Essays on the Economics of Public Health

Medical Decision Making and Epidemics

Corinna Marianne Margarete Hartung

Munich 2021

Essays on the Economics of Public Health

Medical Decision Making and Epidemics

Inaugural-Dissertation
zur Erlangung des Grades

Doctor oeconomiae publicae
(Dr. oec. publ.)

an der
Ludwig–Maximilians–Universität München

2021

vorgelegt von
Corinna Marianne Margarete Hartung

Referent: Prof. Dr. Joachim Winter
Korreferentin: Prof. Dr. Amelie Wuppermann
Promotionsabschlussberatung: 14. Juli 2021

Datum der mündlichen Prüfung: 5. Juli.2021
Namen der Berichterstatter: Prof. Dr. Joachim Winter
Prof. Dr. Amelie Wuppermann
Prof. Dr. Uwe Sunde

Acknowledgments

Throughout my doctoral studies, I have been very fortunate to meet and collaborate with many dedicated and inspiring people which has been a continuous source of motivation. This dissertation would not have been possible without the support of my advisers, colleagues, friends, and my family.

First and foremost, I would like to thank my advisor, Joachim Winter. I am deeply grateful to him for providing the environment, guidance, freedom, and resources that I needed to conduct my research. Joachim continuously supported my research interest in the cross-field between public health and economics and provided me with invaluable advice and encouragement. I am also very grateful to Amelie Wuppermann, my co-advisor, for inspiring discussions and valuable advice. Working together was both very insightful and extremely enjoyable.

I want to thank my co-authors Charles F. Manski, Pavel Obraztcov, Gregory Veramendi, Joachim Winter, and Amelie Wuppermann, who have taught and mentored me. I owe special thanks to Gregory Veramendi. I greatly benefited from his constructive advice, patience, and his contagious enthusiasm for research. I would also like to thank Charles F. Manski who hosted me during my stay at Northwestern University. I learned a lot from his critical questions and his inspiring way of thinking about research questions. Thanks also go to Uwe Sunde for providing valuable feedback on my research and agreeing to serve on my committee. I would also like to express my gratitude to Nicole Maestas who invited me to Harvard Medical School. It was a fantastic and stimulating experience to be surrounded of such a high density of scholars interested in health economics.

Furthermore, I thank the Robert Koch Institute for providing me access to data and for their hospitality during my research stays in Berlin. I gratefully acknowledge financial support from the German Science Foundation through the CRC TRR 190 “Rationality and Competition” and the doctoral program “GRK 1928: Microeconomic Determinants of Individual Productivity”.

My special gratitude goes to my fantastic colleagues and friends at the LMU and MPI who made sure that the years of my PhD were not all work but also lots of fun. I am particularly thankful to Dennis Byrski, Nadine Geiger, Felix Hagemeister, Michael Kaiser, Anne Niedermaier, Felix Pöge, Zhaoxin Pu, Peter Redler, Tobias Rossmann, Brendan Shanks, Fabian Stürmer-Heiber, and Christoph Winter.

I am grateful to my dear friends, who provided me with welcome distractions and gave me the energy to keep going. My deepest gratitude goes to my parents and my sisters. Thanks for your constant support, patience, and for reminding me of a world outside research!

Corinna Hartung
Munich, July 2021

Table of Contents

Preface

"Health is not everything, but without health everything is nothing."
Arthur Schopenhauer, 1788-1860

The World Health Organization proclaims that improving people's health and longevity is a key objective of social policy. It is, however, also a means to fostering individual productivity, well-being, and a country's economic growth (Sachs 2001). The statement of the World Health Organization highlights the contribution of improving population health to economic and social development beyond reducing the prevalence of diseases. The value of individual and population health for economic performance is well documented (see e.g. Almond et al. 2018; Bloom et al. 2020; Prinz et al. 2018; Weil 2007). Health can operate through numerous channels to promote economic well-being, which can in turn create additional resources to invest in health (Bärnighausen et al. 2014). Healthy individuals have a better educational attainment, a better labor market performance, are wealthier and happier (e.g. Baird et al. 2016; Case et al. 2002; Currie 2009; Smith 1999). Moreover, healthy populations tend to save more in anticipation of a longer retirement and attract more foreign direct investment contributing to capital accumulation and technological progress (Alsan et al. 2006; Bärnighausen et al. 2014). Understanding the causes and consequences of disease incidence is, therefore, of relevance not only for promoting public health, but also for sustaining and increasing economic well-being.

This dissertation aims to uncover causal relationships that contribute to a better understanding of the determinants shaping public health, of their economic impacts and their policy implications. I apply microeconometric, experimental, and epidemiological methods to two distinct subfields of health economics: medical decision making and the economics of epidemiology.

Chapter 1 focuses on improving risk prediction and treatment choice in medical decision making by exploring a novel algorithm-based risk prediction tool. Advances in technology have been shown to improve clinical practice by enlarging the set of diagnostics and treatment choices (Sampat 2019). Historically, the development of vaccines and antibiotics is among the most prominent examples illustrating how the availability of new technologies boosts population health and shapes social and economic development (e.g. Bhalotra and Venkataramani 2015; Bütikofer and Salvanes 2018). Today, the use of artificial intelligence and evidence-based algorithms in medical decision making is at the center of attention to enhance accuracy in risk prediction and early detection of

diseases (Bayer and Galea 2015; Chen and Asch 2017; Obermeyer and Emanuel 2016). Although predictive algorithms cannot eliminate medical uncertainty, they may support physicians to make better informed treatment choices and thus, improve allocation of scarce health care resources (Chen and Asch 2017). Particularly in the light of an ageing population and increasing health care costs, early and targeted detection of future diseases appears to be an important policy tool to reduce disease burden and costs related to chronic diseases. Chapter 1 provides empirical evidence on the use of a new risk assessment tool in medical decision making that allows physicians to incorporate their clinical experience as an input factor into an algorithm-based decision aid.

Chapter 2 and 3 focus on the economics of infectious diseases. The emergence and rapid transmission of the Covid-19 pandemic underscore the importance of infectious diseases for society and economic well-being. Less salient, but still substantial are the consequences of established infectious diseases on society and the economy. For instance, in the last century infectious diseases have caused more deaths worldwide than all armed conflicts that occurred during the same time period combined (Adda 2016). In modern societies with better access to health care, common infectious diseases such as seasonal flu are an important cause of morbidity and impose substantial costs on society through multiple channels. These include increased health utilization, loss in hours of schooling and work, long-lasting morbidity, and premature deaths (e.g. Almond and Mazumder 2005; Currie and Schwandt 2013; Schwandt 2017). Seasonal influenza, for instance, costs the United States approximately $16 billion and Germany €800 million each year (Lambert and Fauci 2010; Molinari et al. 2007; Scholz et al. 2019). A better understanding of institutional determinants spurring the spread of infectious diseases and of people's behavioral responses to the risk of infection, is crucial in informing policy to enact targeted containment measures and thus, minimize costs to society. Chapter 2 and 3 jointly strive to gain insight into the interaction between infectious diseases, institutional organizations, and behavioral responses. Chapter 2 investigates the effect of early child care on the spread of influenza. While there is much evidence on a link between increased infection rates and child care attendance (e.g. Ball et al. 2000; Ball et al. 2002; Côté et al. 2010), little is known about a causal relationship. As child care attendance has substantially increased over the last two decades (OECD 2020), it is all the more relevant to understand whether child care facilities are a hotspot for disease propagation or not. The chapter further provides insights on the impact of public policies such as mandatory vaccination before entry into child care and closing child care facilities during local outbreaks on disease transmission patterns. Chapter 3 focuses on the role of public information about local and unexpected outbreaks of Covid-19 as an important policy tool to mitigate the spread of the virus. Economists have long been interested in the contribution of information to human decision making (Akerlof 1970; Arrow 1963; Pauly 1968; Spence 1973; Stigler 1961) and more recently, started to explore how people respond to information about health risks (e.g. Banerjee et al. 2020; Dupas 2011; Kim et al. 2019; Oster 2017; Prina and Royer 2014). Yet it remains unclear whether individuals undertake costly behaviors with corresponding health benefits in response to health information or not. The Covid-19 pandemic is one setting, where behavioral responses have been particularly impor-

tant in limiting the spread of the virus and, hence, the costs of the pandemic. Chapter 3 provides evidence on people's intrinsic behavioral responses caused by public information on changes in the risk of infection.

The remainder of this section provides an overview of each chapter of this dissertation. Each chapter is self-contained and can be read independently. A consolidated bibliography is presented at the end of the dissertation.

CHAPTER 1, which is joint work with Charles F. Manski, Joachim Winter, and Amelie Wuppermann, explores a new decision tool that aims at improving physicians' risk assessment and treatment choice. Physicians often face situations in which they must predict patients' future health outcomes under uncertainty. Precise risk assessment is the basis for informative treatment choice and treatment intensity. To support physicians in their risk estimation, clinical practice guidelines commonly recommend the use of evidence-based risk assessment tools (Goff et al. 2014; Piepoli et al. 2016). These decision aids are based on prediction models derived from clinical data, but typically include only a small set of known risk factors (Conroy 2003; Gail et al. 1989; Goff et al. 2014). In clinical practice, physicians often observe additional risk factors and patient characteristics which may change patient's risk, but are not considered in the existing tool. Physicians may either ignore the additional information and stick to the risk assessment tool or may subjectively include the additional information into risk assessment in an entirely unstructured fashion. Prior literature comparing these options suggests that physicians should rather ignore additional characteristics than attempt to subjectively include them into their risk assessment (Camerer and Johnson 1991; Dawes et al. 1989). Neither of these two options appears to be optimal. Recent work by Manski (2018) provides a theoretical framework using bounded-variation assumptions that allows physicians to combine their clinical experience and their assessment of a patient's additional risk factors with predictions from a decision tool in a structured way. The algorithm, that we call "Personalized Risk Assessment Tool" (PRAT), can be embedded into a decision tool using a sequence of probabilistic questions.

In this chapter, we assess empirically whether and to what extent PRAT improves the accuracy of risk prediction and treatment choices compared to risk assessment (i) without the provision of a decision tool and (ii) under the provision of an existing tool recommended by the current clinical guidelines. To address these questions, we run an online experiment with medical students based on the current European clinical practice guidelines for prevention of cardiovascular diseases. Data from the experiment show that the use of PRAT significantly improves precision in risk assessment compared to a decision aid currently applied in clinical practice. Specifically, we document that the use of PRAT improves accuracy in risk prediction by 37 percentage points compared to risk assessment without using any decision aid and by 8 percentage points compared to using an existing tool. We further provide suggestive evidence that the use of this decision tool enables students to make better treatment choices.

CHAPTER 2 investigates the impact of the staggered expansion of early child care in Germany on the spread of influenza. Recent data from contact tracing show that educational institutions such as universities, schools, and child care facilities are among the most common locations where airborne diseases are spread (Robert Koch Institute 2012, 2013). Simultaneously, the number of children attending formal child care has increased substantially in many OECD countries over the last two decades (OECD 2020). If child care facilities indeed propel the spread of infectious diseases, the expansion of early child care imposes unintended costs on society. Work in medical literature shows that child care attendance is positively related to infection rates (e.g Ball et al. 2000; Ball et al. 2002; Côté et al. 2010). Yet empirical evidence on a causal impact of child care on the spread of infectious diseases is scarce.

This chapter attempts to close that gap by providing causal evidence on the transmission of influenza as a consequence of the expansion of early child care – care offered to children under the age of 3 – in Germany. To identify these effects, I leverage the staggered roll-out of early child care provision in Germany between 2005 and 2016. I further exploit detailed, high-frequency data on the incidence of influenza that allow me to uncover age-specific disease transmission patterns over time and space. The analysis proceeds in three steps: First, I provide reduced-form evidence on an economically and statistically significant effect of child care on infection rates using a difference-in-differences strategy. Second, I extend a semi-parametric model of disease diffusion – the Susceptible-Infected-Resistant Model (SIR) – that builds on epidemiological work. The model allows the estimation of age-specific transmission rates and the identification of policy-relevant margins of heterogeneity. Due to the dynamic structure of the SIR model, a non-classical measurement error in the infection rates may result in biased estimates (Adda 2016). To address this endogeneity concern, I create a novel instrument based on lagged temperature. The results of the SIR model document that transmission rates between children aged 0 to 2 and children aged 3 to 6 significantly increase in response to a rise in early child care coverage rates. I further find profound differences between urban and rural areas: While the expansion in child care significantly increases transmission in urban counties, the data do not provide evidence for a significant effect in rural areas. Finally, I evaluate the effect of two counterfactual policy interventions that aim at limiting the spread of infectious diseases in child care facilities: mandatory vaccination before entry into child care and the closure of child care during local outbreaks. I find that both policies significantly reduce infection rates. While the policies have positive spill-over effects on adults, they mostly benefit children. Evaluating the economic costs and benefits of mandatory vaccination policies shows that the policy intervention would be cost-effective reducing net annual costs by about 20% relative to the status quo.

CHAPTER 3, which is joint work with Pavel Obraztcov, Gregory Veramendi, and Joachim Winter, studies the role of public information about unexpected local outbreaks of Covid-19 in mitigating the spread of the virus. Covid-19 is an *overdispersed* pathogen, where a small fraction of individuals is responsible for a large fraction of the transmission. Epidemiological studies have shown

that only 10-20% of individuals are responsible for between 80-90% of transmission clusters often called "super-spreader" events (see e.g. Baggett et al. 2020; Endo et al. 2020; Hamner et al. 2020; James et al. 2020; Lemieux et al. 2020; Majra et al. 2021; Riou and Althaus 2020). These features imply that the local risk level can change dramatically when a transmission cluster occurs. In this case, having quick, credible, and localized information about outbreaks can allow individuals to change their behavior with respect to the local state of the epidemic and be an important part of the mitigation of the outbreak. Yet it remains unclear to what extent individuals react to information on health risks. Some studies argue that people's behavior is sensitive to information on health risk (Chan et al. 2016; Oster 2017; Philipson 2000), while others find that individuals appear reluctant to undertake costly behaviors with corresponding health benefits in response to health information (Cawley et al. 2020; Dupas 2011; Kim et al. 2019; Oster 2018; Prina and Royer 2014).

In this chapter, we combine high-frequency data on the incidence of Covid-19 and on mobility patterns with facts about the incubation period of Covid-19 and reporting time in Germany to isolate the role of public information on positive cases from other possible confounding explanations. We first develop a simple epidemiological model that allows us to identify unexpected local outbreaks by comparing the observed number of cases to the expected number. Based on the model, we find 259 outbreaks at the county level that are distributed relatively evenly across the counties in Germany and across time. Using an event study design on local unexpected outbreaks, we find that mobility significantly decreases by about 2 to 3% in response to public information about the outbreak. In contrast, private knowledge about people falling sick does not appear to cause a change in behavior. The effect is driven by a voluntary change in behavior, as controlling for non-pharmaceutical interventions makes little differences to our estimates. There are important heterogeneities in the behavioral responses. Responses are stronger in counties with high population density, with more hotel beds per capita, and with a higher share of college educated. These findings are consistent with behavioral changes depending on the relative risk and costs of changing mobility. Overall, the findings underscore the importance of public information as a policy tool for mitigating public health risks.

Chapter 1

Improving Risk Assessment and Treatment Choice in Medical Decision Making*

Abstract

Physicians often face situations in which they must predict patients' health outcomes under uncertainty. Existing decision tools are based on prediction models derived from clinical data, but typically include only a small set of known risk factors. In clinical practice, physicians often observe additional risk factors and patient characteristics which may improve their risk assessment. We explore a new risk assessment tool that allows physicians to combine their clinical experience and their assessment of a patient's risk factors with the predictions from a decision tool in a structured way. Data from an online experiment with medical students show that this decision tool significantly improves precision in risk assessment compared to a decision aid currently applied in clinical practice. We further provide suggestive evidence that the use of this decision tool enables students to make better treatment choices.

*This chapter is based on joint work with Charles F. Manski, Joachim Winter, and Amelie Wuppermann.

1.1 Introduction

Physicians often face situations in which they must predict patients' health outcomes under uncertainty. A patient, for example, may ask the physician: "What is the likelihood that I develop a specific disease in the future? What is the chance that the treatment has severe side effects? Or what is the probability that I survive the next ten years given a cancer diagnosis?" In these scenarios, risk assessment is crucial for treatment choice and treatment intensity. Commonly, physicians seek to optimize patient care by predicting risk conditional on *all observed* patient characteristics. However, clinical practice guidelines may recommend the use of an evidence-based risk assessment tool that predicts health outcomes conditional on *just a subset* of the patient's characteristics. How should the physician proceed with the risk assessment?

To date, the physician has two polar options: First, he[1] may ignore the additional information and stick to the risk assessment tool. Second, he may subjectively include the additional information into risk assessment in an entirely unstructured fashion. A physician using the second option acts as a Bayesian. If the physician has rational expectations, i.e. the physician makes accurate probabilistic predictions conditional on the patient's characteristics, the latter option performs at least as well as an evidence-based prediction (Manski 2018). However, a strand of psychological literature concludes that physicians do not have rational expectations.[2] These studies find that predictions based on clinical judgment perform worse than ones made with risk assessment tools using the same patient characteristics. The gap in performance remains even when physicians have an informational edge, in the sense that they observe additional predictive attributes that are not included in the decision aid, as physicians tend to put too much emphasis on the additional information relative to the information included in an existing decision tool (see e.g. Dawes et al. 1989, Camerer and Johnson 1991). Hence, Dawes et al. (1989) suggest that physicians should rather ignore the additional characteristics than attempt to subjectively include them into their risk assessment.

None of these two options appears to be optimal. Recent work by Manski (2018) proposes a decision-theoretic framework using bounded-variation assumptions as a substantial middle ground between these two polar options. In contrast to existing medical decision tools, the approach suggested by Manski (2018) allows physicians to incorporate their clinical experience in a structured way, as it serves as an input of an algorithm that predicts the patient's risk. The algorithm, that we call "Personalized Risk Assessment Tool" (i.e. PRAT), can be embedded in a decision tool using a sequence of probabilistic questions.

