Vincent Buskens (1999), *Social networks and trust*. ICS dissertation, Utrecht
54. Susanne Rijken (1999), *Educational expansion and status attainment. A cross-national and over-time comparison*. ICS dissertation, Utrecht
55. Mérove Gijsberts (1999), *The legitimization of inequality in state-socialist and market societies, 1987–1996*. ICS dissertation, Utrecht
56. Gerhard G. Van de Bunt (1999), *Friends by choice. An actor-oriented statistical network model for friendship networks through time*. ICS dissertation, Groningen
57. Robert Thomson (1999), *The party mandate: Election pledges and government actions in the Netherlands, 1986–1998*. Amsterdam: Thela Thesis
58. Corine Baarda (1999), *Politieke besluiten en boeren beslissingen. Het draagvlak van het mestbeleid tot 2000*. ICS dissertation, Groningen
59. Rafael Wittek (1999), *Interdependence and informal control in organizations*. ICS dissertation, Groningen
60. Diane Payne (1999), *Policy making in the European Union: An analysis of the impact of the reform of the structural funds in Ireland*. ICS dissertation, Groningen
61. René Veenstra (1999), *Leerlingen – klassen – scholen. Prestaties en vorderingen van leerlingen in het voortgezet onderwijs*. Amsterdam: Thela Thesis
62. Marjolein Achterkamp (1999), *Influence strategies in collective decision making. A comparison of two models*. ICS dissertation, Groningen
63. Peter Mühlau (2000), *The governance of the employment relation. A relational signaling perspective*. ICS dissertation, Groningen
64. Agnes Akkerman (2000), *Verdeelde vakbeweging en stakingen. Concurrentie om leden*. ICS dissertation, Groningen
65. Sandra van Thiel (2000), *Quangocratization: Trends, causes and consequences*. ICS dissertation, Utrecht

66. Rudi Turksema (2000), *Supply of day care*. ICS dissertation, Utrecht
67. Sylvia E. Korupp (2000), *Mothers and the process of social stratification*. ICS dissertation, Utrecht
68. Bernard A. Nijstad (2000), *How the group affects the mind: Effects of communication in idea generating groups*. ICS dissertation, Utrecht
69. Inge F. de Wolf (2000), *Opleidingsspecialisatie en arbeidsmarktsucces van sociale wetenschappers*. ICS dissertation, Utrecht
70. Jan Kratzer (2001), *Communication and performance: An empirical study in innovation teams*. ICS dissertation, Groningen
71. Madelon Kroneman (2001), *Healthcare systems and hospital bed use*. ICS/NIVEL-dissertation, Utrecht
72. Herman van de Werfhorst (2001), *Field of study and social inequality. Four types of educational resources in the process of stratification in the Netherlands*. ICS dissertation, Nijmegen
73. Tamás Bartus (2001), *Social capital and earnings inequalities. The role of informal job search in Hungary*. ICS dissertation, Groningen
74. Hester Moerbeek (2001), *Friends and foes in the occupational career. The influence of sweet and sour social capital on the labour market*. ICS dissertation, Nijmegen
75. Marcel van Assen (2001), *Essays on actor perspectives in exchange networks and social dilemmas*. ICS dissertation, Groningen
76. Inge Sieben (2001), *Sibling similarities and social stratification. The impact of family background across countries and cohorts*. ICS dissertation, Nijmegen
77. Alinda van Bruggen (2001), *Individual production of social well-being. An exploratory study*. ICS dissertation, Groningen
78. Marcel Coenders (2001), *Nationalistic attitudes and ethnic exclusionism in a comparative perspective: An empirical study of attitudes toward the country and ethnic immigrants in 22 countries*. ICS dissertation, Nijmegen
79. Marcel Lubbers (2001), *Exclusionistic electorates. Extreme right-wing voting in Western Europe*. ICS dissertation, Nijmegen
80. Uwe Matzat (2001), *Social networks and cooperation in electronic communities. A theoretical-empirical analysis of academic communication and internet discussion groups*. ICS dissertation, Groningen
81. Jacques P.G. Janssen (2002), *Do opposites attract divorce? Dimensions of mixed marriage and the risk of divorce in the Netherlands*. ICS dissertation, Nijmegen
82. Miranda Jansen (2002), *Waardenoriëntaties en partnerrelaties. Een panelstudie naar wederzijdse invloeden*. ICS dissertation, Utrecht