In this paper, we assess empirically whether and to what extent PRAT improves accuracy of risk prediction and treatment choices compared to risk assessment (i) without the provision of a decision tool and (ii) under the provision of an existing tool recommended by the current clinical

[1] We use only masculine gender for simplification and better readability. It, however, applies likewise to feminine gender.

[2] Dawes et al. (1989) summarize this literature well.

guidelines. To address these questions, we run an online experiment with medical students[3] based on the current European clinical practice guidelines for prevention of cardiovascular diseases.[4]

In the experiment, students assess the risk of cardiovascular mortality of five hypothetical patients (henceforth: vignettes) and make treatment choices. The experiment comprises three treatment groups that differ with respect to the availability and type of the provided decision tool. Students in treatment group 1 (T1) obtain no decision tool, students in treatment group 2 (T2) are provided with a tool currently used in clinical practice called SCORE, that predicts 10-year cardiovascular mortality conditional on five risk factors, i.e. age, gender, smoking habits, systolic blood pressure, and cholesterol level,[5] and students in treatment group 3 (T3) obtain PRAT in addition to SCORE.

The experimental procedure is as follows. Students first learn about the respective decision tool and then are asked to assess the 10-year risk of a fatal cardiovascular event of five vignettes conditional on a set of well-known risk factors for cardiovascular diseases. The set of risk factors includes the *five characteristics* considered in SCORE, namely age, gender, smoking habits, systolic blood pressure and cholesterol level, and *one additional factor* that is in our application obesity. Obesity – defined as a body-mass-index (BMI) equal to or greater than 30kg/m^2 – is well documented to modify the risk of cardiovascular diseases (van Gaal et al. 2006). In T3, we elicit students' beliefs on the vignette's minimum and maximum risk of cardiovascular mortality conditional on vignette's risk factors before they finalize risk estimations. Students' expectations serve as input factor into the risk estimation of PRAT. Since risk assessment is not an objective in itself but the basis for treatment choice and treatment intensity, we include a second stage in the experiment to investigate whether the provision of PRAT alters treatment choice. To this end, we ask students to state whether they would recommend the vignette to change its lifestyle, e.g. diet habits or physical activity, prescribe the intake of medication, and/or refer the vignette to a cardiologist.

To construct the vignettes, we use individual-level data gathered by the Robert Koch Institute that cover a representative sample of the adult population aged between 17 and 79 living in Germany. The data consist of two waves, a baseline wave including information on a broad set of health measures and a mortality follow-up. The data allow us to estimate prediction models for cardiovascular mortality and then, employ the models to predict the vignettes' objective risk of cardiovascular mortality. In doing so, we can compare subjective risk predictions elicited in the online

[3] We recruited 65 students from two German universities (University of Munich and University of Halle-Wittenberg) between March 2020 and May 2020. Medical students enrolled at these universities were invited to participate in the online survey via the official e-mail distributor of the respective medical faculty. The programming of the experiment was conducted by CentERdata, University of Tilburg.

[4] Cardiovascular diseases subsume a group of disorders of heart and blood vessels, such as high blood pressure, heart attacks, strokes, as well as peripheral vascular diseases.

[5] The use of SCORE is recommended by the latest European clinical guidelines (Piepoli et al. 2016). SCORE exemplifies a common clinical quandary, as it considers only a small set of observable risk factors for cardiovascular diseases. If the physician observes additional risk modifying factors, such as a family history of premature cardiovascular death or a diagnosis of diabetes, the clinical guidelines recommend to adjust risk predictions. There is, however, no guidance how risk adjustment should actually be realized (Piepoli et al. 2016).

experiment with objective risk measures and thus, evaluate student's accuracy in risk prediction conditional on the provided decision tool.

Further, to elicit students' beliefs on vignettes' risk, we use an elicitation procedure proposed by Giustinelli et al. (2019). This approach allows us to distinguish between students holding precise beliefs (i.e. exact probabilities) and ones holding imprecise beliefs (i.e. a range of probabilities) about the vignette's cardiovascular mortality.[6]

We find that medical students substantially overestimate objective risk without the support of a decision tool. Providing students with SCORE helps them to make significantly more precise risk predictions compared to risk estimation without the use of a decision tool. The provision of PRAT in addition to SCORE further significantly improves accuracy of risk prediction. More specifically, average deviation of subjective risk estimation from objective risk without decision tool is 37 percentage points, with SCORE 8 percentage points and with PRAT 4 percentage points. We show that the improvement in risk estimation achieved by the provision of PRAT can be explained by at least two mechanisms: First, PRAT helps students to bound the vignette's maximum risk significantly better than using only SCORE. Second, PRAT supports students holding imprecise beliefs on the risk of cardiovascular mortality to state tightened probability intervals. This suggests that students feel more certain about reporting expectations of the patient's cardiovascular mortality when using PRAT. Turning to the effect of decision tools on treatment choice, we find that the provision of SCORE and PRAT support students to make treatment choices better aligned with treatment recommendations made by current clinical practice guidelines.

This paper contributes to several strands of literature. First, our findings relate to a growing strand of literature that investigates whether algorithmic predictions lead to efficiency gains in human decision making. These studies examine the potential of algorithms to mitigate errors and systematic biases in human judgment across a variety of domains.[7] Research in that field can be broadly divided into two subcategories: First, a body of literature that addresses a "man versus machine" question by comparing algorithmic predictions with human judgment. Many algorithms have been shown to outperform existing prediction routines across a variety of different settings (e.g. Chen and Asch 2017; Dawes et al. 1989; Kleinberg et al. 2018; Obermeyer and Emanuel 2016; Rose 2018). Second, a few studies evaluate the performance of risk assessment tools when they are placed in human hands. These studies view prediction aids as a supplement to human discretion rather than as a replacement of human judgment and allow for human discretion. In this setting, the performance of the tool does not only depend on its algorithm, but also on how the

[6]The procedure proposed by Giustinelli et al. (2019) starts by asking subjects to report precise probabilities as single numbers between 0 and 100 percent and then, uses two follow-up questions to learn about the nature of people's beliefs. The first follow-up question asks whether the stated probability was intended to be an exact number or was rounded/approximated. When the response is rounded/approximated, subjects are asked to report in a second follow-up question an exact precise probability or an imprecise probability, expressed as an interval.

[7]Hoffman et al. (2018) investigate the use of prediction aids in the hiring process, Kleinberg et al. (2018) and Stevenson and Doleac (2019) in juridical decisions on sentencing and Chen and Asch (2017) and Obermeyer and Emanuel (2016) in medical decision making.

decision maker applies the aid in practice. Stevenson and Doleac (2019), for example, show that the performance of a prediction tool may vary substantially, whether it is evaluated solely on the basis of its algorithm or based on the decision maker's use of the decision aid.[8] This paper adds to previous literature in studying a tool that not only allows for human discretion, but also considers human beliefs as one input factor of the algorithm.

Second, the paper adds to a literature that studies the relationship between survey responses on subjective expectations and individual's perceived uncertainty about future events.[9] While most economic research maintains the assumption that survey respondents hold precise subjective beliefs on uncertain events, a small but growing literature on subjective expectations attempts to understand the nature of people's beliefs underlying what they report in a survey (e.g. Giustinelli et al. 2019; Manski and Molinari 2010). Previous literature finds that respondents, who perceive higher uncertainty about future events, tend to round to focal values, such as 0%, 50% or 100% (Manski and Molinari 2010) and/or compress towards 50:50 (Enke and Graeber 2021). Our survey design allows us to distinguish between subjects with precise expectations and ones with imprecise expectations on cardiovascular mortality risk uncovering that indeed about one third of the students holds imprecise beliefs.

Third, the paper complements a number of experimental studies in health economics that use vignettes to study medical decision making (e.g. Brosig-Koch et al. 2019; Cutler et al. 2019; Hoffmann et al. 2014).[10] Cutler et al. (2019), for example, use vignettes to study whether physicians' beliefs on end-of-life care explain regional variation in health care spending. Our paper adds a new layer to previous studies by constructing vignettes based on real patient data. This feature allows us not only to compare subjects' response behavior across treatment status in the experiment, but also to compare subjective risk measures to objective ones. The approach is adopted from studies evaluating individuals' subjective health perception compared to objective risk levels (e.g. Khwaja et al. 2009; Khwaja et al. 2007; Lundborg and Lindgren 2004; Winter and Wuppermann 2014), but is so far not used to evaluate physicians' risk estimation.

Finally, the algorithm underlying PRAT can easily be applied to other diseases than cardiovascular diseases. An accurate prediction of cardiovascular diseases is a major challenge of society, as cardiovascular diseases are a leading cause of morbidity and mortality in Western societies and also cause considerable economic burden to society (Piepoli et al. 2016). It is well-established that

[8] In their study, Stevenson and Doleac (2019) demonstrate that a prediction tool, that outperforms juridical decision making, if judges had completely complied, does not reduce crime recidivism in practice. The authors conclude that their findings can be explained by judicial discretion in its use.

[9] Hurd (2009) and Manski (2004, 2018) summarize the measurement and analysis of subjective expectations.

[10] Vignettes are a validated and often used tool to evaluate treatment choice in clinical practice (Peabody et al. 2000; Peabody et al. 2004). The gold standard to evaluate medical decision making is the use of standardized patients where trained actors consult physicians and record their performance. This procedure is, however, very expensive and, thus, it is commonly only used in very small samples. In contrast, vignettes (i.e. written case simulations) can be easily administered, are less costly and can be applied to large sample sizes. Peabody et al. (2000) and Peabody et al. (2004), comparing the use of standardized patients to the use of vignettes, conclude that the use of vignettes is an accurate measure of clinical practice.

cardiovascular diseases can be prevented by early interventions, such as a change in lifestyle and/or medication. Hence, improving risk prediction of cardiovascular diseases is of great importance to foster population health and also to reduce economic burden.

The paper proceeds as follows. Section 1.2 describes the theoretical considerations that build on the decision-theoretic model of Manski (2018). Section 1.3 presents the design of the experiment. In Section 1.4, we outline the empirical strategy and describe the sample of medical students. Section 1.5 presents the results of the experiment. In Section 1.6, we discuss the experimental design and the results. Section 1.7 concludes.

1.2 Theoretical Considerations based on Manski (2018)

In this section, we outline the theoretical framework of the risk assessment tool PRAT based on recent work by Manski (2018). Section 1.2.1 and Section 1.2.2 summarize the decision-theoretic considerations of Manski (2018). More specifically, Section 1.2.1 describes the identification problem that arises when a physician seeks to optimize patient care by predicting a patient's health outcome y, such as disease development or life span, conditional on the observed patient's characteristics x and w. The characteristics x and w are known to be informative predictors of the patient's health outcome y. An evidence-based decision tool exists that predicts the patient's health outcome y conditional on a *subset* of the observed covariates x, but not conditional on *all* observed characteristics (x,w). In other words, the decision tool predicts $P(y|x)$, but not $P(y|x, w)$. We discuss partial identification of $P(y|x, w)$ given knowledge of $P(y|x)$ and $P(w|x)$ without further structural assumptions and compare it to a set-up with structural assumptions that embody some a priori knowledge of $P(y|x, w)$. Sufficiently strong assumptions may point identify the risk conditional on all observed characteristics, but have low credibility. Hence, a central issue is the tension between the strength of identifying power and the credibility of assumptions. More specifically, stronger assumptions have more identifying power, but less credibility (Manski 2019).[11] Manski (2018) proposes in his work the use of bounded-variation assumptions as a middle course between imposing assumptions that yield point identification and making no structural assumptions. Identification based on bounded-variation assumptions allows tightened identification compared to an approach without any structural assumptions and imposes weaker and thus, more credible assumptions than ones that point identify risk. Section 1.2.2 summarizes briefly the decision-theoretic framework of using bounded-variation assumptions proposed by Manski (2018). Section 1.2.3 describes how the theoretical framework can be used as a risk assessment tool that physicians may easily apply in clinical practice.

[11] The tension between the strength of assumptions and their credibility alludes to "The Law of Decreasing Credibility" that is, "the credibility of inference decreases with the strength of the assumptions maintained" (Manski 2019, p.54).

1.2.1 Risk Assessment without Structural Assumptions and with Strong Structural Assumptions

Suppose a physician wants to predict a patient's health outcome y conditional on the observed patient's characteristics (x=k, w=j). An evidence-based risk assessment tool exists that predicts the patient's health outcome y conditional on a set of characteristics x, but does not consider one additional risk factor w.

The Law of Total Probability relates $P(y|x = k)$ with $P(y|x = k, w = j)$ and $P(y|x = k, w \neq j)$, as specified in Equation 1.1:

$$P(y|x = k) = P(w = j|x = k)P(y|x = k, w = j) + P(w \neq j|x = k)P(y|x = k, w \neq j) \tag{1.1}$$

Knowledge of $P(y|x = k)$ alone reveals nothing about $P(y|x = k, w = j)$. If $P(w = j|x = k)$ is zero, any distribution $P(y|x = k, w = j)$ satisfies Equation 1.1. If $P(w = j|x = k)$ is positive, partial conclusions about $P(y|x = k, w = j)$ may be drawn if one knows $P(w = j|x = k)$. It is plausible to assume that surveys and administrative data exist that allow to derive the distribution $P(w = j|x = k)$ for clinical observable characteristics (x, w).

Hence, the identification problem we address is inference on $P(y|x = k, w = j)$ given knowledge of $P(y|x = k)$ and $P(w = j|x = k)$. We make the following two assumptions: First, we assume that x and w take values in finite spaces X and W and y takes values in the binary set Y $\{0, 1\}$. Put differently, the health outcome is binary ($y = 1$ patient dies due to a specific disease, $y = 0$ otherwise). Second, we assume that the existing, evidence-based risk assessment tool is accurate, in the sense that it predicts $P(y|x)$ correctly. This assumption simplifies the analysis and is commonly maintained by physicians in clinical practice (Manski 2018).[12]

Without imposing any structural assumptions the identification region for $P(y = 1|x = k, w = j)$ can be estimated in the following way: We first solve Equation 1.1 for $P(y = 1|x = k, w = j)$ and obtain Equation 1.2.

$$P(y = 1|x = k, w = j) = \left[\frac{P(y = 1|x = k) - P(y = 1|x = k, w \neq j)P(w \neq j|x = k)}{P(w = j|x = k)}\right] \tag{1.2}$$

[12]Existing tools may indeed not be fully accurate. One concern refers to the methodology that is used to calculate risk assessment tools in the first place: Commonly, risk assessment tools are estimated using prospective studies in that some individuals are treated, while others are not and no information on actual treatment or health behavior is available. More precisely, the outcome of interest may measure future health status unconditional on future treatment status. With historical patient outcomes, there is, however, no way to avoid that we estimate averages across historical treatment choices as made in the population we study. This is a left-out variables problem, and regression with a left-out variable can be thought of as averaging.

In a subsequent step, we set $P(y = 1|x = k, w \neq k)$ equal to 1 to identify the lower bound of $P(y = 1|x = k, w = j)$. Analogously, we set $P(y = 1|x = k, w \neq k)$ equal to 0 to specify the upper bound of $P(y = 1|x = k, w = j)$. Equation 1.3 defines the identification region for $P(y = 1|x = k, w = j)$:

$$P(y = 1|x = k, w = j) \in [0, 1] \cap$$
$$\left[\frac{P(y = 1|x = k) - P(w \neq j|x = k)}{P(w = j|x = k)}, \frac{P(y = 1|x = k)}{P(w = j|x = k)}\right] \tag{1.3}$$

Equation 1.3 identifies risk based on knowledge of $P(y|x = k)$ and $P(w = j|x = k)$. However, without imposing any structural assumptions, the predicted interval may have a large interval width and thus, drawing informative conclusion may be hampered. To increase identifying power, the literature discusses two approaches that impose structural assumptions strong enough to yield point identification: One assumes the existence of an instrumental variable and the other assumes a parametric model for $P(y|x, w)$ (see Manski 2018, Section 2.2 for more details on risk assessment with strong structural assumptions). These approaches impose rather strong assumptions implying that physicians have knowledge of the entire distribution of $P(y|x = k, w = j)$. This, however, may not be plausible and thus, conclusions drawn from these approaches may suffer from low credibility.

Hence, summing up, neither identification without imposing structural assumptions nor one with strong structural assumptions may comfort a physician, that wants to optimize patient care by predicting patient's risk, as conclusions drawn from these approaches are either not very informative or have low credibility.

1.2.2 Risk Assessment using Bounded-Variation Assumptions

Manski (2018) proposes an alternative approach that partially identifies $P(y|x, w)$ by combining knowledge of $P(y|x)$ and $P(w|x)$ with weaker structural assumptions than the ones that yield point identification of $P(y|x, w)$. The central assumption in this approach is that physicians hold beliefs about a patient's minimum and maximum health risk y conditional on (x, w) due to their clinical expertise. Hence, Manski's approach is not having physicians conjecture the entire distribution of $P(y|x, w)$, but rather to place bounds on features of $P(y|x, w)$. Intuitively, it means that a physician being asked to pin down a patient's health risk conditional on a set of observed characteristics may plausibly be able to state that the patient's risk falls into a range, e.g. say between 20% and 40% or above 80%. The smaller the reported interval width, the more certain the physician is about his risk assessment. A physician that has no a priori knowledge about a

patient's risk may state that the risk is between 0% and 100%, while a physician being certain about a patient's risk, may choose a point estimate.