83. Anne Rigt Poortman (2002), *Socioeconomic causes and consequences of divorce*. ICS dissertation, Utrecht
84. Alexander Gattig (2002), *Intertemporal decision making*. ICS dissertation, Groningen
85. Gerrit Rooks (2002), *Contract en conflict: Strategisch management van inkooptransacties*. ICS dissertation, Utrecht
86. Károly Takács (2002), *Social networks and intergroup conflict*. ICS dissertation, Groningen
87. Thomas Gautschi (2002), *Trust and exchange, effects of temporal embeddedness and network embeddedness on providing and dividing a surplus*. ICS dissertation, Utrecht
88. Hilde Bras (2002), *Zeeuwse meiden. Dienen in de levensloop van vrouwen, ca. 1850–1950*. Amsterdam: Aksant Academic Publishers
89. Merijn Rengers (2002), *Economic lives of artists. Studies into careers and the labour market in the cultural sector*. ICS dissertation, Utrecht
90. Annelies Kassenberg (2002), *Wat scholieren bindt. Sociale gemeenschap in scholen*. ICS-dissertation, Groningen
91. Marc Verboord (2003), *Moet de meester dalen of de leerling klimmen? De invloed van literatuuronderwijs en ouders op het lezen van boeken tussen 1975 en 2000*. ICS dissertation, Utrecht
92. Marcel van Egmond (2003), *Rain falls on all of us (but some manage to get more wet than others): Political Context and Electoral Participation*. ICS dissertation, Nijmegen
93. Justine Horgan (2003), *High performance human resource management in Ireland and the Netherlands: Adoption and effectiveness*. ICS dissertation, Groningen
94. Corine Hoeben (2003), *LETS' be a community. Community in local exchange trading systems*. ICS dissertation, Groningen
95. Christian Steglich (2003), *The framing of decision situations. Automatic goal selection and rational goal pursuit*. ICS dissertation, Groningen
96. Johan van Wilsem (2003), *Crime and context. The impact of individual, neighborhood, city and country characteristics on victimization*. ICS dissertation, Nijmegen
97. Christiaan Monden (2003), *Education, inequality and health. The impact of partners and life course*. ICS dissertation, Nijmegen
98. Evelyn Hello (2003), *Educational attainment and ethnic attitudes. How to explain their relationship*. ICS dissertation, Nijmegen
99. Marnix Croes en Peter Tammes (2004), *Gif laten wij niet voortbestaan. Een onderzoek naar de overlevingskansen van joden in de Nederlandse gemeenten, 1940–1945*. Amsterdam: Aksant Academic Publishers
100. Ineke Nagel (2004), *Cultuurdeelname in de levensloop*. ICS dissertation, Utrecht

101. Marieke van der Wal (2004), *Competencies to participate in life. Measurement and the impact of school*. ICS dissertation, Groningen
102. Vivian Meertens (2004), *Depressive symptoms in the general population: A multifactorial social approach*. ICS dissertation, Nijmegen
103. Hanneke Schuurmans (2004), *Promoting well-being in frail elderly people. Theory and intervention*. ICS dissertation, Groningen
104. Javier Arregui (2004), *Negotiation in legislative decision-making in the European Union*. ICS dissertation, Groningen
105. Tamar Fischer (2004), *Parental divorce, conflict and resources. The effects on children's behaviour problems, socioeconomic attainment, and transitions in the demographic career*. ICS dissertation, Nijmegen
106. René Bekkers (2004), *Giving and volunteering in the Netherlands: Sociological and psychological perspectives*. ICS dissertation, Utrecht
107. Renée van der Hulst (2004), *Gender differences in workplace authority: An empirical study on social networks*. ICS dissertation, Groningen
108. Rita Smaniotto (2004), *'You scratch my back and I scratch yours' versus 'Love thy neighbour'. Two Proximate Mechanisms of Reciprocal Altruism*. ICS dissertation, Groningen
109. Maurice Gesthuizen (2004), *The life-course of the low-educated in the Netherlands: Social and economic risks*. ICS dissertation, Nijmegen
110. Carlijne Philips (2005), *Vakantiegemeenschappen. Kwalitatief en kwantitatief onderzoek naar gelegenheid- en refreshergemeenschap tijdens de vakantie*. ICS dissertation, Groningen
111. Esther de Ruijter (2005), *Household outsourcing*. ICS dissertation, Utrecht
112. Frank van Tubergen (2005), *The integration of immigrants in cross-national perspective: Origin, destination, and community effects*. ICS dissertation, Utrecht
113. Ferry Koster (2005), *For the time being. Accounting for inconclusive findings concerning the effects of temporary employment relationships on solidary behavior of employees*. ICS dissertation, Groningen
114. Carolien Klein Haarhuis (2005), *Promoting anti-corruption reforms. Evaluating the implementation of a World Bank anti-corruption program in seven African countries (1999–2001)*. ICS dissertation, Utrecht
115. Martin van der Gaag (2005), *Measurement of individual social capital*. ICS dissertation, Groningen
116. Johan Hansen (2005), *Shaping careers of men and women in organizational contexts*. ICS dissertation, Utrecht
117. Davide Barrera (2005), *Trust in embedded settings*. ICS dissertation, Utrecht