Formally speaking, a physician may assume that $P(y = 1|x = k, w = j)$ or $P(y = 1|x = k, w \neq j)$ falls within specific bounds, say $[a(x = k, w \neq j), b(x = k, w \neq j)]$ and $[a(x = k, w = j), b(x = k, w = j)]$, where

$$a(x = k, w \neq j) \leq P(y = 1|x = k, w \neq j) \leq b(x = k, w \neq j) \tag{1.4a}$$

$$a(x = k, w = j) \leq P(y = 1|x = k, w = j) \leq b(x = k, w = j) \tag{1.4b}$$

$P(y = 1|x = k, w \neq j)$ may take all values within the interval specified in Equation 1.4a. Hence, the identification region for $P(y = 1|x = k, w = j)$ is:

$$P(y = 1|x = k, w = j) \in [a(k, j), b(k, j)] \cap$$

$$\left[\frac{P(y = 1|x = k) - b(k, \neq j)P(w \neq j|x = k)}{P(w = j|x = k)}, \frac{P(y = 1|x = k) - a(k, \neq j)P(w \neq j|x = k)}{P(w = j|x = k)}\right] \tag{1.5}$$

Note that bounds on $P(y = 1|x = k, w = j)$ help to identify $P(y = 1|x = k, w \neq j)$ and vice versa.

The advantage of using bounded-variation assumptions is twofold: First, it results in a tightened interval of feasible probabilities compared to solving the Law of Total Probability without any assumptions (compare Equations 1.3 and 1.5). Second, the assumption that physicians are able to place bounds on minimum and maximum risk is substantially weaker and thus, more credible than assumptions strong enough to yield point identification.

Overall, risk assessment using bounded-variation assumptions allows physicians to flexibly limit the magnitudes of risk assessments and the extent to which they alter with patient characteristics. Thus, it enables physicians to express quantitative risk evaluations in a structured way (Manski 2018).

1.2.3 Implementation of PRAT for Cardiovascular Diseases

The decision tool PRAT presents an application of the theoretical framework described in Section 1.2.2. The risk predicted by PRAT rests on three pillars: (i) An evidence-based risk assessment tool exists that predicts $P(y|x)$. (ii) The conditional probability $P(w|x)$ can be estimated using an auxiliary data set. (iii) Physicians hold beliefs on patients' minimum and maximum health risk y conditional on patients' characteristics (x,w). The beliefs, i.e. $a(x = k, w \neq j)$; $b(x = k, w \neq j)$;

$a(x = k, w = j)$; $b(x = k, w = j)$, serve as input factors into the algorithm of PRAT. Equation 1.6a and 1.6b specify the algorithm in our specific setup.

$$P(y = 1|x = k, w = j) \in [a(k, j), b(k, j)] \cap [\text{Lower Bound}, \text{Upper Bound}]$$

$$\text{Lower Bound} = \frac{\overbrace{P(y = 1|x = k)}^{\text{SCORE}} - \overbrace{\boxed{b(k, \neq j)}}^{\substack{\text{physician's} \\ \text{expectations}}} \overbrace{P(w \neq j|x = k)}^{\substack{\text{estimated in a} \\ \text{separate data set}}}}{\underbrace{P(w = j|x = k)}_{\substack{\text{estimated in a} \\ \text{separate data set}}}} \tag{1.6a}$$

$$\text{Upper Bound} = \frac{\overbrace{P(y = 1|x = k)}^{\text{SCORE}} - \overbrace{\boxed{a(k, \neq j)}}^{\substack{\text{physician's} \\ \text{expectations}}} \overbrace{P(w \neq j|x = k)}^{\substack{\text{estimated in a} \\ \text{separate data set}}}}{\underbrace{P(w = j|x = k)}_{\substack{\text{estimated in a} \\ \text{separate data set}}}} \tag{1.6b}$$

We apply PRAT to predict a patient's 10-year risk of cardiovascular mortality. A number of risk assessment tools exists that predict fatal/non-fatal cardiovascular events conditional on a pre-defined, typically small number of risk factors.[13] The difference between existing tools and PRAT is that PRAT allows physicians to incorporate their expectations about the patient's risk into the algorithm. Hence, PRAT can be regarded as an extension of traditional risk assessment tools, as it combines the risk predicted by an existing tool with physicians' expectations.

To estimate the risk predicted by PRAT we use three distinct components: First, we use the risk assessment tool SCORE that assesses the 10-year risk of cardiovascular mortality conditional on a set of risk factors as one component of the algorithm of PRAT (see Section 1.3.1 for more information on SCORE). It is noteworthy that any other existing risk assessment tool such as PROCAM or Framingham, could be equally used within the framework of PRAT.[14] Second, to

[13] See Appendix Table 3 in Goff et al. (2014) for an overview of existing evidence-based risk assessment tools currently used to assess the risk of cardiac events.

[14] We decided to use SCORE in this study, as it is the risk assessment tool for cardiovascular diseases recommended by the latest European clinical guidelines (Piepoli et al. 2016).

estimate $P(w = j|x = k)$, we use individual-level data on a broad set of health measures, covering a representative sample of the adult population living in Germany (see Section 1.3.2 for more information). Third, PRAT considers physician's beliefs on patient's minimum and maximum risk of cardiovascular mortality conditional on patient's characteristics in the risk prediction. We elicit beliefs on minimum and maximum risk in the online experiment.

From the applicant's perspective, the usage of PRAT is a two-step procedure: First, physicians are asked to report their beliefs on a patient's minimum and maximum health risk. Subsequently, the algorithm of PRAT combines physicians' expectations with the risk predicted by SCORE and with the patient's probability of the additional risk factor w given the set of risk factors x. Second, physicians are informed about the interval predicted by PRAT and thus, may assess a patient's risk conditional on risk factors (x,w).

1.3 Design of the Experiment

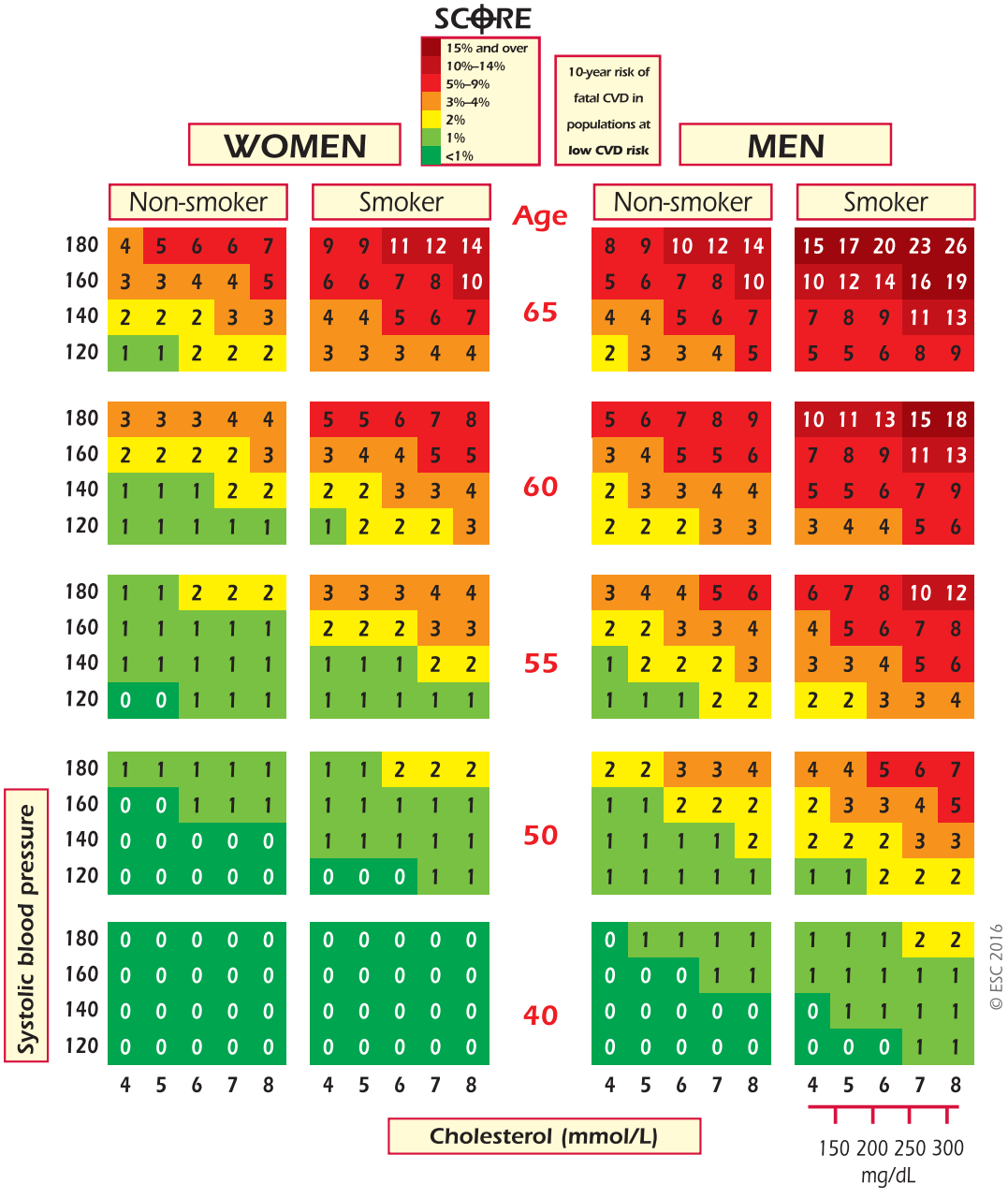
This section summarizes the design of the online experiment with medical students, conducted between March 2020 and May 2020. In Section 1.3.1, we describe the risk assessment tool, SCORE, that assesses the 10-year risk of cardiovascular mortality conditional on five, predefined risk factors. The use of SCORE is recommended by the latest European clinical guidelines (Piepoli et al. 2016) and thus, we benchmark the performance of PRAT against SCORE. Section 1.3.2 outlines the construction of the vignettes using survey data. Section 1.3.3 explains the procedure of the experiment in detail.

1.3.1 Background Information on SCORE

In 2003 the European Society of Cardiology developed a risk scoring system tool that aims to improve risk assessment of cardiovascular diseases, as well as to enable physicians to base treatment choices on evidence-based and personalized risk assessment. The decision tool SCORE (Systematic COronary Risk Evaluation) considers five risk factors, namely age, gender, current smoking status, systolic blood pressure and serum cholesterol level, that are well-known to modify the likelihood of a cardiovascular event (Conroy 2003).[15] SCORE is supposed to be used for primary prevention of cardiovascular diseases, i.e. to assess the risk of apparently healthy individuals, not of individuals with established cardiovascular diseases, chronic kidney disease or any risk mod-

[15]Risk levels of SCORE are estimated from 12 European cohort studies that cover 205,178 individuals. Ten-year risk is calculated using a Weibull proportional hazard model in which age was used as a measure of exposure time to risk rather than as a risk factor. Conroy (2003) estimate risk separately for coronary and non-coronary heart disease, as weights assigned to different risk factors may be different for the two types of cardiovascular diseases. This may, in turn, affect the shape of the lifetime hazard functions. Separate risk charts have been developed for low and high-risk countries in Europe. Countries are defined as "low-risk", if age-adjusted cardiovascular mortality rates for individuals aged 45–74 in 2012 were below 225/100,000 for males and 175/100,000 for females. As Germany is considered as low-risk country, we refer to those estimates in this paper. Finally, note that SCORE was re-calibrated separately for many European countries (e.g. Diederichs et al. 2018; Lindman et al. 2007).

Figure 1.1: SCORE Chart taken from the Clinical Practice Guidelines



Notes: The chart from the European clinical practice guidelines (2016) reports the 10-year risk of fatal cardiovascular diseases based on the following 5 risk factors, i.e. age, gender, smoking behavior, systolic blood pressure, and total cholesterol level (Piepoli et al. 2016, p.2328).

ifying pre-conditions (e.g. cancer diagnosis, diabetes with severe organ damage). Moreover, its recommended application is limited to the age range of 40 to 65 years (Piepoli et al. 2016).

To simplify the use of SCORE in clinical practice, the guidelines provide a chart that illustrates the 10-year risk of a fatal cardiovascular disease event for 400 different combinations of risk factors (Figure 1.1). Hence, a physician may easily read a patient's 10-year risk of cardiovascular mortality conditional on the patient's individual set of risk factors from that chart. For instance, a male 65-year old non-smoker, who has a systolic blood pressure of 160mmHg and a cholesterol level of 6mmol/l, has on average a 7% risk of dying due to a cardiovascular event in the next 10 years.

The guidelines classify the risk into three broad categories which are intended to assist physicians in their treatment choice: Individuals with a SCORE below 5% are classified as low-to-moderate risk persons and physicians should offer lifestyle advice to maintain their low-to-moderate risk status. High risk persons (calculated SCORE $\geq$5% and $<$10%) qualify for intensive lifestyle advice and may be candidates for drug treatment, while for very high risk persons (calculated SCORE $\geq$10%) drug treatment is more frequently required (Piepoli et al. 2016). The clinical guidelines emphasize that the thresholds are not universally applicable, but should be reconsidered for each patient individually.

It is well known that other attributes than those considered in SCORE may modify the probability of a cardiovascular death. Prominent examples for attributes, that may change the risk of a fatal cardiovascular event, but are not used as predictors in SCORE, are the diagnosis of diabetes, obesity, a family history of cardiovascular events, and socio-economic characteristics (see e.g. Clark et al. 2009; Sattelmair et al. 2011; van Gaal et al. 2006). The clinical practice guidelines list these risk factors and recommend physicians to reconsider the risk predicted by SCORE in the presence of any additional risk modifying factors. There is, however, no guidance how physicians should revise patient's risk in the context of additional risk factors (Piepoli et al. 2016).

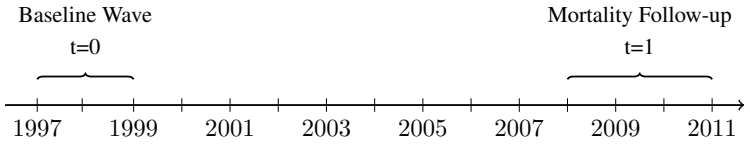
1.3.2 Construction of the Vignettes

In the online experiment, medical students are asked to evaluate the risk of five vignettes. To construct the vignettes, we use individual-level data on socio-demographic characteristics, health measures, and the event of death provided by the Robert Koch Institute (henceforth: RKI study). The RKI study covers a representative sample of the population aged 17 to 79 living in Germany.

The RKI study comprises two waves – a baseline wave and a mortality follow-up (see Figure 3.1). The baseline wave was conducted between October 1997 and March 1999 and it contains information on health behavior, health outcomes and lab parameters. Importantly, the data include information on each of the five risk factors considered in the risk assessment tool SCORE (Thefeld et al. 1999). A follow-up study was conducted about 12 years after the baseline wave, eliciting

information on the event of death. For participants, who died between the two waves, information on date and cause of death was collected (Wolf et al. 2014; Wolf et al. 2012).[16]

Figure 1.2: RKI Study – Timeline of the Baseline Wave and the Mortality Follow-up



Notes: The figure illustrates the timeline of the baseline wave and the mortality follow-up. The baseline wave of the RKI study took place between 1997 and 1999. The mortality follow-up was conducted between 2008 and 2011.

Importantly, the data provide us with information on participants' individual risk factors for cardiovascular diseases and death caused by cardiovascular diseases. Therefore, the data enable us to estimate risk prediction models for cardiovascular mortality. More precisely, we use Probit regressions to calculate the objective 10-year risk of cardiovascular mortality (i) conditional on the risk factors considered in SCORE (i.e. age, gender, blood pressure, cholesterol level and smoking habits) and (ii) conditional on the characteristics included in SCORE *plus* obesity. Subsequently, we calculate the objective risks for five vignettes with different combinations of risk factors using the parameters estimated in the risk prediction models (see Appendix Section A1 for more details).

To estimate the prediction models, we construct a sample that is in terms of age structure and risk factors as similar as possible to the sample used to estimate SCORE. Hence, in line with Conroy (2003), we exclude participants with prior history of heart attacks from the sample. We further exclude participants if the time span between the interview of the baseline wave and the mortality-follow-up was less than 10 years and participants did not decease within that period.[17] This leaves us with 6,274 participants with complete information on the risk factors considered in the study and on the event of a cardiovascular death (defined by ICD-10 code, I00 to I99). About 1.66% of the sample died due to a cardiovascular disease within 10-years after the first interview in the baseline wave.

Table 1.1 gives an overview of the vignettes. Panel A summarizes the vignettes' characteristics considered in the risk assessment process. Panel B illustrates the information that a physician may read from the current practical guidelines (CPG). It covers the risk predicted by SCORE and

[16]The mortality follow-up includes the information provided on the official death certificates. In Germany, physicians have to report a causal chain of diagnosis leading to death on the certificate. Specifically, physicians have to state (i) the immediate cause of death (*unmittelbar zum Tode führende Krankheit*) (ii) previously diagnosed diseases leading to death (*vorangegangene Ursachen*) (iii) chronic clinical conditions (*Grundleiden*) (iv) other relevant diagnosis (*andere wesentliche Krankheiten*). Typically, physicians state more than one ICD-code in each of the categories. Hence, to define the *main* cause of death we use the international validated system called IRIS. For further information see Federal Institute for Drugs and Medical Devices (2021).

[17]Note that it is necessary to exclude these participants to predict 10-year mortality risk, as we have no information whether participants actually died or not within 10 years.