118. Mattijs Lambooi (2005), *Promoting cooperation. Studies into the effects of long-term and short-term rewards on cooperation of employees*. ICS dissertation, Utrecht
119. Lotte Vermeij (2006), *What's cooking? Cultural boundaries among Dutch teenagers of different ethnic origins in the context of school*. ICS dissertation, Utrecht
120. Mathilde Strating (2006), *Facing the challenge of rheumatoid arthritis. A 13-year prospective study among patients and cross-sectional study among their partners*. ICS dissertation, Groningen
121. Jannes de Vries (2006), *Measurement error in family background variables: The bias in the intergenerational transmission of status, cultural consumption, party preference, and religiosity*. ICS dissertation, Nijmegen
122. Stefan Thau (2006), *Workplace deviance: Four studies on employee motives and self-regulation*. ICS dissertation, Groningen
123. Mirjam Plantinga (2006), *Employee motivation and employee performance in child care. The effects of the introduction of market forces on employees in the Dutch child-care sector*. ICS dissertation, Groningen
124. Helga de Valk (2006), *Pathways into adulthood. A comparative study on family life transitions among migrant and Dutch youth*. ICS dissertation, Utrecht
125. Henrike Elzen (2006), *Self-management for chronically ill older people*. ICS dissertation, Groningen
126. Ayşe Güveli (2007), *New social classes within the service class in the Netherlands and Britain. Adjusting the EGP class schema for the technocrats and the social and cultural specialists*. ICS dissertation, Nijmegen
127. Willem-Jan Verhoeven (2007), *Income attainment in post-communist societies*. ICS dissertation, Utrecht
128. Marieke Voorpostel (2007), *Sibling support: The exchange of help among brothers and sisters in the Netherlands*. ICS dissertation, Utrecht
129. Jacob Dijkstra (2007), *The effects of externalities on partner choice and payoffs in exchange networks*. ICS dissertation, Groningen
130. Patricia van Echtelt (2007), *Time-greedy employment relationships: Four studies on the time claims of post-Fordist work*. ICS dissertation, Groningen
131. Sonja Vogt (2007), *Heterogeneity in social dilemmas: The case of social support*. ICS dissertation, Utrecht
132. Michael Schweinberger (2007), *Statistical methods for studying the evolution of networks and behavior*. ICS dissertation, Groningen
133. István Back (2007), *Commitment and evolution: Connecting emotion and reason in long-term relationships*. ICS dissertation, Groningen

134. Ruben van Gaalen (2007), *Solidarity and ambivalence in parent-child relationships*. ICS dissertation, Utrecht
135. Jan Reitsma (2007), *Religiosity and solidarity - Dimensions and relationships disentangled and tested*. ICS dissertation, Nijmegen
136. Jan Kornelis Dijkstra (2007) *Status and affection among (pre)adolescents and their relation with antisocial and prosocial behavior*. ICS dissertation, Groningen
137. Wouter van Gils (2007), *Full-time working couples in the Netherlands. Causes and consequences*. ICS dissertation, Nijmegen
138. Djamila Schans (2007), *Ethnic diversity in intergenerational solidarity*. ICS dissertation, Utrecht
139. Ruud van der Meulen (2007), *Brug over woelig water: Lidmaatschap van sportverenigingen, vriendschappen, kennissenkringen en veralgemeend vertrouwen*. ICS dissertation, Nijmegen
140. Andrea Knecht (2008), *Friendship selection and friends' influence. Dynamics of networks and actor attributes in early adolescence*. ICS dissertation, Utrecht
141. Ingrid Doorten (2008), *The division of unpaid work in the household: A stubborn pattern?* ICS dissertation, Utrecht
142. Stijn Ruiter (2008), *Association in context and association as context: Causes and consequences of voluntary association involvement*. ICS dissertation, Nijmegen
143. Janneke Joly (2008), *People on our minds: When humanized contexts activate social norms*. ICS dissertation, Groningen
144. Margreet Frieling (2008), *'Joint production' als motor voor actief burgerschap in de buurt*. ICS dissertation, Groningen
145. Ellen Verbakel (2008), *The partner as resource or restriction? Labour market careers of husbands and wives and the consequences for inequality between couples*. ICS dissertation, Nijmegen
146. Gijs van Houten (2008), *Beleidsuitvoering in gelaagde stelsels. De doorwerking van aanbevelingen van de Stichting van de Arbeid in het CAO-overleg*. ICS dissertation, Utrecht
147. Eva Jaspers (2008), *Intolerance over time. Macro and micro level questions on attitudes towards euthanasia, homosexuality and ethnic minorities*. ICS dissertation, Nijmegen
148. Gijs Weijters (2008), *Youth delinquency in Dutch cities and schools: A multilevel approach*. ICS dissertation, Nijmegen
149. Jessica Pass (2009), *The self in social rejection*. ICS dissertation, Groningen
150. Gerald Mollenhorst (2009), *Networks in contexts. How meeting opportunities affect personal relationships*. ICS dissertation, Utrecht
151. Tom van der Meer (2009), *States of freely associating citizens: Comparative studies into the impact of state institutions on social, civic and political participation*. ICS dissertation, Nijmegen