Table 1.1: Characteristics of the Vignettes

	Vignettes				
	1	**2**	**3**	**4**	**5**
Panel A: Vignettes' Risk Factors					
Gender	male	male	female	male	female
Age	65	65	62	64	56
Blood Pressure [mmHg]	145	165	180	145	180
Cholesterol Level [mmol/l]	6.3	4	5.4	5	5
Smoking Status (yes/no)	yes	yes	yes	no	no
Obese (yes/no)	yes	no	no	yes	yes
Panel B: Information from CPG					
Risk Predicted by SCORE	9%	10%	5%	4%	1%
Risk Category	very high	very high	high	high	low-to-moderate
Panel C: CVD Mortality Risk in RKI Study					
Risk predicted in RKI study $P(y = 1\|x = k)$	8.6%	9.45%	4.95%	4.23%	1.09%
Risk Predicted in RKI Study $P(y = 1\|x = k, w = \text{obese})$	10.15%	10.88%	5.66%	5.01%	1.25%
Risk Predicted in RKI Study $P(y = 1\|x = k, w = \text{not obese})$	8.28%	8.92%	4.49%	3.94%	0.93%
Panel D: Probability of Obesity conditional on Risk Factors in RKI Study					
Risk Predicted in RKI Study $P(w = \text{obese}\|x = k)$	25.08%	32.37%	45.15%	25.07%	44.65%
Risk Predicted in RKI study $P(w = \text{not obese}\|x = k)$	74.92%	67.63%	54.85%	74.93%	55.35%

Notes: The table summarizes the characteristics/risk factors of the five vignettes used in the study. The parameter k denotes the set of risk factors considered in the estimation tool SCORE (age, gender, blood pressure, cholesterol level, and smoking habits).

the risk category provided in the clinical practice guidelines. Panel C presents the 10-year risk of cardiovascular mortality predicted using the RKI study. We estimate risk levels (i) conditional on the five risk factors included in SCORE if obesity is not considered as separate risk factor (ii) conditional on the five risk factors included in SCORE if the vignette is assumed to be obese, and (iii) conditional on the five risk factors included in SCORE if the vignette is assumed to be not obese. The results show that being obese is associated with an increased risk of cardiovascular mortality of roughly 25% compared to being not obese. This indicates that obesity changes relative risks of cardiovascular mortality substantially. Panel D describes the probability of being obese given the characteristics considered in SCORE. The estimated values are calculated using a Probit regression model (see Appendix Section A1 for more details).

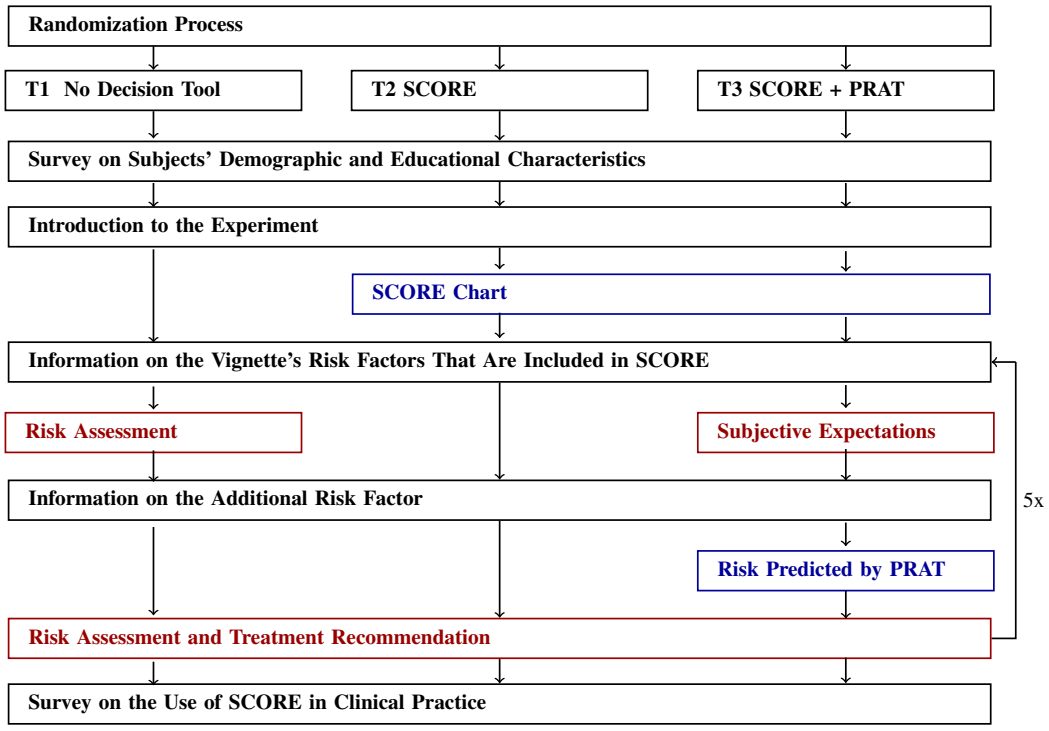
1.3.3 Experimental Procedure

Medical students[18] are randomly drawn into one of three treatment groups that differ with respect to the provided decision tool. Subjects in T1 are asked to assess the vignette's risk without the

[18]The term "subjects" and "respondents" are interchangeable and used throughout the paper to describe medical students that take part in our experiment. Individuals that participated in the RKI study are denoted as "participants".

provision of any decision tool, while subjects in T2 are provided with the SCORE chart (see Figure 1.1). Subjects in T3 obtain the SCORE chart and additionally get informed about the risk predicted by PRAT.

Figure 1.3: Flow Chart - Experimental Procedure



Notes: The figure summarizes the experimental procedure. Cells highlighted in blue display stages of the experiment where subjects obtain information on the risk predicted by SCORE or PRAT. Cells highlighted in red present stages where subjects are asked to assess the vignettes' risk.

Figure 1.3 summarizes the experimental procedure. After subjects are randomly assigned to a treatment group, the survey starts with a questionnaire eliciting information on subjects' socio-demographic, educational, and job characteristics (e.g. age, gender, location of residence, year of studies, and work experience). In the subsequent part of the survey, subjects first assess the vignettes' risk and then, make treatment recommendations. To avoid ordering effects, the sequence of the vignettes presented to the subjects is randomized. The exact procedure of the survey depends on the treatment group and is described below. The survey ends with a short questionnaire asking subjects to report about prior knowledge and use of the risk assessment tool SCORE.

Risk Assessment in T1 and T2

Subjects in T1 do not obtain any decision tool to support their risk assessment. During the experiment they are asked to assess each vignette's risk twice – first, conditional on the five risk factors considered by SCORE and second, conditional on six risk factors, i.e. those included in SCORE plus obesity. Subjects in T2 are introduced to the risk assessment tool SCORE and obtain the SCORE chart to support their decision.[19] Subsequently, subjects in T2 are asked to evaluate each vignette's risk once conditional on six risk factors, i.e. the factors considered by SCORE plus obesity.

To elicit expectations of the risk of cardiovascular mortality, we build on Giustinelli et al. (2019) and devise an elicitation procedure that allows us to distinguish between subjects holding precise beliefs and ones holding imprecise beliefs. To this end, we start by asking subjects about a percent-chance of cardiovascular mortality and then, use two follow-up questions to learn whether subjects hold precise or imprecise beliefs about the risk of cardiovascular mortality. The first follow-up question asks whether the reported probability was intended to be an exact number or was rounded/approximated. When the response is rounded/approximated, subjects are asked to state in a second follow-up question the exact precise probability or an imprecise probability, stated as a range (see Appendix Figure A.1 for a graphical illustration). The initial question asking for a percent-chance of cardiovascular mortality is worded as follows:

> ***Initial Question:*** *What is the 10-year risk that the patient described to you above dies due to a cardiovascular disease? Please choose a number on a scale of 0 to 100.*

This question is equivalent to the standard expectation questions used in a number of surveys (e.g. Health and Retirement Survey) that ask subjects to state precise probabilities as single numbers between 0% and 100%, where 0% means, that there is no chance that the patient will die within the next 10 years and 100% means that the patient will certainly die due to a cardiovascular disease within the subsequent 10 years. To differentiate between subjects with precise probabilistic beliefs and ones with imprecise probabilistic beliefs, subjects are asked to state whether the reported probability was intended to be an exact number or was rounded/approximated. The first follow-up question is worded as follows:

> ***Follow-up Question 1:*** *When you said [X] percent just now, did you mean this as an exact number or were you rounding or approximating?*
> Possible answers: (i) Exact number (ii) Rounding or approximating

Subjects stating that their initial response was exact are directly asked to make a treatment recommendation, while those reporting that they were rounding or approximating are asked a second follow-up question.

[19] To ensure that subjects understand how to read the vignette's risk from the chart, they can only continue with the survey if they answer a test question correctly.

Follow-up Question 2: *Now please try without rounding or approximating your answer. What is the 10-year risk that the patient dies due to a cardiovascular disease? If you are uncertain about the chances, you may state an interval that includes the vignette's risk. For example, you may say something like "less than 9%," "between 13% and 20%" or "greater than 50%."*
Possible answers: (i) Exact number (ii) Risk below [X] (iii) Risk above [X] (iv) Risk between [X] and [Y]

Risk Assessment in T3

Subjects in T3 are provided with two risk assessment tools, namely SCORE and PRAT. At the beginning of the experiment, subjects in T3 are introduced to the risk assessment tool SCORE and obtain the SCORE chart (as subjects in T2). Subsequently, subjects in T3 learn about the vignette's characteristics that are considered in the algorithm of SCORE (i.e. age, gender, systolic blood pressure, cholesterol level, smoking habits) and are asked to state the vignette's minimum and maximum risk of cardiovascular mortality conditional on two counterfactual scenarios; first, if the vignette was obese; second, if the vignette was not obese. Put differently, subjects at that stage of the survey do not know whether the vignette is obese or not, but are asked to report maximum and minimum risk of cardiovascular mortality for both potential scenarios.[20] The wording of the question is as follows:

Elicitation of Subjective Bounds on Risk of Cardiovascular Mortality: *Consider whether and to what extent obesity may affect the risk that the patient described above dies due to a cardiovascular disease in the next 10 years. Presumably, you may not be able to answer the question precisely. However, based on the risk predicted by SCORE you may be able to state an interval, that contains the patient's risk. For example, you may say something like "less than 9%", "between 13% and 20%" or "greater than 50%". Please choose a number on a scale of 0 to 100 to respond to the following statements.*
Please report the ***maximum 10-year risk*** *that the patient described above dies due to any cardiovascular disease, if you have the information that* ***the patient is (not) obese:***
☐ %
Please report the ***minimum 10-year risk*** *that the patient described above dies due to any cardiovascular disease, if you have the information that the* ***patient is (not) obese****:*
☐ %

[20] Note that this is necessary as $P(y|x = k, w = \text{obese})$ is identified by PRAT based on beliefs on $P_{\text{min}}(y|x = k, w = \text{not obese})$ and on $P_{\text{max}}(y|x = k, w = \text{not obese})$ and vice versa.

Subjects' beliefs on minimum and maximum risk are used as input factors into the algorithm of PRAT. It is, further, to note that we impose the assumption that the probability of cardiovascular mortality is monotonously increasing in adverse risk factors. Hence, obesity increases the risk of cardiovascular diseases relative to the risk predicted by SCORE that is calculated by averaging across obese and non-obese individuals. In the same vein, not being obese decreases the risk of cardiovascular diseases relative to the risk calculated based on SCORE. Formally speaking, we impose the assumption that, if the vignette is not obese, the maximum risk of cardiovascular mortality conditional on the risk factors included in SCORE must be equal or lower than the respective risk predicted by SCORE. Vice versa, if the vignette is obese, the minimum risk of cardiovascular mortality conditional on the risk factors covered by SCORE, must be equal or larger than the risk predicted by SCORE. Equation 1.7 specifies the monotonicity assumption:

$$\begin{aligned} P_{min}(x = k, w = \text{not obese}) \leq P_{max}(x = k, \text{not obese}) \\ \leq P_{\text{SCORE}} \leq P_{min}(x = k, w = \text{obese}) \leq P_{max}(x = k, w = \text{obese}) \end{aligned} \tag{1.7}$$

Subjects reporting minimum and maximum risks not in line with the monotonicity assumption, may revise their response three times. If the third attempt does not meet the criteria (Equation 1.7), subjects are directly led to assess the next vignette's risk. Subjects stating bounds, that align with the monotonicity assumption, learn in the next step of the experiment whether the vignette is obese or not. Further, they get informed about the risk predicted by PRAT. The information on the risk predicted by PRAT is provided by the following wording: "The algorithm PRAT considers in its calculations the risk level predicted by SCORE and your estimates on the patient's minimum and maximum risk level. In contrast to SCORE, PRAT takes obesity as additional risk factor into account. PRAT predicts that the patient's 10-year cardiovascular mortality risk is between [Lower Bound]% and [Upper Bound]%".

Hence, subjects in T3 obtain information on the risk predicted by SCORE and PRAT, before they are asked to evaluate the vignette's 10-year risk of cardiovascular mortality. The question is worded as follows:

> ***Risk Assessment using SCORE and PRAT:*** *What is the 10-year risk that the patient described above dies due to a cardiovascular disease? You may either choose an exact number, for example 5% or if you are not certain, you may state an interval that includes the vignette's risk. For example, you may say something like "less than 9%", "between 13% and 20%" or "greater than 50%". Please choose a number on a scale of 0 to 100 to respond to the following statements.*
> Possible answers: (i) Exact number (ii) Risk below [X] (iii) Risk above [X] (iv) Risk between [X] and [Y]

Treatment Recommendations

After assessing the vignette's risk, subjects are asked to make treatment recommendations. More precisely, we ask subjects whether they recommend (i) a lifestyle advice such as a change in diet, increased physical activity (ii) the prescription of blood pressure lowering drugs (i.e. antihypertensive drugs) and/or of cholesterol lowering drugs (i.e. statins), and (iii) a referral to a cardiologist on a yes/no scale (see Appendix A1 for the exact wording).

1.4 Data and Empirical Strategy

1.4.1 Sample Description and Balancing Tests

We provide a description of the sample of medical students and data below.

Sample We invited medical students enrolled at two German universities (University of Munich and University of Halle-Wittenberg) to participate in the online survey via the official e-mail distributor of the respective medical faculty. The experiment was conducted between March 2020 and May 2020. Overall, 239 students started with the survey, 129 of them (53%) completed the first part of the survey, a questionnaire on demographics and clinical experience, and 65 (27%) students finished the complete survey (see Appendix Table A.2 for a detailed overview of the sample structure).

In the analysis, we focus on the sample of students that completed the survey. Hence, our sample comprises 65 medical students: 25 in T1, 24 in T2, and 16 in T3. Table 1.2 provides an overview of subjects' socio-demographic characteristics (Panel A) and clinical experience (Panel B) separately by treatment group. The students participating in the survey are on average 24 years old, roughly 30% of them are male, and attend the 6^{th} semester of the study curriculum of human medicine. About 80% of the students report to reside in Bavaria.[21] To learn about students' clinical experience, we elicit information on work experience in health care prior to their studies, as well as on internships in internal medicine during the first (i.e. preclinical) and second (i.e. clinical) period of their studies (Panel B). In our sample, about 17% of the students worked prior to their studies in health care and about 60% conducted an internship in internal medicine in the preclinical and clinical period. Panel C in Table 1.2 gives information on the survey design. It took subjects on average 40 minutes to complete the survey.

[21] We did not elicit whether a student is enrolled at the University of Munich or Halle-Wittenberg. Thus, we take reported information on the location of residency as a proxy for being a student at the University of Munich or Halle-Wittenberg.

Table 1.2: Sample Description and Balance

	Treat. Group 1		Treat. Group 2		Treat. Group 3		Diff. 1 − 2	Diff. 1 − 3	Diff. 2 − 3
	Mean	S.D.	Mean	S.D.	Mean	SD	p-value	p-value	p-value
Panel A: Sociodemographic Characteristics									
Year of Birth	1996.32	2.32	1994.63	5.36	1995.88	3.22	0.15	0.61	0.41
Gender (male=1)	0.32	0.48	0.17	0.38	0.44	0.51	0.22	0.46	0.06
Semester	6.76	2.91	6.42	3.02	6.06	2.72	0.69	0.45	0.71
Location of Residence (Bavaria=1)	0.80	0.41	0.71	0.46	1.00	0.00	0.47	0.06	0.02
Panel B: Clinical Experience									
Work experience in health care prior to studies (=1)	0.08	0.28	0.25	0.44	0.19	0.40	0.11	0.32	0.65
Preclincal Period - Internship in Internal Medicine (=1)	0.71	0.46	0.52	0.51	0.60	0.51	0.20	0.50	0.65
Clincal Period - Internship in Internal Medicine (=1)	0.68	0.48	0.59	0.50	0.57	0.51	0.54	0.51	0.91
Panel C: Survey Design									
Duration Survey (min)	48.91	77.38	31.06	16.41	38.89	34.86	0.29	0.63	0.35
Observations	25		24		16		49	41	40

Notes: The table summarizes characteristics of the sample of students that completed the survey and tests for balance in observable characteristics between subjects in each of these treatment groups. Columns 1 to 6 report the mean and standard deviation of the set of variables included in the analysis by treatment group. Columns 7 to 9 display the respective p-values of the differences in means between the three treatment groups.

Selective Drop-outs To investigate whether we face selective drop-outs, we run a number of specification tests. To this end, we use two definitions to describe the sample of drop-outs: (i) subjects that started the survey, but stopped the questionnaire at any point before completion (174 subjects, attritors definition 1) (ii) subjects that finished the first part of the survey, a questionnaire on sociodemographic and educational characteristics, and thus, showing general interest in the topic, but left the survey before completion (64 subjects, attritors definition 2).

First, we compare characteristics of students who did not finish the survey with ones that completed it. We find no significant differences in sociodemographic characteristics between drop-outs and the students completing the survey with one exception, namely students living in Bavaria are more likely to finish the experiment compared to students residing outside Bavaria. Interestingly, we find that students with more clinical experience (measured by internships and work experience prior to studies) are more likely to drop out than students with less experience (see Appendix Table A.3). Second, comparing characteristics of drop-outs by treatment group, we find no significant differences by treatment status at a 5% significance level for all characteristics except one. Drop-outs in T1 and T3 are significantly more likely to live in Bavaria than those in T2 (see attritors definition 1, Appendix Table A.4). Using the sample of attritors described by definition 2, there are no significant differences between drop-outs by treatment group (see Appendix Table A.5).