152. Manuela Vieth (2009), *Commitments and reciprocity in trust situations. Experimental studies on obligation, indignation, and self-consistency*. ICS dissertation, Utrecht
153. Rense Corten (2009), *Co-evolution of social networks and behavior in social dilemmas: Theoretical and empirical perspectives*. ICS dissertation, Utrecht
154. Arieke J. Rijken (2009), *Happy families, high fertility? Childbearing choices in the context of family and partner relationships*. ICS dissertation, Utrecht
155. Jochem Tolsma (2009), *Ethnic hostility among ethnic majority and minority groups in the Netherlands. An investigation into the impact of social mobility experiences, the local living environment and educational attainment on ethnic hostility*. ICS dissertation, Nijmegen
156. Freek Bucx (2009), *Linked lives: Young adults' life course and relations with parents*. ICS dissertation, Utrecht
157. Philip Wotschack (2009), *Household governance and time allocation. Four studies on the combination of work and care*. ICS dissertation, Groningen
158. Nienke Moor (2009), *Explaining worldwide religious diversity. The relationship between subsistence technologies and ideas about the unknown in pre-industrial and (post-)industrial societies*. ICS dissertation, Nijmegen
159. Lieke ten Brummelhuis (2009), *Family matters at work. Depleting and enriching effects of employees' family lives on work outcomes*. ICS dissertation, Utrecht
160. Renske Keizer (2010), *Remaining childless. Causes and consequences from a life course perspective*. ICS dissertation, Utrecht
161. Miranda Sentse (2010), *Bridging contexts: The interplay between family, child, and peers in explaining problem behavior in early adolescence*. ICS dissertation, Groningen
162. Nicole Tieben (2010), *Transitions, tracks and transformations. Social inequality in transitions into, through and out of secondary education in the Netherlands for cohorts born between 1914 and 1985*. ICS dissertation, Nijmegen
163. Birgit Pauksztat (2010), *Speaking up in organizations: Four studies on employee voice*. ICS dissertation, Groningen
164. Richard Zijdeman (2010), *Status attainment in the Netherlands, 1811-1941. Spatial and temporal variation before and during industrialization*. ICS dissertation, Utrecht
165. Rianne Kloosterman (2010), *Social background and children's educational careers. The primary and secondary effects of social background over transitions and over time in the Netherlands*. ICS dissertation, Nijmegen
166. Olav Aarts (2010), *Religious diversity and religious involvement. A study of religious markets in Western societies at the end of the twentieth century*. ICS dissertation, Nijmegen
167. Stephanie Wiesmann (2010), *24/7 negotiation in couples transition to parenthood*. ICS dissertation, Utrecht

168. Borja Martinovic (2010), *Interethnic contacts: A dynamic analysis of interaction between immigrants and natives in Western countries*. ICS dissertation, Utrecht
169. Anne Roeters (2010), *Family life under pressure? Parents' paid work and the quantity and quality of parent-child and family time*. ICS dissertation, Utrecht
170. Jelle Sijtsema (2010), *Adolescent aggressive behavior: Status and stimulation goals in relation to the peer context*. ICS dissertation, Groningen
171. Kees Keizer (2010), *The spreading of disorder*. ICS dissertation, Groningen
172. Michael Mäs (2010), *The diversity puzzle. Explaining clustering and polarization of opinions*. ICS dissertation, Groningen
173. Marie-Louise Damen (2010), *Cultuurdeelname en CKV. Studies naar effecten van kunsteducatie op de cultuurdeelname van leerlingen tijdens en na het voortgezet onderwijs*. ICS dissertation, Utrecht
174. Marieke van de Rakt (2011), *Two generations of crime: The intergenerational transmission of convictions over the life course*. ICS dissertation, Nijmegen
175. Willem Huijnk (2011), *Family life and ethnic attitudes. The role of the family for attitudes towards intermarriage and acculturation among minority and majority groups*. ICS dissertation, Utrecht
176. Tim Huijts (2011), *Social ties and health in Europe. Individual associations, cross-national variations, and contextual explanations*. ICS dissertation, Nijmegen
177. Wouter Steenbeek (2011), *Social and physical disorder. How community, business presence and entrepreneurs influence disorder in Dutch neighborhoods*. ICS dissertation, Utrecht
178. Miranda Vervoort (2011), *Living together apart? Ethnic concentration in the neighborhood and ethnic minorities' social contacts and language practices*. ICS dissertation, Utrecht
179. Agnieszka Kanas (2011), *The economic performance of immigrants. The role of human and social capital*. ICS dissertation, Utrecht
180. Lea Ellwardt (2011), *Gossip in organizations. A social network study*. ICS dissertation, Groningen
181. Annemarije Oosterwaal (2011), *The gap between decision and implementation. Decision making, delegation and compliance in governmental and organizational settings*. ICS dissertation, Utrecht
182. Natascha Notten (2011), *Parents and the media. Causes and consequences of parental media socialization*. ICS dissertation, Nijmegen
183. Tobias Stark (2011), *Integration in schools. A process perspective on students' interethnic attitudes and interpersonal relationships*. ICS dissertation, Groningen
184. Giedo Jansen (2011), *Social cleavages and political choices. Large-scale comparisons of social class, religion and voting behavior in Western democracies*. ICS dissertation, Nijmegen

185. Ruud van der Horst (2011), *Network effects on treatment results in a closed forensic psychiatric setting*. ICS dissertation, Groningen
186. Mark Levels (2011), *Abortion laws in European countries between 1960 and 2010. Legislative developments and their consequences for women's reproductive decision-making*. ICS dissertation, Nijmegen
187. Marieke van Londen (2012), *Exclusion of ethnic minorities in the Netherlands. The effects of individual and situational characteristics on opposition to ethnic policy and ethnically mixed neighbourhoods*. ICS dissertation, Nijmegen
188. Sigrid M. Mohnen (2012), *Neighborhood context and health: How neighborhood social capital affects individual health*. ICS dissertation, Utrecht
189. Asya Zhelyazkova (2012), *Compliance under controversy: Analysis of the transposition of European directives and their provisions*. ICS dissertation, Utrecht
190. Valeska Korff (2012), *Between cause and control: Management in a humanitarian organization*. ICS dissertation, Groningen
191. Maike Gieling (2012), *Dealing with diversity: Adolescents' support for civil liberties and immigrant rights*. ICS dissertation, Utrecht
192. Katya Ivanova (2012), *From parents to partners: The impact of family on romantic relationships in adolescence and emerging adulthood*. ICS dissertation, Groningen
193. Jelmer Schalk (2012), *The performance of public corporate actors: Essays on effects of institutional and network embeddedness in supranational, national, and local collaborative contexts*. ICS dissertation, Utrecht
194. Alona Labun (2012), *Social networks and informal power in organizations*. ICS dissertation, Groningen
195. Michał Bojanowski (2012), *Essays on social network formation in heterogeneous populations: Models, methods, and empirical analyses*. ICS dissertation, Utrecht
196. Anca Minescu (2012), *Relative group position and intergroup attitudes in Russia*. ICS dissertation, Utrecht
197. Marieke van Schellen (2012), *Marriage and crime over the life course. The criminal careers of convicts and their spouses*. ICS dissertation, Utrecht
198. Mieke Maliepaard (2012), *Religious trends and social integration: Muslim minorities in the Netherlands*. ICS dissertation, Utrecht
199. Fransje Smits (2012), *Turks and Moroccans in the Low Countries around the year 2000: determinants of religiosity, trend in religiosity and determinants of the trend*. ICS dissertation, Nijmegen
200. Roderick Sluiter (2012), *The diffusion of morality policies among Western European countries between 1960 and 2010. A comparison of temporal and spatial diffusion patterns of six morality and eleven non-morality policies*. ICS dissertation, Nijmegen

201. Nicoletta Balbo (2012), *Family, friends and fertility*. ICS dissertation, Groningen.
202. Anke Munniksma (2013), *Crossing ethnic boundaries: Parental resistance to and consequences of adolescents' cross-ethnic peer relations*. ICS dissertation, Groningen
203. Anja Abendroth (2013), *Working women in Europe. How the country, workplace, and family context matter*. ICS dissertation, Utrecht
204. Katia Begall (2013), *Occupational hazard? The relationship between working conditions and fertility*. ICS dissertation, Groningen
205. Hidde Bekhuis (2013), *The popularity of domestic cultural products: Cross-national differences and the relation to globalization*. ICS dissertation, Utrecht
206. Lieselotte Blommaert (2013), *Are Joris and Renske more employable than Rashid and Samira? A study on the prevalence and sources of ethnic discrimination in recruitment in the Netherlands using experimental and survey data*. ICS dissertation, Utrecht
207. Wiebke Schulz (2013), *Careers of men and women in the 19th and 20th centuries*. ICS dissertation, Utrecht
208. Ozan Aksoy (2013), *Essays on social preferences and beliefs in non-embedded social dilemmas*. ICS dissertation, Utrecht
209. Dominik Morbitzer (2013), *Limited farsightedness in network formation*. ICS dissertation, Utrecht
210. Thomas de Vroome (2013), *Earning your place: The relation between immigrants' economic and psychological integration in the Netherlands*. ICS dissertation, Utrecht
211. Marloes de Lange (2013), *Causes and consequences of employment flexibility among young people. Recent developments in the Netherlands and Europe*. ICS dissertation, Nijmegen
212. Roza Meuleman (2014), *Consuming the nation. Domestic cultural consumption: Its stratification and relation with nationalist attitudes*. ICS dissertation, Utrecht
213. Esther Havekes (2014), *Putting interethnic attitudes in context. The relationship between neighbourhood characteristics, interethnic attitudes and residential behaviour*. ICS dissertation, Utrecht
214. Zoltán Lippényi (2014), *Transitions toward an open society? Intergenerational occupational mobility in Hungary in the 19th and 20th centuries*. ICS dissertation, Utrecht
215. Anouk Smeekes (2014), *The presence of the past: Historical rooting of national identity and current group dynamics*. ICS dissertation, Utrecht
216. Michael Savelkoul (2014), *Ethnic diversity and social capital. Testing underlying explanations derived from conflict and contact theories in Europe and the United States*. ICS dissertation, Nijmegen
217. Martijn Hogerbrugge (2014), *Misfortune and family: How negative events, family ties, and lives are linked*. ICS dissertation, Utrecht