Balance Tests Table 1.2 shows that balance across intervention status was achieved for all observable characteristics other than location of residence. We find that the fraction of respondents living in Bavaria is significantly smaller in T2 than in T3 (i.e. p-value<0.05, see Table 1.2, Columns 7 to 9).

1.4.2 Empirical Strategy

To identify the causal impact of the provision of PRAT on the accuracy in risk prediction and treatment choices compared to decision making (i) without the provision of any decision tool and (ii) under the provision of SCORE, we estimate models of the following kind:

$$Y_{i,j} = \alpha + \beta_{T1} T1_i + \beta_{T2} T2_i + X_i' \gamma + \delta_j + \epsilon_{i,j} \tag{1.8}$$

where the parameter j denotes the vignette scenario (i.e. j = 1,., 5) and i refers to the subject responding the survey. The binary variables $T1_{\mathrm{i}}$ and $T2_{\mathrm{i}}$ define the treatment status and are equal to 1 if subjects are provided with no decision tool (T1) or with SCORE (T2). Treatment effects are measured relative to being provided with SCORE plus PRAT (T3, baseline category). In some model specifications, we control for a set of covariates, X_{i}, that describes the subject's sociodemographic characteristics, as well as his clinical experience. To account for potential ordering effects, we vary the sequence of the vignettes at random. Further, the parameter δ_j controls for ordering fixed effects. Standard errors are clustered at the subject's level. As the treatment status is randomly assigned, $T1_{\mathrm{i}}$ and $T2_{\mathrm{i}}$ are by construction orthogonal to X_{i} and $\epsilon_{\mathrm{i,j}}$. Hence, β_{T1} and β_{T2} identify the average treatment effect of using no decision tool (T1) or SCORE (T2) relative to being provided with SCORE plus PRAT (T3) even without controlling for X_{i}.

To evaluate risk prediction, we use as outcome variable $Y_{i,j}$, the difference between subjective and objective risk measures as a proxy for accuracy in risk assessment. Throughout the paper, we maintain the following two assumptions: First, responses that subjects give after probing express their "true" expectations, be they precise or imprecise.[22] Second, our measure of objective risk is the "true" underlying risk of a vignette. Hence, the smaller the absolute difference between subjective and objective risk measure, the more accurate is the subjective risk assessment. Following the first assumption, we use post-probe responses to evaluate subjective risk assessment. As post-probe responses include precise probabilities, expressed as exact numbers, and imprecise probabilities, stated as range, we create three measures of subjective risks: (i) the midpoint of a stated range, (ii) the upper bound of a stated interval, and (iii) the lower bound of a stated interval. If precise probabilities are reported, all three measures fall together. More precisely, to evaluate

[22] We follow Giustinelli et al. (2019) in assuming that a subject's initial response is a swift and likely error-ridden measure of his underlying belief, while a subject's post-probe response is an error-free measure stated after some reflection.

accuracy in risk assessment, we consider three outcome variables that are defined as the absolute difference between each of the three described subjective risk measures and the objective risk. This procedure allows us to disentangle whether the provision of decision tools enables subjects to identify more precisely average risk (midpoints), maximum risk (upper bound), and/or minimum risk (lower bound). To assess the impact of the availability and the type of a decision tool on treatment choices, we run Probit regression models with treatment choice as binary outcome variables.

1.5 Results

This section summarizes the empirical findings. In Section 1.5.1, we discuss the impact of using PRAT on the accuracy in risk assessment compared to risk prediction without a decision aid and with using SCORE. In Section 1.5.2, we analyze whether PRAT impacts treatment choice. Section 1.5.3 discusses whether the availability and the type of decision tool affects subjects' nature of reporting precise or imprecise probabilities in surveys. In Section 1.5.4, we compare the performance of decision tools without human discretion to risk assessment based on clinical judgment.

1.5.1 Accuracy in Risk Prediction

To begin, we study the causal impact of the provision of PRAT on the accuracy in risk prediction of cardiovascular mortality conditional on six risk factors (i.e. age, gender, systolic blood pressure, cholesterol, smoking status, obesity) compared to decision making (i) without the provision of any decision tool and (ii) under the provision of SCORE,

Descriptive Statistics Figure 1.4 illustrates the empirical distribution of the difference between subjective and objective risk measures, by treatment group. We find that medical students without any decision tool (T1) substantially overestimate objective risk. Expectations on the vignette's risk deviate by plus 85pp to minus 5pp from the objective risk measure showing a roughly uniform distribution within that range. Subjects being provided with SCORE (T2) state expectations that deviate from the objective risk on a comparable range. However, in stark contrast to the findings in T1, the empirical distribution of the difference between subjective and objective risk in T2 is compressed around zero. Providing subjects with PRAT in addition to SCORE (T3) enables subjects to further close the gap between subjective and objective risk. Table 1.3 presents the summary statistics of the difference between subjective and objective risk measures.

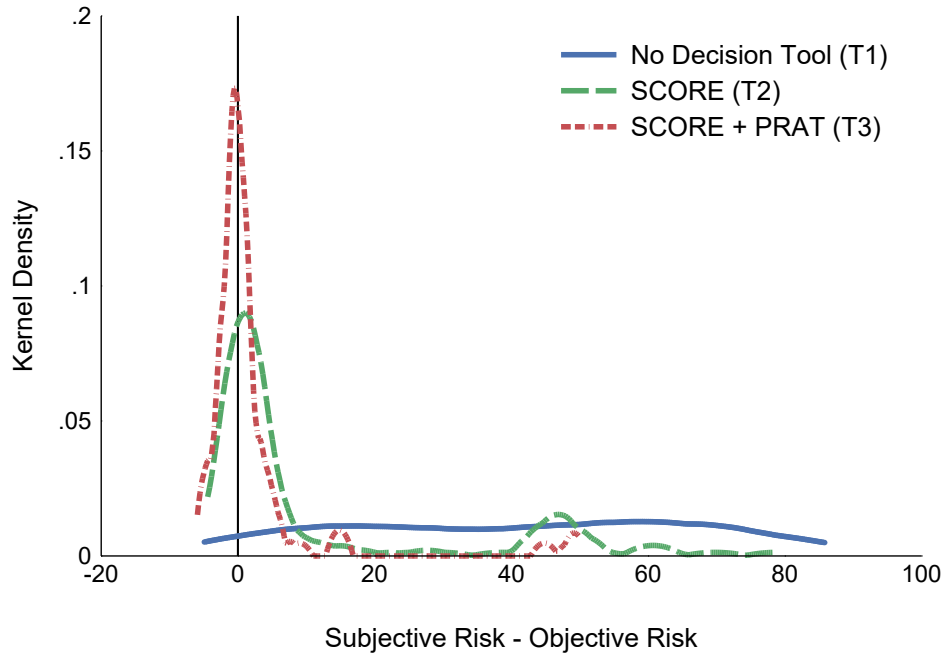
Overall, the descriptive statistics indicate that the provision of the decision tools SCORE and PRAT helps medical students to significantly improve risk assessment.

Table 1.3: Summary Statistics: Difference between Subj. Risk and Obj. Risk

	No Decision Tool (T1)	**SCORE (T2)**	**SCORE + PRAT (T3)**
Mean	41.29	11.27	2.18
S.D.	26.01	19.96	9.89
Median	43.75	1.75	0.11
Range	(-4.92; 85.85)	(-4.42; 78.75)	(-5.92; 49.85)
Observations	125	120	77

Notes: The table presents mean, standard deviation, median and the range of the difference between subjective and objective risk measures, by treatment group. If subjects report probabilities as a range, we use the midpoint of the range as measure for the subjective probability.

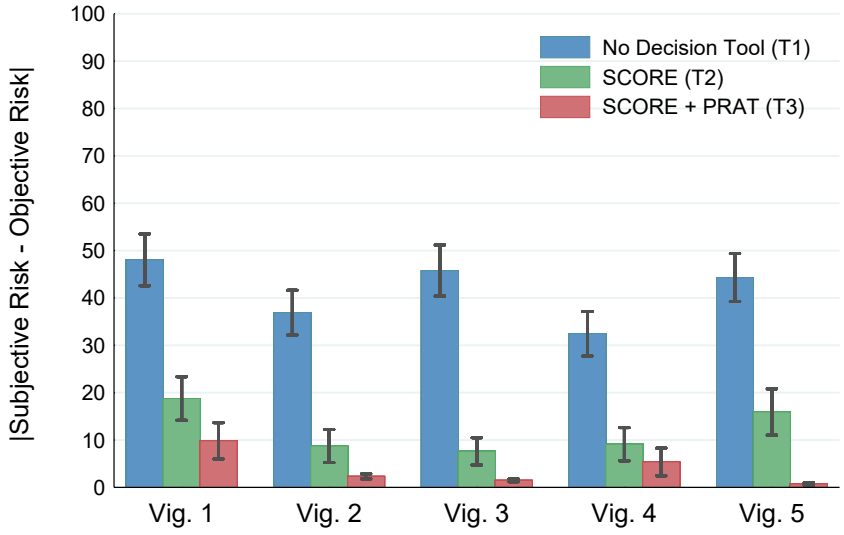
Figure 1.4: Empirical Distribution of the Difference between Subj. Risk and Obj. Risk, by Treatment Group



Notes: The figure displays the kernel density of the difference between subjective and objective risk measures, by treatment group. If subjects report probabilities as a range, we use the midpoint of the range as measure for the subjective probability. The black, vertical line at a value of zero displays the point at which subjective risk is equal to the objective risk measure. Calculations are based on an Epanechnikov kernel function.

Average Treatment Effects Next, we turn to study the average treatment effect of the provision of decision tools on the accuracy in risk assessment. Figure 1.5 presents the mean absolute difference between subjective and objective risk assessment in percentage points by treatment status and vignette. The figure illustrates that the provision with SCORE helps students to make significantly more precise risk predictions than without any decision tool. The provision of PRAT in addition to SCORE further improves accuracy in risk prediction. The pattern looks similar across all five vignettes.

Figure 1.5: Average Treatment Effect, by Treatment Group and Vignette



Notes: The figure displays the mean absolute difference between subjective and objective risk measures, shown separately by treatment group and vignette. If subjects report probabilities as a range, we use the midpoint of the range as measure for the subjective probability. The error bars present the respective standard errors.

So far, we have been using the midpoint of a range as a measure of subjective beliefs, if a subject reported imprecise probabilities. However, taking the midpoint is clearly only one of an infinity large set of estimates within the range of reported probabilities. Hence, we further analyze the impact of PRAT on the accuracy of the predicted upper (maximum risk) and lower bound (minimum risk). To this end, we take (i) the absolute difference between the upper bound of a stated range and the objective risk, and (ii) the absolute difference between the lower bound of a stated range and the objective risk as measures for accuracy in risk estimation. This procedure allows us to disentangle whether the provision of decision tools enables subjects to identify more precisely average risk (midpoints), maximum risk (upper bound), and/or minimum risk (lower bound).

Table 1.4 presents the empirical findings of mean linear regressions in that we use different measures of risk accuracy as a function of the treatment status, a set of covariates and order fixed effects. In all specifications, we find that the provision of SCORE plus PRAT (baseline category, T3) increases accuracy in risk prediction compared to assessment without any decision tool (T1) at a 1% significance level. Further, the provision of PRAT significantly improves predicted average risk, as well as the predicted maximum risk compared to using solely SCORE (T2). We don't, however, find that the use of PRAT improves the precision of the estimated lower bound of a probability range compared to using SCORE.

Table 1.4: Treatment Effects of the Provision of Decision Tools on Accuracy in Risk Prediction

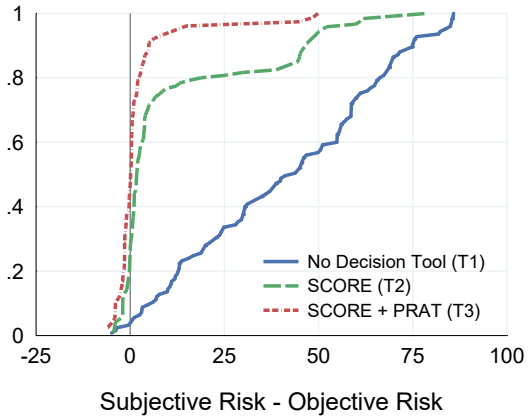
	Absolute Difference between Subjective and Objective Risk								
	Midpoint			**Lower Bound**			**Upper Bound**		
No Decision Tool (T1)	37.44***	37.47***	38.49***	35.98***	36.02***	36.46***	37.64***	37.67***	39.43***
	(4.75)	(4.78)	(5.24)	(4.52)	(4.55)	(4.98)	(5.54)	(5.57)	(6.13)
SCORE (T2)	7.97**	8.00**	7.97*	3.66	3.69	2.62	12.63**	12.66**	14.01**
	(3.33)	(3.37)	(4.41)	(2.59)	(2.62)	(3.71)	(5.07)	(5.11)	(6.15)
Constant	4.058***	4.035***	19.39	3.103***	3.078***	399.2	6.339***	6.315***	-473.8
	(1.12)	(1.13)	(1083.50)	(0.53)	(0.52)	(917.92)	(2.10)	(2.11)	(1393.95)
Observations	322	322	322	322	322	322	322	322	322
Adjusted R^2	0.383	0.380	0.396	0.451	0.446	0.456	0.238	0.236	0.253
Order FE	no	yes	yes	no	yes	yes	no	yes	yes
Controls	no	no	yes	no	no	yes	no	no	yes

Notes: The table summarizes the average treatment effects of using no decision tool (Row 1) or SCORE (Row 2) relative to being provided with SCORE plus PRAT (baseline category). Accuracy in risk prediction is measured by the absolute difference between subjective and objective risk measures. Columns 1 to 3 use the midpoint as measure for subjective risk, if a range is stated. Columns 4 to 6 use lower bounds and Columns 7 to 10 use upper bound as measure for subjective risk, if subjects report a probability interval. We control for order fixed effects and a set of covariates (i.e. age, gender, location of residence, semester of studies work experience in health care prior to studies). Standard errors are clustered at the subject's level. Significance levels at 10%, 5% and 1% are reported by ***, ** and *, respectively.

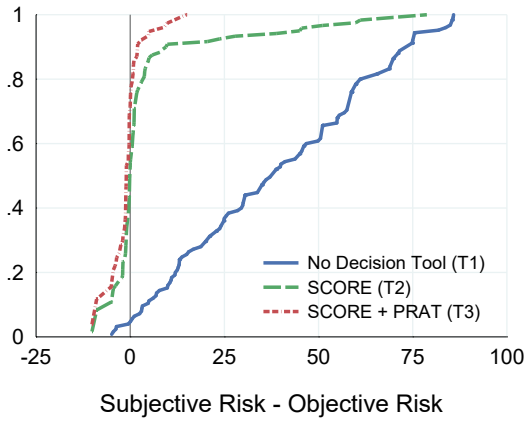
Cumulative Distribution and Quantile Regressions We further aim to study the causal effect of PRAT on risk assessment beyond average effects. To this end, we first investigate the cumulative distribution functions of the difference between subjective and objective risks for each measure of subjective expectations, stratified by treatment group. Figure 1.6 provides four key insights: (1) Without the provision of any decision tool, students overestimate risk on an almost uniform distribution. Only 10% of the responses fall into a range of +/-5 pp around 0. This holds for each of the three measures for subjective risk. (2) Risk estimations in T2 are significantly more precise. About 70% of the responses on average risks are in a range of +/-5pp around 0 (79% lower bound; 67% upper bound). The findings indicate that SCORE is a well-working decision

Figure 1.6: Cumulative Distribution

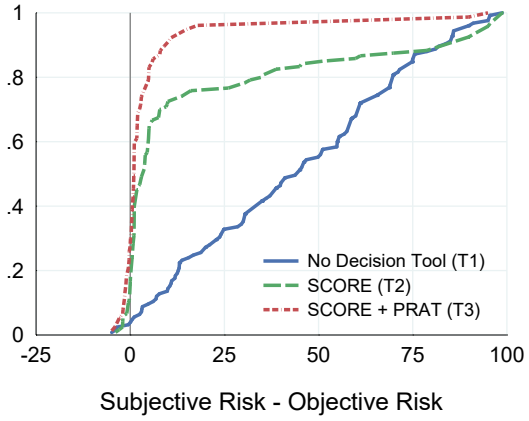
(a) Midpoint



(b) Lower Bound



(c) Upper Bound



Notes:. The figure illustrates the cumulative distribution functions of the differences between subjective and objective risk measures, by treatment group and different measures for the subjective risk. The blue line represents the cumulative distribution of accuracy in risk prediction in T1 (no decision tool), the green line in T2 (SCORE), and the red line in T3 (SCORE + PRAT), respectively.

aid that enables students to make more informative risk assessments. (3) The use of PRAT in addition to SCORE helps students to assess the vignettes' risk even more precisely. We find that about 90% of the responses on average risks fall in a range of +/-5pp around 0 (80% lower bound, 81% upper bound). (4) The gap in accuracy between risk assessment using SCORE versus PRAT is especially large for the predicted upper bound of a vignette's risk. This indicates that PRAT particularly supports subjects to predict the maximum risk more precisely.

To quantify our findings, we run quantile regressions using 10^{th}, 25^{th}, 50^{th}, 75^{th}, 90^{th} quantile (see Appendix Table A.6). In line, with the insights derived from Figure 1.6, we find that using PRAT particularly increases precision in risk assessment at the 90^{th} quantile and in estimating the upper bound of a vignette's risk.

1.5.2 Treatment Choice

The results in Section 1.5.1 show that using PRAT enables students to improve accuracy in their risk assessment. Since risk prediction is not an objective per se, but rather the first stage of a decision making process, we focus in this section on the question whether and to what extent the provision of PRAT alters treatment recommendations compared to treatment choices (i) without a decision tools and (ii) with using SCORE.