218. Gina Potarca (2014), *Modern love. Comparative insights in online dating preferences and assortative mating*. ICS dissertation, Groningen
219. Mariska van der Horst (2014), *Gender, aspirations, and achievements: Relating work and family aspirations to occupational outcomes*. ICS dissertation, Utrecht
220. Gijs Huitsing (2014), *A social network perspective on bullying*. ICS dissertation, Groningen
221. Thomas Kowalewski (2015), *Personal growth in organizational contexts*. ICS dissertation, Groningen
222. Manu Muñoz-Herrera (2015), *The impact of individual differences on network relations: Social exclusion and inequality in productive exchange and coordination games*. ICS dissertation, Groningen
223. Tim Immerzeel (2015), *Voting for a change. The democratic lure of populist radical right parties in voting behavior*. ICS dissertation, Utrecht
224. Fernando Nieto Morales (2015), *The control imperative: Studies on reorganization in the public and private sectors*. ICS dissertation, Groningen
225. Jellie Sierksma (2015), *Bounded helping: How morality and intergroup relations shape children's reasoning about helping*. ICS dissertation, Utrecht
226. Tinka Veldhuis (2015), *Captivated by fear. An evaluation of terrorism detention policy*. ICS dissertation, Groningen
227. Miranda Visser (2015), *Loyalty in humanity. Turnover among expatriate humanitarian aid workers*. ICS dissertation, Groningen
228. Sarah Westphal (2015), *Are the kids alright? Essays on postdivorce residence arrangements and children's well-being*. ICS dissertation, Utrecht
229. Britta Rüschhoff (2015), *Peers in careers: Peer relationships in the transition from school to work*. ICS dissertation, Groningen
230. Nynke van Miltenburg (2015), *Cooperation under peer sanctioning institutions: Collective decisions, noise, and endogenous implementation*. ICS dissertation, Utrecht
231. Antonie Knigge (2015), *Sources of sibling similarity. Status attainment in the Netherlands during modernization*. ICS dissertation, Utrecht
232. Sanne Smith (2015), *Ethnic segregation in friendship networks. Studies of its determinants in English, German, Dutch, and Swedish school classes*. ICS dissertation, Utrecht
233. Patrick Präg (2015), *Social stratification and health. Four essays on social determinants of health and wellbeing*. ICS dissertation, Groningen
234. Wike Been (2015), *European top managers' support for work-life arrangements*. ICS dissertation, Utrecht
235. André Grow (2016), *Status differentiation: New insights from agent-based modeling and social network analysis*. ICS dissertation, Groningen

236. Jesper Rözer (2016), *Family and personal networks. How a partner and children affect social relationships*. ICS dissertation, Utrecht
237. Kim Pattiselanno (2016), *At your own risk: The importance of group dynamics and peer processes in adolescent peer groups for adolescents' involvement in risk behaviors*. ICS dissertation, Groningen
238. Vincenz Frey (2016), *Network formation and trust*. ICS dissertation, Utrecht
239. Rozemarijn van der Ploeg (2016), *Be a buddy, not a bully? Four studies on social and emotional processes related to bullying, defending, and victimization*. ICS dissertation, Groningen
240. Tali Spiegel (2016), *Identity, career trajectories and wellbeing: A closer look at individuals with degenerative eye conditions*. ICS dissertation, Groningen
241. Felix Tropf (2016), *Social Science Genetics and Fertility*. ICS dissertation, Groningen
242. Sara Geven (2016), *Adolescents problem behavior in school: the role of peer networks*. ICS dissertation, Utrecht
243. Josja Rokven (2016), *The victimization-offending relationship from a longitudinal perspective*. ICS dissertation, Nijmegen
244. Maja Djundeva (2016), *Healthy ageing in context: Family welfare state and the life course*. ICS dissertation, Groningen
245. Mark Visser (2017), *Inequality between older workers and older couples in the Netherlands. A dynamic life course perspective on educational and social class differences in the late career*. ICS dissertation, Nijmegen
246. Beau Oldenburg (2017), *Bullying in schools: The role of teachers and classmates*. ICS dissertation, Groningen
247. Tatang Muttaqin (2017), *The education divide in Indonesia: Four essays on determinants of unequal access to and quality of education*. ICS dissertation, Groningen
248. Margriet van Hek (2017), *Gender inequality in educational attainment and reading performance. A contextual approach*. ICS dissertation, Nijmegen
249. Melissa Verhoef (2017), *Work schedules, childcare and well-being. Essays on the associations between modern-day job characteristics, childcare arrangements and the well-being of parents and children*. ICS dissertation, Utrecht
250. Timo Septer (2017), *Goal priorities, cognition and conflict: Analyses of cognitive maps concerning organizational changes*. ICS dissertation, Groningen
251. Bas Hofstra (2017), *Online Social Networks: Essays on Membership, Privacy, and Structure*. ICS dissertation, Utrecht
252. Yassine Khoudja (2018), *Women's labor market participation across ethnic groups: The role of household conditions, gender role attitudes, and religiosity in different national context*. ICS dissertation, Utrecht

253. Joran Laméris (2018), *Living together in diversity. Whether, why and where ethnic diversity affects social cohesion*. ICS dissertation, Nijmegen
254. Maaïke van der Vleuten (2018), *Gendered Choices. Fields of study of adolescents in the Netherlands*. ICS dissertation, Utrecht
255. Mala Sondang Silitonga (2018), *Corruption in Indonesia: The impact of institutional change, norms, and networks*. ICS dissertation, Groningen
256. Manja Coopmans (2018), *Rituals of the past in the context of the present. The role of Remembrance Day and Liberation Day in Dutch society*. ICS dissertation, Utrecht
257. Paul Hindriks (2018), *The Struggle for Power: Attitudes towards the political participation of ethnic minorities*. ICS dissertation, Utrecht
258. Nynke Niezink (2018), *Modeling the dynamics of networks and continuous behavior*. ICS dissertation, Groningen
259. Simon de Bruijn (2018), *Reaching agreement after divorce and separation. Essays on the effectiveness of parenting plans and divorce mediation*. ICS dissertation, Utrecht
260. Susanne van 't Hoff-de Goede (2018), *While you were locked up. An empirical study on the characteristics, social surroundings and wellbeing of partners of prisoners in The Netherlands*. ICS dissertation, Utrecht
261. Loes van Rijsewijk (2018), *Antecedents and Consequences of Helping among Adolescents*. ICS dissertation, Groningen
262. Mariola Gremmen (2018), *Social network processes and academic functioning. The role of peers in students' school well-being, academic engagement, and academic achievement*. ICS dissertation, Groningen
263. Jeanette Renema (2018), *Immigrants' support for welfare spending. The causes and consequences of welfare usage and welfare knowledgeability*. ICS dissertation, Nijmegen
264. Suwatin Miharti (2018), *Community health centers in Indonesia in the era of decentralization. The impact of structure, staff composition and management on health outcomes*. ICS dissertation, Groningen
265. Chaïm la Roi (2019), *Stigma and stress: Studies on attitudes towards sexual minority orientations and the association between sexual orientation and mental health*. ICS dissertation, Groningen
266. Jelle Lössbroek (2019), *Turning grey into gold. Employer-employee interplay in an ageing workforce*. ICS dissertation, Utrecht
267. Nikki van Gerwen (2019), *Employee cooperation through training. A multi-method approach*. ICS dissertation, Utrecht
268. Paula Thijs (2019), *Trends in cultural conservatism: The role of educational expansion, secularisation, and changing national contexts*. ICS dissertation, Nijmegen

269. Renske Verweij (2019), *Understanding childlessness: Unravelling the link with genes and the socio-environment*. ICS dissertation, Groningen
270. Niels Blom (2019), *Partner relationship quality under pressing work conditions. Longitudinal and cross-national investigation*. ICS dissertation, Nijmegen
271. Müge Simsek (2019), *The dynamics of religion among native and immigrant youth in Western Europe*. ICS dissertation, Utrecht
272. Leonie van Breeschoten (2019), *Combining a career and childcare: The use and usefulness of work-family policies in European organizations*. ICS dissertation, Utrecht
273. Roos van der Zwan (2019), *The political representation of ethnic minorities and their vote choice*. ICS dissertation, Nijmegen
274. Ashwin Rambaran (2019), *The classroom as context for bullying: A social network approach*. ICS dissertation, Groningen
275. Dieko Bakker (2019), *Cooperation and social control: Effects of preferences, institutions, and social structure*. ICS dissertation, Groningen
276. Femke van der Werf (2019), *Shadow of a rainbow? National and ethnic belonging in Mauritius*. ICS dissertation, Utrecht
277. Robert Krause (2019), *Multiple Imputation for Missing Network Data*. ICS dissertation, Groningen
278. Take Sipma (2020), *Economic insecurity and populist radical right voting*. ICS dissertation, Nijmegen
279. Mathijs Kros (2020), *The nature of negative contact: studies on interethnic relations in western societies*. ICS dissertation, Utrecht
280. Lonneke van den Berg (2020), *Time to leave: Individual and contextual explanations for the timing of leaving home*. ICS dissertation, Amsterdam.
281. Marianne Hooijsma (2020), *Clashrooms: Interethnic peer relationships in schools*. ICS dissertation, Groningen
282. Marina Tulin (2020), *Blind Spots in Social Resource Theory: Essays on the Creation, Maintenance and Returns of Social Capital*. ICS dissertation, Amsterdam
283. Tessa Kaufman (2020), *Toward Tailored Interventions: Explaining, Assessing, and Preventing Persistent Victimization of Bullying*. ICS dissertation, Groningen
284. Lex Thijssen (2020), *Racial and Ethnic Discrimination in Western Labor Markets: Empirical Evidence from Field Experiments*. ICS dissertation, Utrecht
285. Lukas Norbutas (2020), *Trust on the Dark Web: An analysis of illegal online drug markets*. ICS dissertation, Utrecht
286. Tomáš Diviák (2020), *Criminal Networks: Actors, Mechanisms, and Structures*. ICS dissertation, Groningen

287. Tery Setiawan (2020), *Support for Interreligious Conflict in Indonesia*. ICS dissertation, Nijmegen
288. Vera de Bel (2020), *The ripple effect in family networks: Relational structures and well-being in divorced and non-divorced families*. ICS dissertation, Groningen
289. Diego Palacios (2020), *How Context and the Perception of Peers' Behaviors Shape Relationships in Adolescence: A Multiplex Social Network Perspective*. ICS dissertation, Groningen
290. Saskia Glas (2020), *Where are the Muslim Feminists?: Religiosity and Support for Gender Equality in the Arab region*. ICS dissertation, Nijmegen
291. Tomas Turner-Zwinkels (2020), *A New Macro-Micro Approach to the Study of Political Careers: Theoretical, Methodological and Empirical Challenges and Solutions*. ICS dissertation, Groningen
292. Lotte Scheeren (2020), *Not on the Same Track? Tracking Age and Gender Inequality in Education*. ICS dissertation, Amsterdam
293. Joris Broere (2020), *Essays on how social network structure affects asymmetric coordination and trust*. ICS dissertation, Utrecht
294. Marcus Kristiansen (2021), *Contact with Benefits: How Social Networks affect Benefit Receipt Dynamics in the Netherlands*. ICS dissertation, Utrecht
295. Judith Kas (2021), *Trust and reputation in the peer-to-peer platform economy*. ICS dissertation, Utrecht
296. Andrea Forster (2021), *Navigating Educational Institutions: Mechanisms of Educational Inequality and Social Mobility in Different Educational Systems*. ICS dissertation, Amsterdam
297. Jannes ten Berge (2021), *Technological change and work: The relation between technology implementation within organizations and changes in workers' employment*. ICS dissertation, Utrecht
298. Jolien Geerlings (2021), *Teaching in culturally diverse classrooms: The importance of dyadic relations between teachers and children*. ICS dissertation, Utrecht
299. Kirsten van Houdt (2021), *Stepfamilies in adulthood: Solidarity between parents and adult children*. ICS dissertation, Amsterdam
300. Suzanne de Leeuw (2021), *The intergenerational transmission of educational attainment after divorce and remarriage*. ICS dissertation, Amsterdam
301. Fleur Goedkoop (2021), *Involvement in bottom-up energy transitions: The role of local and contextual embeddedness*. ICS dissertation, Groningen
302. Eva Vriens (2021), *Mutualism in the 21st century: The why, when, and how behind successful risk-sharing institutions*. ICS dissertation, Utrecht
303. Ardita Muja (2021), *From school to work: The role of the vocational specificity of education in young people's labor market integration*. ICS dissertation, Nijmegen