Before turning to the empirical analysis, we seek to emphasize a few theoretical considerations: First, the clinical guidelines classify patients' risk into three broad categories that are used to derive treatment recommendations (for more details see Section 1.3.1). For patients in the low-to-moderate risk category the guidelines recommend to prescribe only *some* of the medical treatments that the students are asked to evaluate in the survey, while for patients in the very high risk category the guidelines recommend to apply *all* of the treatment options covered in the survey. Second, the results in the previous section reveal that subjects primarily overestimate and hardly underestimate the risk of cardiovascular mortality (see Figure 1.4). The use of decision tools helps subjects to restrict the magnitude of overestimated risk levels. Hence, we expect that decision making without using algorithms leads, if at all, to overtreatment, in the sense that subjects recommend a larger set of medical treatments than is advised by the guidelines at a particular objective risk level. This effect, however, can not be studied for vignettes in the very high risk category, since the guidelines advise to prescribe in this risk category all treatment options considered in the survey already at the vignette's objective risk level. Therefore, we analyze the impact of using decision tools on treatment choices for each risk category separately. Finally, it is worthwhile to mention that we intentionally did not inform subjects about the guideline's risk categories and respective treatment recommendations to avoid anchoring.

Descriptive Statistics Table 1.5 and Table 1.6 summarize the fraction of vignettes that students advise to undergo a particular medical treatment (i.e. drug and/or behavioral treatment) or re-

fer to a cardiologist, by treatment status and risk category (see Table 1.1 for information on the classification of vignettes into different risk categories).

We find almost no variation in the recommended intake of antihypertensive drugs and lifestyle advice across treatment groups and risk categories. Students recommend the prescription of antihypertensive medication and a change in lifestyle in almost 100% of the vignette scenarios. The results indicate that irrespective of the availability or the type of the decision tool, medical students recommend the prescription of antihypertensive medication for vignettes with low-to-moderate cardiovascular risk, while clinical guidelines do not advice to prescribe medication in this risk category. This finding, however, may be explained by the fact that the vignette with low-to-moderate risk levels has a fairly small total risk of cardiovascular mortality of only 1%, but suffers from high blood pressure (i.e. systolic blood pressure of 180mmHg). The results on hypertensive treatment raise a very interesting question, namely how much emphasis decision makers put on the risk dimension that is targeted by a specific treatment (here: blood pressure) relative to the total risk of a health outcome when making treatment choices.

Next, we investigate the impact of the use of decision tools on the recommended intake of statins. In line with the theoretical considerations, we find that the use of decision tools lowers the share of vignettes in the low-to-moderate risk category that is recommended to take statins, whereas treatment recommendations in the very high risk category do not differ across treatment status.

Table 1.5: Treatment Recommendations - Drug and Behavioral Treatment

	Chol. Drug			**BP. Drug**			**BP. and Chol. Drug**			**Lifestyle Advice**		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
Low to Moderate Risk (in %)	48	29.17	28.57	92	100	100	48	29.17	23.53	100	100	92.86
High Risk (in %)	62	39.58	48.39	86	97.92	93.55	58	39.58	41.18	100	100	100
Very High Risk (in %)	56	54.17	56.25	90	95.83	96.88	50	52.08	50	100	97.92	100

Notes: The table reports the fraction of vignettes that was recommended to undergo a particular medical treatment within each risk category and by treatment group. "BP. Drug" denotes medication that lowers the blood pressure and "Chol. Drug" describes medication that aims at lowering the cholesterol level.

Finally, we investigate the likelihood that medical students recommend a vignette to see a cardiologist. The results show that the fraction of vignettes in the low-to-moderate risk category that is referred to a specialist decreases from T1 to T3 (60% in T1, 33% in T2, and 14% in T3). In contrast, the share of vignettes in the very high risk category that is advised to see a specialist increases from T1 to T3 (48% in T1, 60% in T2, and 75% in T3). The pattern can be explained by increasing precision in risk prediction from T1 to T3.

Table 1.6: Treatment Recommendations - Referral to a Specialist

	T1			T2			T3		
	Yes	No	Don't Know	Yes	No	Don't Know	Yes	No	Don't Know
Low to Moderate Risk (in %)	60.00	32.00	8.00	33.33	62.50	4.17	14.29	71.43	14.29
High Risk (in %)	44.00	48.00	8.00	43.75	52.08	4.17	48.39	32.26	19.35
Very High Risk (in %)	48.00	42.00	10.00	60.42	35.42	4.17	75.00	18.75	6.25

Notes: The table reports the fraction of vignettes referred to a specialist, stratified by treatment group and risk category.

Regression Analysis To quantify the impact of using PRAT on treatment choice, we run binary outcome regression models using treatment recommendations as outcome variables (see Appendix Table A.7).

Considering the total sample of vignettes (irrespective of the risk category), we find at most a weak significant impact of the provision of SORE or PRAT on treatment choices (Columns 1 to 4). The results, however, may hide important effects as the provision of decision tools may decrease the likelihood of medical interventions (i.e. medication and/or referral to a specialist) for low-to-moderate risk vignettes, while it may not affect treatment recommendations for vignettes in the high/very high risk category. To this end, we run the analysis for each risk category separately.

In the low-to-moderate risk category, we find that students in T1 (no tool) are significantly more likely to recommend medical interventions compared to students in T3 (PRAT). Students in T2 (SCORE) also tend to recommend more often medical treatments than students in T3, but not at a significant level (Columns 5 to 8). In the high and very high risk category we find as expected that neither the availability nor the type of decision tool affects the treatment recommendations (Columns 9 to 16).

Overall, we want to emphasize that our sample size is quite small, particularly once we stratify by risk category and thus, we can at most give suggestive evidence for the impact of PRAT on treatment choices due to a lack of statistical power (see Section 1.6 for a detailed discussion).

1.5.3 Precise and Imprecise Subjective Probabilities

In Section 1.5.1, we focused on post-probe responses to evaluate accuracy in risk assessment. In this section, we discuss the impact of decision tools on people's nature of reporting precise and imprecise expectations. To this end, we first analyze whether the provision and the type of decision tool affect subject's choice to express precise or imprecise expectations of cardiovascular mortality. Second, we investigate whether the availability of decision tools alters the relationship between initial and post-probe responses.

Patterns of Precise and Imprecise Expectations Panel A in Table 1.7 shows the empirical distribution of three probabilistic response types – (i) those who state that their initial response were exact numbers (ii) those who initially round/approximate and who report precise numbers after probing (iii) those who initially round/approximate and state probability intervals after probing. These statistics are shown separately for subjects in T1, T2, and T3.

About 60 to 80% of the subjects in T1 express precise probabilities after probing, while about half of the subjects in T2 and only about one third of the subjects in T3 report precise estimates. One explanation for the particularly high fraction of subjects that express precise probabilities in T1 (also compare with Table 1 in Giustinelli 2019) may be the survey design of our experiment - subjects in T1 learn that once they state that they meant an "exact number" in the initial question, no follow-up questions are asked. Since the survey is quite long[23] and, subjects in T1 are probed for each vignettes twice (i.e. conditional on 5 and 6 risk factors), they may fatigue of being probed and thus, report that they meant an exact number to complete the survey faster.[24] Subjects, however, do not get feedback on the accuracy in risk assessment during the experiment. Thus, learning effects on the individual performance can be excluded. Besides, the survey design does not allow us to define the fraction of subjects in T3 that initially round/approximate and state an exact number in the post-probe question.[25] Overall, we find no significant differences in the response type between T2 and T3, while subjects in T1 are more likely to report exact numbers than subjects in T2 and T3 (see Appendix Table A.8, Columns 1 to 3).

Next, we analyze whether the provision of algorithm-based decision tools affects the extent of imprecisely stated beliefs among subjects that answered the post-probe question with a probability range. To do so, we compare the empirical distribution of the interval widths of stated probability ranges among students with imprecise probabilities across treatment groups.

Panel B in Table 1.7 summarizes the 1st decile, median, and 9th decile of the respective distribution (see Appendix Figure A.2 for a graphical illustration of the empirical distribution of the interval width, per treatment group). We find that the interval width decreases in the provision of decision tools. More specifically, the average interval width without the provision of a decision tool amounts to 20pp, using SCORE decreases it to 14pp, and under provision of PRAT it shrinks further to 4pp. The decrease in the interval width suggests that subjects feel more certain about reporting expectations of the vignette's cardiovascular mortality, once decision tools are made available and, that certainty in predictions increases in the number of observed vignettes' characteristics that are covered by the available decision tool. Summing up, we observe that the provision

[23] On average students need 40min to complete the survey; see Table 1.2, Panel C

[24] To investigate whether subjects respond to the survey design, we calculate the fraction of individuals in T1 that state precise probabilities for the first and second vignette shown to them. We find that a smaller share of subjects state precise numbers for the first and second vignette (i.e. 5 RF: 48% exact point, 4% rounding & 48% interval; 6 RF: 72% exact point & 28% interval) than for the full sample of 5 vignettes (see Table 1.7).

[25] In the survey, subjects are first asked to state minimum and maximum risk with and without obesity and then after obtaining the bounds calculated by PRAT may directly choose to report an exact number or an interval (see Section 1.3.3).

Table 1.7: Classification of Response to Risk Assessments into Probabilistic Response Types and Width of Interval among Subjects with Imprecise Probability

	T1		T2	T3
	5 Risk Factors	**6 Risk Factors**		
Panel A: Response Type				
Exact Point (in %)	64.00	81.60	56.67	33.77
Rounding (in %)	3.20	1.60	0.83	–
Interval (in %)	32.80	16.80	42.50	66.23
Observations	125	125	120	77
Panel B: Interval Width				
1st Decile	10	10	4	1
Median	20	20	14	4
9th Decile	70	65	96	15
Observations	41	21	51	51

Notes: Panel A summarizes the response type by treatment group. Panel B describes the empirical distribution of the interval widths among subjects that answered the post-probe question with a probability range.

of PRAT significantly lowers the stated interval width and thus, the extent of imprecision in risk assessment (see Appendix Table A.8, Columns 4 to 6).

Relationship between Initial and Post-Probe Expectations Next, we turn to investigate the relationship between initial and post-probe expectations of cardiovascular mortality, stratified by response type and treatment status.

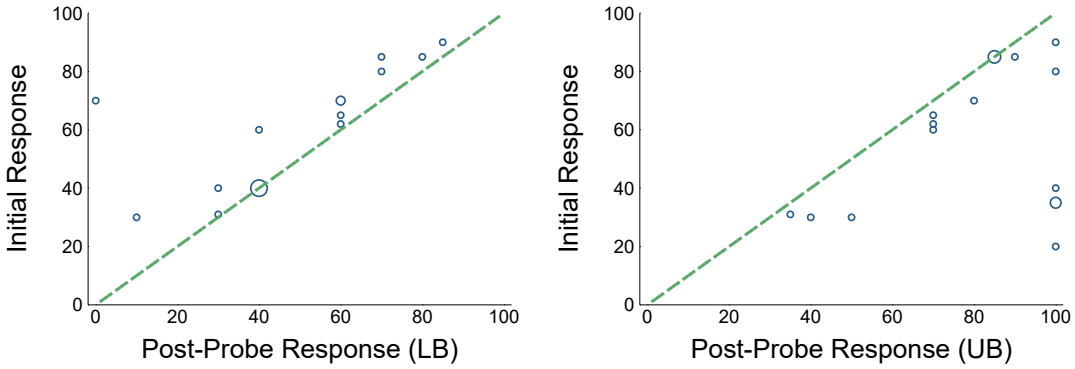
Panel A in Appendix Table A.9 and Appendix Table A.10 summarize the respective empirical distribution of initial and post-probed expectations.[26] Figure 1.7 graphically illustrates the relationship between initial and post-probe response among subjects that state imprecise probabilities. Since the results in Appendix Table A.9, Appendix Table A.10 and Figure 1.7 convey the same main findings, we discuss them together.

We first focus on the effect of the availability of decision tools on the relationship between initial and post-probe expectations among subjects that stated imprecise probabilities. We find that subjects in T2 and T3 are substantially more likely to report a lower bound of 0% after probing than subjects in T1 (5% in T1, 39% in T2, and 31% in T3). The provision of SCORE does not alter the share of subjects stating an upper bound of 100% compared to T1 ($\sim$ 30% in T1 and T2). In contrast, subjects in T3 are significantly less likely to report 100% as upper bound ($\sim$ 6% in T3).

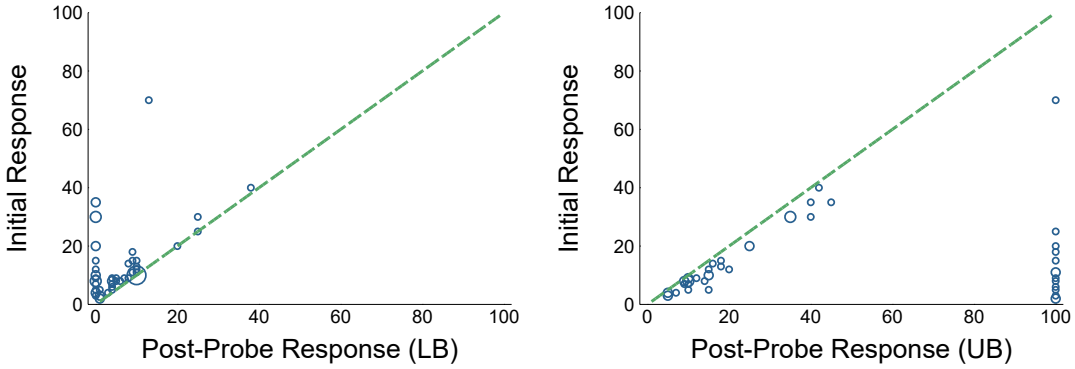
[26] As subjects rarely rounded/approximated in the initial question and stated a precise number in the post-probe question, we do not consider this case in Table A.9 and Table A.10.

Figure 1.7: Relationship Initial versus Post-Probe Response

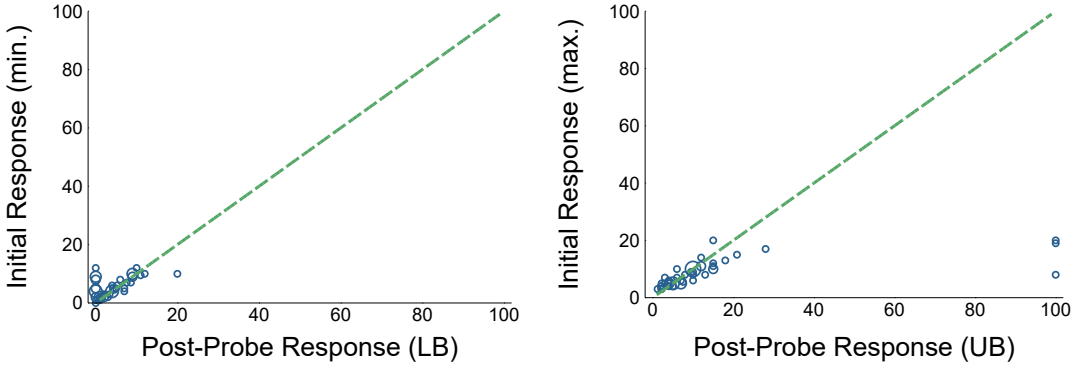
(a) No Decision Tool (T1)



(b) SCORE (T2)



(c) SCORE + PRAT (T3)



Notes: The figure illustrates the relationship between initial responses and post-probe expectations among subjects that stated a range in the probing question. LB (UB) denotes the lower (upper) bound of the stated interval. In T1 and T2, the initial response is a point estimate, while it is a range in T3.

Taking into account that the vignettes' objective risk is at most 11%, the results suggest that the provision of PRAT significantly improves assessing the maximum risk level.

Turning to the impact of decision tools on precisely reported probabilities, we find that students in T2 and T3 use the value predicted by SCORE as an anchor. In 15% (27%) of the responses in T2 (T3), subjects ignore additional information on obesity and stick to the risk calculated by SCORE, and about 75% of precisely reported responses in T2 and T3 fall into an interval defined by the risk of SCORE plus/minus 5pp.

1.5.4 Actuarial versus Clinical Judgment

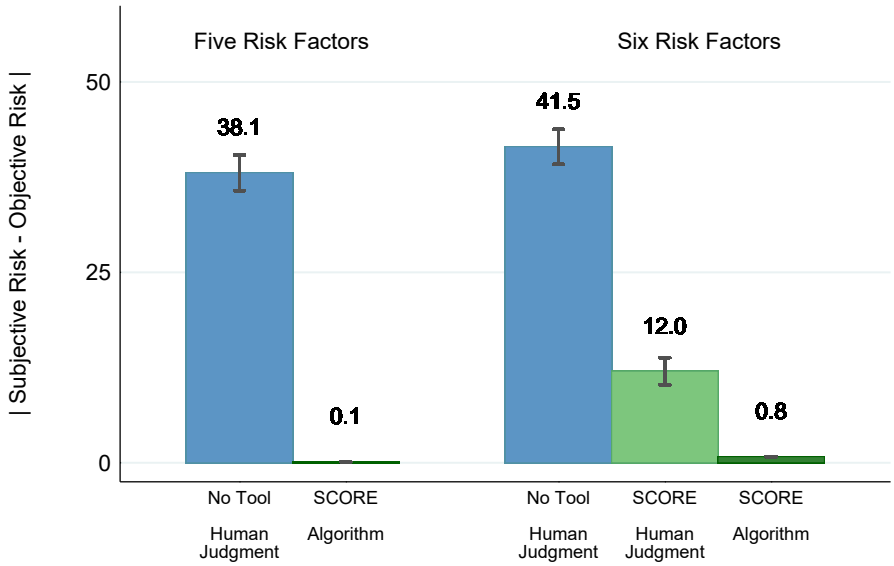
In the previous sections, we evaluated the performance of SCORE and PRAT placed in the hands of humans. This section discusses our results in the light of prior literature that compares actuarial or algorithm based predictions with informal clinical judgment (see e.g. Camerer and Johnson 1991; Dawes et al. 1989; Kleinberg et al. 2018). This body of research addresses a type of "man versus machine" question rather than questioning how humans use predictive algorithms. We add to this strand of literature by assessing the performance of the accuracy in risk prediction of SCORE and PRAT, if subjects had fully complied with the risk predicted by SORE or PRAT. To this end, we compare risk predictions with and without subject's discretional choice.