304. Siyang Kong (2021), *Women and work in contemporary China: The effect of market transition on women's employment, earnings, and status attainment*. ICS dissertation, Utrecht
305. Marijn Keijzer (2022), *Opinion dynamics in online social media*. ICS dissertation, Groningen
306. Sander Kunst (2022), *The educational divide in openness towards globalisation in Western Europe*. ICS dissertation, Amsterdam
307. Nella Geurts (2022), *Puzzling pathways: The integration paradox among migrants in Western Europe*. ICS dissertation, Nijmegen
308. Dragana Stojmenovska (2022), *Men's place: The incomplete integration of women in workplace authority*. ICS dissertation, Amsterdam
309. Bram Hogendoorn (2022), *Divorce and inequality: Stratification in the risk and consequences of union dissolution*. ICS dissertation, Amsterdam
310. Tom Nijs (2022), *This place is ours: Collective psychological ownership and its social consequences*. ICS dissertation, Utrecht
311. Nora Storz (2022), *'This land is ours—but is it also theirs?' Collective ownership beliefs and reconciliation in territorial conflict regions*. ICS dissertation, Utrecht
312. Tara Koster (2022), *Parenting and fairness in diverse families*. ICS dissertation, Utrecht
313. Danelien van Aalst (2022), *Elements Contributing to Teachers' Role in Bullying*. ICS dissertation, Groningen
314. Wybren Nooitgedagt (2022), *Who owns the country? Collective psychological ownership and intergroup relations in settler societies*. ICS dissertation, Utrecht
315. Marija Dangubić (2022), *Rejecting Muslim minority practices: Principles and prejudices*. ICS dissertation, Utrecht
316. Sara Cvetkovska (2022), *Lines in the shifting sand: The implications of being tolerated*. ICS dissertation, Utrecht
317. Maaike Hornstra (2022), *Going beyond the dyad: Adult intergenerational closeness after divorce and remarriage*. ICS dissertation, Amsterdam
318. Wouter Kiekens (2022), *Sexual and gender minority youth's mental health and substance use: Disparities, mechanisms, and protective factors*. ICS dissertation, Groningen
319. Carlijn Bussemakers (2022), *Adversity and educational inequality: The interplay between adverse experiences and parental resources for children's educational attainment*. ICS dissertation, Nijmegen
320. Evi Velthuis (2022), *'To tolerate or not to tolerate?' Reasons for tolerance of minority group practices among majority members in the Netherlands and Germany*. ICS dissertation, Utrecht
321. Hendrik Nunner (2023), *Co-evolution of social networks and infectious diseases*. ICS dissertation, Utrecht

Since the beginning of the COVID-19 pandemic in early 2020 until the time of finishing this dissertation in July 2022, the World Health Organization has registered almost 560 Million cases and well over 6 Million deaths from COVID-19. Although a number of effective vaccines have been developed and large-scale vaccination campaigns have been implemented globally, only 58 of 194 WHO member states have met the 70% vaccine coverage goal. Thus, mathematical models of disease spread remain one of the most important tools to understand the spread of such dynamics and can thus support public health policymaking. Network models of disease spread, the type of model addressed in this dissertation, have the advantage of accounting for social structure. However, they typically neglect the changes in our social relationships, despite the evidence that people avoid public places or infectious others during times of increased risk of infection. This dissertation presents, to our knowledge, the first model for the co-evolution of social networks and infectious diseases considering theory from social sciences, health psychology and epidemiology. The Networking during Infectious Diseases Model (NIDM) describes networking decisions as a trade-off between the benefits, costs, and potential harms of infections that a social relation creates. Agent-based simulations suggest that disease avoidance may reduce infection numbers significantly, depending on risk perceptions and network structure. A large-scale networking experiment, however, suggests that when disease avoidance is prioritized over maintaining social relationships, the cost might be a loss of social cohesion.

Hendrik Nunner obtained his bachelor's degree in Media Informatics at Fachhochschule Wedel, Germany, and his research master in Artificial Intelligence at Utrecht University, Netherlands. He conducted the present study as part of his Ph.D. research at the Interuniversity Center for Social Science Theory and Methodology (ICS), the department of Sociology, the research group Cooperative Relations, and the Centre for Complex Systems Studies (CCSS) at Utrecht University, as well as the Infectious Disease Modelling group at UMC Utrecht.