To begin, we analyze risk assessment maintaining the current status quo (i.e. without the availability of PRAT) in that a physician may stick to the existing tool SCORE that considers five, predefined risk factors or may base prediction solely on clinical judgment. The results confirm previous findings in psychological research: First, risk predictions based on algorithms outperform ones made by clinical judgment using the *same* patient attributes (see Figure 1.8, left bars). Second, the gap in performance remains even when subjects have an informational edge, in the sense that they observe additional predictive attributes that are not included in the decision aid. In line with the so-called "broken leg hypothesis", we observe that subjects put too much emphasis on the additional attribute relative to the remaining risk factors (Dawes et al. 1989) (see Figure 1.8, right bars). Overall, the results based on a sample of medical students align with previous findings that argue that decision makers should rather ignore the additional characteristics than attempt to subjectively include them into risk assessment, as individuals tend to put too much emphasis on the additional information relative to the information included in the existing tool.

Next, we investigate whether the use of PRAT leads to efficiency gains based on the performance of the algorithm. A classical "man versus machine" comparison is impossible in the case of PRAT, as human beliefs serve as input into the predictive algorithm. Therefore, we assess accuracy in risk prediction, if subjects had fully complied to the midpoint of the interval predicted by PRAT. To calculate the prediction based on PRAT, we consider the following three settings: (i) We take subjects' beliefs about the vignette's maximum and minimum risk *elicited in the survey*. (ii) We assume that subjects have *no knowledge* on the vignette's maximum or minimum risk. Without any

knowledge, subjects may choose a value of 0 for vignette's minimum risk and the risk predicted by SCORE as maximum risk if the vignette is not obese. Analogously, subjects may choose a value of 1 as vignette's maximum risk, and the risk predicted by SCORE as minimum risk, if the vignette is obese. (iii) We calculate PRAT based on a sample of students with *advanced knowledge* on cardiovascular mortality risk. In line with the theoretical considerations based on Manski (2018), it is plausible to assume that subjects with advanced knowledge may state beliefs on the vignette's minimum and maximum risk with smaller interval width. Hence, in this setting, we calculate PRAT based on a subsample of medical students that stated beliefs with an interval width smaller than the 25th percentile of the empirical distribution.

Figure 1.8: Risk Prediction under Status-Quo Conditions (without PRAT)

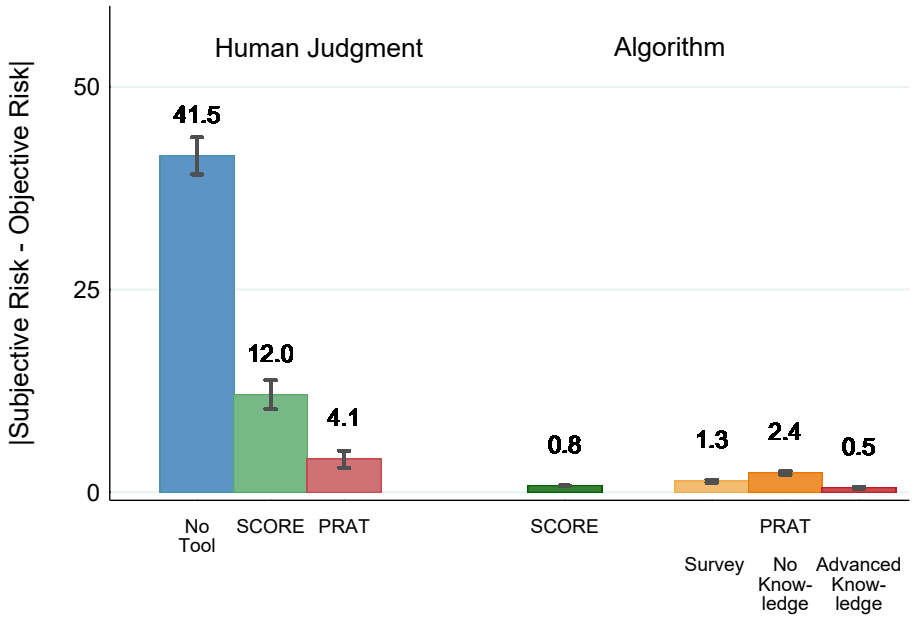


Notes: The bars display the mean absolute difference between the risk of cardiovascular mortality predicted by evidence-based algorithms (SCORE) or clinical judgment and the objective risk conditional on 5 or 6 risk factors. The errors bars present the respective standard errors.

Figure 1.9 summarizes the average absolute difference between risk predicted based on clinical judgment and/or evidence-based tools and the objective risk level. In scenarios with human discretion, the provision of PRAT leads to significant efficiency gains compared to risk assessment without decision aid or with SCORE (as described in Section 1.5.1). However, if subjects had fully complied to the risk predicted by SCORE and thus, ignored the additional information on obesity, they would have performed even better than by using PRAT with discretional choice. The performance of PRAT without human discretion (i.e. subjects stick to the midpoint of the predicted interval) depends substantially on people's knowledge how to place bounds on the vignette's min-

imum and maximum risk. We find that advanced knowledge on cardiovascular mortality risk improves accuracy in risk prediction by about 2pp compared to a setting without any knowledge on cardiovascular mortality risk. Overall, the results suggest that risk assessment based on the algorithm without discretional choice outperforms prediction under discretion.

Figure 1.9: Actuarial vs. Clinical Judgment



Notes: The bars display the mean absolute difference between the risk of cardiovascular mortality predicted by evidence-based algorithms (SCORE/PRAT) or clinical judgment and the objective risk. The errors bars present the respective standard errors.

1.6 Discussion

This study provides evidence that the provision of PRAT in addition to SCORE improves risk assessment. The study is, however, subject to several limitations. First, the sample consists of medical students and it is relatively small. Medical students have some clinical experience, but certainly less than practicing physicians. Therefore, it is plausible that physicians assess the risk of cardiovascular mortality more precisely than medical students. This would hold across all three treatment groups and thus, it is not clear whether and to which extent PRAT enables physicians *with profound clinical experience* to improve accuracy of risk prediction (i) compared to risk assessment without the provision of a decision tool and (ii) under the provision of SCORE. While our data show that the use of PRAT has a positive effect on risk assessment in a sample of medical

students, additional experiments with practicing physicians are needed to evaluate the use of PRAT in clinical practice.

Second, we study the use of PRAT in a fairly simplified setup, as the risk factors are well defined and we provide subjects with only *one* risk factor additionally to those included in SCORE. In clinical practice, however, physicians are often confronted with more than one risk modifying factor in addition to those included in an existing decision tool. This complicates placing bounds within the framework of PRAT, particularly if the effect of the additional risk factors for cardiovascular mortality goes in different directions (e.g. if a patient has diabetes, but is not obese). In this scenario, the physician must weigh the relative risk of diabetes against the relative risk of not being obese. This adds another layer of complexity to the decision making process and it may be worthwhile to assess PRAT in the hands of physicians under the provision of more than one additional risk factor.

A related question that remains open in our study is whether and to which extent the improvement in risk accuracy in T3 can be explained by the provision of the risk predicted by PRAT or by triggering some reflection process through the elicitation mode. More specifically, the elicitation procedure in T3 requires decision makers to think in a structured and quantitative way about maximum and minimum risk. This, in turn, may induce some reflection process that may explain alone without the actual information on the risk predicted by PRAT an increased accuracy in risk prediction. One approach to disentangle these effects, is to add an additional treatment group to the experiment in that physicians are asked to report their beliefs on the minimum and maximum risk of cardiovascular mortality, but they are not informed about the risk predicted by PRAT.

Furthermore, it might add value to elicit treatment recommendations not as discrete choice (yes/no), but rather as probability on a scale from 0 to 100 in the survey among physicians. This would allow us to investigate the impact of using risk assessment tools on treatment choices in a more granular manner.

Finally, it is worthwhile to comment on the estimation of the decision tool SCORE and its recommended application: The sample used to estimate SCORE includes individuals with established cardiovascular diseases except for those with previous history of heart attacks, as well as individuals with chronic diseases that relate to the risk of cardiovascular mortality; e.g. diabetes or chronic kidney disease (Conroy 2003). The use of SCORE, however, is explicitly recommended only to apparently healthy individuals, i.e. people without established cardiovascular diseases, chronic kidney disease or any risk modifying pre-condition, such as cancer, diabetes with severe organ damage (Piepoli et al. 2016). Hence, the calculated risk level is based on a systematically different sample than the sample it is supposed to be applied to. More precisely, it is plausible to assume that SCORE on average overestimates risk for apparently healthy individuals. Analogously, the estimation of SCORE includes individuals aged between 19 and 80 years, whereas its application is intended to an age range from 40 to 65 years. Overall, as the sample used to estimate SCORE in the first place and the sample of patients to whom SCORE is supposed to be applied to may differ,

SCORE may not predict risk fully accurately. Despite of these limitations, we use SCORE in our analysis, as it is the risk assessment tool currently recommended by the European guidelines on cardiovascular disease prevention. However, we emphasize that PRAT could easily be applied to any other existing risk assessment tool.

1.7 Conclusion

Improving risk prediction and treatment choice is a key priority in health policy. Particularly in preventive health care accurate risk predictions may reduce disease prevalence and thus, individual and also economic burden. Algorithm-based decision tools are one approach that may enable physicians to make more informative and evidence-based choices.

We present findings from a new risk assessment tool called "Personalized Risk Assessment Tool" (PRAT) that allows physicians to incorporate their clinical experience on the impact of additional risk factors in a structured way. Leveraging an online experiment with medical students, we document that the new tool significantly improves precision in risk assessment compared to SCORE, a decision aid currently applied in clinical practice. We uncover two mechanisms that explain the increased accuracy in risk estimation achieved through the provision of PRAT: First, PRAT helps students to bound the vignette's maximum risk significantly better than using only SCORE. Second, PRAT supports students holding imprecise beliefs on the risk of cardiovascular mortality to state tightened probability intervals. This indicates that students feel more certain about reporting expectations of the patient's cardiovascular mortality when using PRAT. We further provide suggestive evidence that the use of PRAT enables students to make treatment choices better aligned with treatment recommendations in clinical practice guidelines.

Overall, our findings underscore the contribution of algorithm-based prediction tools to improving risk assessment in health care.

Appendix A

A1 Design of the Experiment

Supplementary Information on the Construction of the Vignettes

Table A.1: Risk Prediction Models

	Dep. Var.: CVD Mortality		**Dep. Var.: Obesity**
	5 Risk Factors	**6 Risk Factors**	
Currently Smoking (=1)	0.269**	0.279**	-0.045
	(0.119)	(0.119)	(0.042)
Cholesterol Level (in mmol/l)	0.021	0.024	0.029*
	(0.039)	(0.039)	(0.016)
Systolic Blood Pressure (in mmHg)	0.005**	0.005**	0.014***
	(0.002)	(0.002)	(0.001)
Age (in years)	0.061***	0.061***	0.008***
	(0.006)	(0.007)	(0.002)
Male (=1)	0.308***	0.322***	-0.101***
	(0.098)	(0.099)	(0.038)
Obesity (=1)		0.113	
		(0.104)	
Constant	-6.788***	-6.824***	-3.242***
	(0.513)	(0.523)	(0.135)
Observations	6274	6274	6274
Pseudo R2	0.2702	0.2713	0.0738

Notes: The table presents the estimated coefficients derived in the risk prediction models using Probit regression. Calculations are based on data from the RKI study (i.e. baseline wave and mortality follow-up). The dependent variable is in Columns 1 and 2 cardiovascular mortality and in Column 3 obesity. Significance levels at 10%, 5% and 1% are reported by ***, ** and *, respectively.

To calculate the vignettes' objective risk level, we insert the vignettes' risk factors presented in Table 1.1 into Equation A.1 and A.2 and predict the vignettes' risk level conditional on five and six risk factors, respectively.

$$\begin{aligned} Pr(Y = 1|x = k) = \Phi(&-6.788 + 0.269 * D_{\text{Smoking}} + 0.021 * \text{Cholesterol Level} \\ &+0.005 * \text{Blood Pressure} + 0.061 * \text{Age} + 0.308 * D_{\text{Male}}) \end{aligned} \tag{A.1}$$

$$\begin{aligned} Pr(Y = 1|x = k, w = j) = \Phi(&-6.824 + 0.279 * D_{\text{Smoking}} + 0.024 * \text{Cholesterol Level} \\ &+0.005 * \text{Blood Pressure} + 0.061 * \text{Age} + 0.322 * D_{\text{Male}} + 0.113 * D_{\text{Obese}}) \end{aligned} \tag{A.2}$$

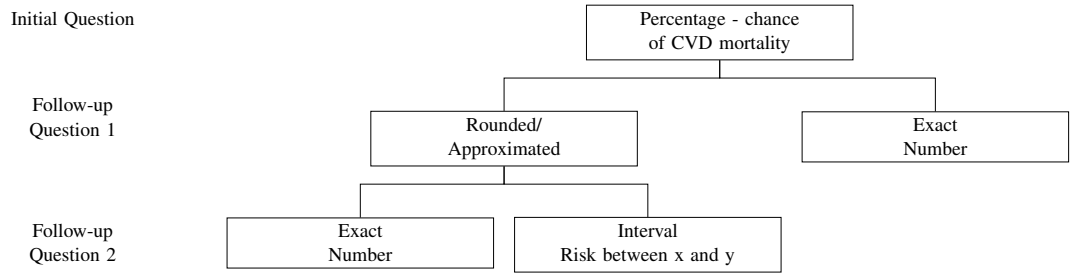
In addition, we use the data from the RKI study to calculate the vignettes' probability of being obese conditional on age, gender, cholesterol level, systolic blood pressure, and smoking habits. This provides us with the probability $P(w = \text{obese}|k)$ which is used as an input factor in the algorithm of PRAT (see Equation 1.6a and 1.6b). In doing so, we insert the vignettes' risk factors into Equation A.3.

$$\begin{aligned} Pr(w = 1|x = k) = \Phi(&-3.242 - 0.045 * D_{\text{Smoking}} + 0.029 * \text{Cholesterol Level} \\ &+0.014 * \text{Blood Pressure} + 0.008 * \text{Age} - 0.101 * D_{\text{Male}}) \end{aligned} \tag{A.3}$$

APPENDIX A

Supplementary Information on the Experimental Procedure

Figure A.1: Elicitation of Probabilistic Beliefs



Notes: The figure presents the elicitation procedure of probabilistic beliefs used to distinguish between subjects holding precise beliefs (i.e. subjects stated an exact number in the post-probe questions) and subjects holding imprecise beliefs (i.e. subjects stated a probability interval in the post-probe question).

Supplementary Information on Survey Questions

Survey Questions on Treatment Recommendations: *Based on your risk assessment, which treatment(s) would you recommend the patient?*

1. *Medication to lower the blood pressure (yes/no)*
2. *Medication to lower the cholesterol level (yes/no)*
3. *Lifestyle advice (yes/no)*

***Referral**: Would you refer the patient to a cardiologist? (yes/no/don't know)*

A2 Data and Sample Description

Table A.2: Sample Flow and Attrition by Treatment Group

	No Decision Tool (T1)	**SCORE (T2)**	**SCORE + PRAT (T3)**
Started with survey	83	81	75
Completed survey on demographics	46	36	47
Started with 1st vignette	40	29	31
Completed 1st vignette	33	29	21
Started with 2nd vignette	29	27	21
Completed 2nd vignette	29	27	15[27]
Started with 3rd vignette	26	24	17
Completed 3rd vignette	26	24	16
Started with 4th vignette	25	24	17
Completed 4th vignette	25	24	16
Started with 5th vignette	25	24	17
Completed 5th vignette	25	24	16
Survey on SCORE	25	24	17
No. of Completed Survey	25	24	16 [28]
No. of Completed Vignettes in Completed Surveys	125	120	77 [29]
No. of Completed Vignettes in Total Sample	138	128	84

Notes: The table reports the total number of medical students that completed different stages of the survey, stratified by treatment group.

[27] If participants did not respond according to the following monotonicity assumption: $P_{\text{min}}(x = k, w = \text{not obese}) \leq P_{\text{max}}(x = k, w = \text{not obese}) \leq P_{\text{SCORE}} \leq P_{\text{min}}(x = k, w = \text{obese}) \leq P_{\text{max}}(x = k, w = \text{obese})$ within 3 attempts, they were directly led to assess the subsequent vignettes' risk (i.e. participants may not complete survey on risk assessment for one particular vignette but may continue with survey).

[28] One participant never answered according to the monotonicity assumption. Hence, the participant completed the survey without assessing any of the 5 vignettes' risk.

[29] Thirteen participants assessed risk in line with the monotonicity assumption within 3 attempts for all 5 vignettes.

APPENDIX A

Table A.3: Characteristics of Attritors

	Sample		Attr. Def. 1		Attr. Def. 2		Diff. Sample - Attr. 1	Diff. Sample - Attr. 2
	Mean	S.D.	Mean	S.D.	Mean	S.D.	p-value	p-value
Year of Birth (in years)	1995.58	3.92	1994.99	3.71	1994.86	3.72	0.35	0.28
Gender (male = 1)	0.29	0.46	0.35	0.48	0.34	0.48	0.50	0.53
Semester	6.46	2.87	6.73	2.52	6.48	2.52	0.56	0.98
Location of Residence (Bavaria =1)	0.82	0.39	0.35	0.48	0.83	0.38	0.00	0.85
Preclincal Period - Internship in Cardiology/Internal Medicine	0.61	0.49	0.50	0.50	0.52	0.50	0.21	0.29
Clincal Period - Internship in Cardiology/Internal Medicine	0.62	0.49	0.84	0.37	0.84	0.37	0.01	0.01
Work experience	0.17	0.38	0.38	0.49	0.38	0.49	0.01	0.01
Observations	65		174		64		239	129

Notes: Columns 1 and 2 show the mean and standard deviation in the sample that completed the survey. Columns 3 to 6 present the mean and standard deviation in the sample of attritors that left the survey before completion. We use two alternative definitions for the sample of attritors: (i) subjects that started the survey, but stopped the questionnaire at any point before completion, i.e. Definition 1, (ii) subjects that finished the first part of the survey but left the survey before completion, i.e. Definition 2. Columns 7 and 8 show the p-values of the differences in means between the different samples.

Table A.4: Characteristics of Attritors by Treatment Group, Definition 1

	Attrition T1		Attrition T2		Attrition T3		Diff. T1-T2	Diff. T1-T3	Diff. T2-T3
	Mean	S.D.	Mean	S.D.	Mean	S.D.	p-value	p-value	p-value
Year of Birth	1995.85	2.67	1994.94	4.38	1994.28	4.04	0.40	0.09	0.60
Gender	0.26	0.45	0.29	0.47	0.45	0.51	0.81	0.13	0.30
Semester	6.80	2.87	6.60	2.69	6.74	2.21	0.83	0.93	0.85
Preclincal Period of Studies - Internship in Cardiology/Internal Medicine	0.52	0.51	0.46	0.52	0.50	0.51	0.73	0.87	0.82
Clincal Period of Studies - Internship in Cardiology/Internal Medicine	0.88	0.34	0.88	0.35	0.80	0.41	1.00	0.55	0.64
Residence of Location	0.41	0.50	0.23	0.42	0.41	0.50	0.03	0.94	0.04
Work experience in health care prior to studies	0.24	0.44	0.33	0.49	0.48	0.51	0.57	0.08	0.38
Observations	58		57		59		115	117	116

Notes: The table compares sociodemographic characteristics of attritors by treatment status using definition 1 for the sample of attritors. Columns 1 to 6 report the mean and the standard deviation in the sample of attritors, stratified by treatment status. Columns 7 to 9 show the p-values of the differences in means between the different samples.

Table A.5: Characteristics of Attritors by Treatment Group, Definition 2

	Attrition T1		**Attrition T2**		**Attrition T3**		**Diff. T1-T2**	**Diff. T1-T3**	**Diff. T2-T3**
	Mean	S.D.	Mean	S.D.	Mean	S.D.	p-value	p-value	p-value
Year of Birth	1995.90	2.91	1994.92	4.06	1994.13	4.01	0.42	0.09	0.57
Gender	0.29	0.46	0.17	0.39	0.45	0.51	0.46	0.24	0.09
Semester	6.29	2.83	6.09	2.88	6.74	2.21	0.86	0.52	0.44
Preclincal Period of Studies - Internship in Cardiology/Internal Medicine	0.52	0.51	0.55	0.52	0.50	0.51	0.91	0.87	0.80
Clincal Period of Studies - Internship in Cardiology/Internal Medicine	0.88	0.34	0.88	0.35	0.80	0.41	1.00	0.55	0.64
Residence of Location	0.90	0.30	0.83	0.39	0.77	0.43	0.56	0.23	0.68
Work experience in health care prior to studies	0.24	0.44	0.33	0.49	0.48	0.51	0.57	0.08	0.38
Observations	21		12		31		33	52	43

Notes: The table compares sociodemographic characteristics of attritors by treatment status using definition 2 for the sample of attritors. Columns 1 to 6 report the mean and the standard deviation in the sample of attritors, stratified by treatment status. Columns 7 to 9 show the p-values of the differences in means between the different samples.

A3 Results

Table A.6: Treatment Effects of the Provision of Decision Tools on Accuracy in Risk Prediction, Quantile Regression

	Abs. Diff. between Subj. & Obj. Risk								
	Midpoint			Lower Bound			Upper Bound		
10th Quantile									
No Decision Tool (T1)	5.830**	5.830**	5.571**	4.670**	4.670**	4.709**	5.790***	5.790**	5.489**
	(2.31)	(2.34)	(2.23)	(1.81)	(1.83)	(2.02)	(2.22)	(2.41)	(2.33)
SCORE (T2)	0.170	0.170	-0.114	-0.100	-0.100	-0.0710	0.200	0.200	-0.202
	(0.18)	(0.18)	(0.27)	(0.08)	(0.08)	(0.18)	(0.19)	(0.19)	(0.28)
Constant	0.250***	0.250***	60.28	0.250***	0.250***	-9.957	0.290**	0.290**	-9.764
	(0.07)	(0.06)	(61.73)	(0.04)	(0.06)	(23.92)	(0.12)	(0.12)	(83.58)
25th Quantile									
No Decision Tool (T1)	18.26***	18.26***	17.85***	14.48***	14.48***	14.06***	17.90***	17.90***	17.81***
	(4.24)	(4.10)	(4.21)	(3.57)	(3.65)	(3.56)	(4.33)	(4.55)	(4.49)
SCORE (T2)	0.500*	0.500*	-0.304	0.240	0.240	-0.387	0.160	0.160	-0.274
	(0.30)	(0.27)	(0.39)	(0.30)	(0.30)	(0.34)	(0.25)	(0.25)	(0.38)
Constant	0.490**	0.490**	130.8*	0.510**	0.510**	-7.870	0.850***	0.850***	140.0
	(0.20)	(0.19)	(74.90)	(0.23)	(0.24)	(80.52)	(0.17)	(0.17)	(133.39)
50th Quantile									
No Decision Tool (T1)	42.26***	42.26***	42.06***	37.43***	37.43***	37.08***	43.50***	43.50***	43.86***
	(4.51)	(4.53)	(4.94)	(4.51)	(4.21)	(4.39)	(4.48)	(4.58)	(4.77)
SCORE (T2)	0.860	0.860	1.175	0.430	0.430	-0.125	2.260***	2.260***	3.459*
	(0.61)	(0.53)	(0.98)	(0.49)	(0.45)	(0.57)	(0.81)	(0.82)	(1.97)
Constant	1.490***	1.490***	-55.11	1.320***	1.320***	197.4	1.490***	1.490***	-93.15
	(0.24)	(0.20)	(135.98)	(0.33)	(0.32)	(311.61)	(0.31)	(0.32)	(278.37)
75th Quantile									
No Decision Tool (T1)	57.33***	57.33***	63.44***	54.26***	54.26***	56.77***	63.90***	63.90***	67.29***
	(3.11)	(3.15)	(3.86)	(3.16)	(2.99)	(3.27)	(3.70)	(3.76)	(3.54)
SCORE (T2)	4.830	4.830	21.78**	0.500	0.500	2.947	11.23	11.23	27.29***
	(7.67)	(7.25)	(9.63)	(1.83)	(1.63)	(2.14)	(10.46)	(10.65)	(8.95)
Constant	3.750***	3.750***	-11.82	4.490***	4.490***	387.1	4.850***	4.850***	-1048.9
	(0.63)	(0.60)	(1017.09)	(1.07)	(1.00)	(557.53)	(1.01)	(0.95)	(1333.76)
90th Quantile									
No Decision Tool (T1)	69.07***	69.07***	68.83***	65.93***	65.93***	67.45***	73.76***	73.76***	74.45***
	(5.60)	(5.87)	(5.49)	(2.81)	(2.83)	(3.10)	(8.71)	(9.38)	(8.66)
SCORE (T2)	41.07***	41.07***	34.20***	1.230	1.230	15.28*	79.86***	79.86***	74.55***
	(5.68)	(5.49)	(5.65)	(10.13)	(10.17)	(7.78)	(14.26)	(15.40)	(13.81)
Constant	5.920	5.920	132.6	8.920***	8.920***	-638.9	9.990	9.990	-2240.3
	(4.55)	(4.85)	(1604.59)	(1.00)	(0.98)	(1035.68)	(7.56)	(8.57)	(2667.54)
Observations	322	322	322	322	322	322	322	322	322
Order FE	no	yes	yes	no	yes	yes	no	yes	yes
Controls	no	no	yes	no	no	yes	no	no	yes

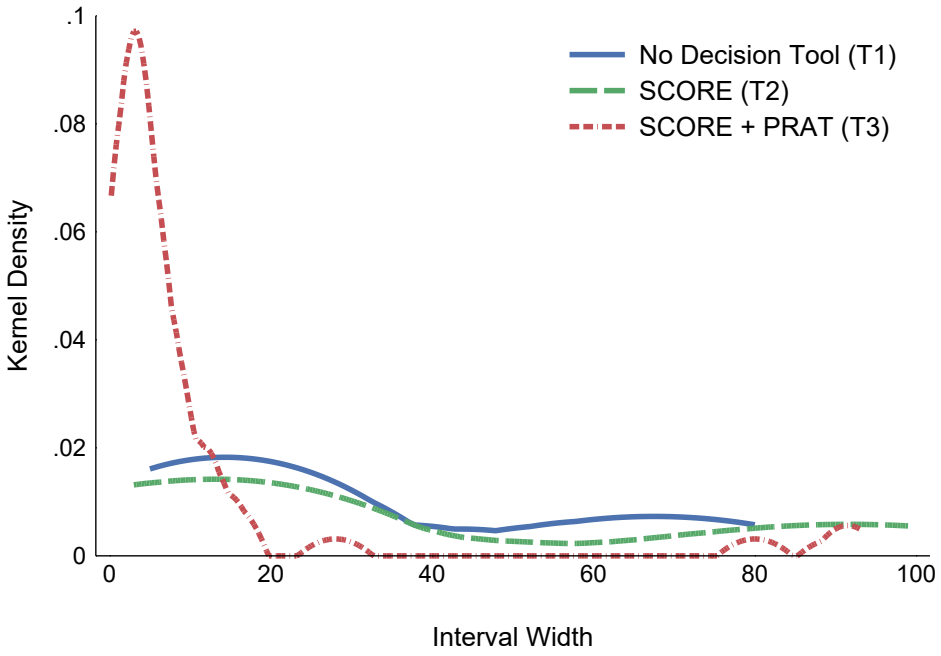
Notes: The table summarizes the results of a quantile regression that investigates the effects of using no decision tool (Row 1) or SCORE (Row 2) relative to being provided with SCORE plus PRAT (baseline category) on accuracy in risk assessment. Accuracy in risk prediction is measured by the absolute difference between subjective and objective risk measures. Columns 1 to 3 use the midpoint as measure for subjective risk if a range is stated. Columns 4 to 6 use the lower bound and Columns 7 to 10 the upper bound as measures for subjective risk if a probability interval is stated. We control for order fixed effects and a set of covariates (i.e. age, gender, location of residence, semester of studies, work experience in health care prior to studies). Each column presents a separate regression. Results for all quantiles within one column were jointly estimated. Standard errors are bootstrapped with 400 replications. Significance levels at 10%, 5% and 1% are reported by ***, ** and *, respectively.

Table A.7: Treatment Recommendations

	Full Sample				Low-to-Moderate Risk Category			High Risk Category				Very High Risk Category			
	BP. Drug	**Chol. Drug**	**BP.+ Chol Drug**	**Referral**	**Chol. Drug**	**BP.+ Chol Drug**	**Referral**	**BP. Drug**	**Chol. Drug**	**BP.+ Chol Drug**	**Referral**	**BP. Drug**	**Chol. Drug**	**BP.+ Chol Drug**	**Referral**
No Decision Tool (T1)	-0.0616	0.154*	0.141	-0.0158	0.305**	0.305**	0.561***	-0.0505	0.207*	0.185	-0.0845	-0.0783	0.0850	0.0547	-0.241
	(0.05)	(0.09)	(0.09)	(0.11)	(0.14)	(0.14)	(0.15)	(0.07)	(0.13)	(0.12)	(0.13)	(0.07)	(0.08)	(0.08)	(0.15)
SCORE (T2)	0.0209	0.0441	0.0566	-0.0169	0.0991	0.0991	0.218	0.0658*	0.00927	0.0193	-0.0462	-0.0175	0.0838	0.0854	-0.115
	(0.04)	(0.10)	(0.10)	(0.11)	(0.15)	(0.15)	(0.14)	(0.04)	(0.14)	(0.14)	(0.14)	(0.06)	(0.09)	(0.09)	(0.14)
Observations	322	322	322	296	63	63	58	129	129	129	117	130	130	130	121
Log-Likelihood	-66.122	-209.030	-209.309	-192.676	-35.399	-35.399	-29.757	-30.553	-76.655	-77.482	-72.462	-24.897	-84.558	-85.317	-68.619
Pseudo R2	0.118	0.063	0.060	0.056	0.144	0.144	0.249	0.131	0.143	0.133	0.106	0.172	0.054	0.053	0.135

Notes: The table summarizes the results of a Probit regression model that investigates the effects of using no decision tool (Row 1) or SCORE (Row 2) relative to being provided with SCORE plus PRAT (baseline category) on treatment choices. Average marginal effects are reported. Dependent variables: recommended intake of blood pressure lowering drugs (BP Drug=1); recommended intake of cholesterol lowering drugs (Chol. Drug=1); recommended intake of blood pressure lowering drugs and cholesterol lowering drugs (BP Drug and Chol. Drug=1); recommended referral to a cardiologist (Referral=1). Columns 1 to 4 include the full sample. Columns 5 to 8 refer to vignettes with low-to-moderate risk. BP drug is in this risk category not considered, as 100% of vignette scenarios are recommended to take antihypertensive medication in T2 and T3. Columns 9 to 12 include vignettes with high risk and Columns 13 to 16 with very high risk. All regressions control for order fixed effects and a set of control variables (i.e. age, gender, location of residence, semester of studies, work experience prior to health care). Significance levels at 10%, 5% and 1% are reported by ***, ** and *, respectively.

Figure A.2: Empirical Distribution of the Interval Width, by Treatment Group



Notes: The figure displays the kernel density of interval width among subjects stated imprecise probabilities, by treatment group. Calculations are based on an Epanechnikov kernel function.

Table A.8: Treatment Effects of the Provision of Decision Tools on Response Type and Interval Width

	Response Type			**Interval Width**		
No Decision Tool (T1)	-1.346***	-1.375***	-1.210***	19.40*	21.90**	22.80*
	(0.38)	(0.38)	(0.39)	(10.27)	(10.15)	(12.30)
SCORE (T2)	-0.600	-0.613	-0.363	24.56***	25.68***	25.01***
	(0.37)	(0.37)	(0.41)	(7.74)	(8.12)	(8.50)
Constant	-0.435	-0.755**	-83.35	10.12***	9.233***	-2370.0*
	(0.29)	(0.33)	(50.98)	(3.09)	(3.18)	(1233.35)
Observations	322	322	322	123	123	123
Adjusted R^2				0.124	0.171	0.187
Pseudo R^2	0.112	0.126	0.192			
Order FE	no	yes	yes	no	yes	yes
Controls	no	no	yes	no	no	yes

Notes: The table summarizes the effects of using no decision tool (Row 1) or SCORE (Row 2) relative to the use of SCORE plus PRAT (baseline category) on the probabilistic response type and the stated interval width among subjects that report imprecise probabilities. Columns 1 to 3 display the results of an ordered probit regression in that the discrete outcome variable defines the probabilistic response type (1 = subjects that use in their initial response exact numbers; 2 = subjects that round/approximate initially and report precise numbers after probing; 3 = subjects that round/approximate initially and state probability intervals after probing). Columns 4 to 6 illustrate the results from an OLS regression. The dependent variable is defined as the interval width among subjects that state imprecise probabilities of cardiovascular mortality risk. We control for order fixed effects and a set of covariates (i.e. age, gender, location of residence, semester of studies, work experience in health care prior to studies). Standard errors are clustered at the subject's level. Significance levels at 10%, 5% and 1% are reported by ***, ** and *, respectively.

Table A.9: Treatment Groups 1 and 2 – Initial Response and Post-Probe Response

	Treatment Group 1								Treatment Group 2			
	5 Risk Factors				6 Risk Factors				6 Risk Factors			
	Exact Number	Imprecise Number			Exact Number	Imprecise Number			Exact Number	Imprecise Number		
	Initial = Post-Probe	Initial	Post-Probe		Initial = Post Probe	Initial	Post-Probe		Initial = Post Probe	Initial	Post-Probe	
			LB	UB			LB	UB			LB	UB
Panel A: Description of reported values - Response Distribution												
1st decile	5	10	0	20	9	30	20	40	2.5	4	0	9
Mean	31.5	51.75	40.41	69.43	44.04	55.86	47.14	76.67	13.36	13.96	5.27	39.96
Median	38.06	55	40	80	40	60	40	80	6	9	3	18
9th decile	72.5	90	80	100	80	85	80	100	50	30	10	100
0 (share in %)	0	0	14.63	0	0	0	4.76	0	0	0	39.22	0
50 (share in %)	7.50	7.32	4.88	0	4.90	0	0	4.76	2.94	0	0	0
100 (share in %)	1.25	0	0	31.71	0	0	0	33.33	0	0	0	27.45
SCORE (share in %)	1.25	7.32	2.44	0	1.96	0	0	0	14.71	0	27.45	15.69
SCORE +/- 5pp. (share in %)	13.75	7.32	12.2	0	9.8	0	0	0	72.06	62.75	60.78	25.49
Panel B: Accuracy in Risk Prediction												
Obj. risk within a 5 points distance	16.25	-	-		11.76	-	-		83.82	-	-	
Obj. risk within a 10 points distance	23.75	-	-		17.65	-	-		88.24	-	-	
Obj. risk inside interval (share in %)	-	-	19.51		-	-	4.76		-	-	88.23	
Obj. risk above UB (share in %)	-	-	0		-	-	0		-	-	0	
Obj. risk above LB; mean distance betw. obj. risk & UB	-	-	-		-	-	-		-	-	-	
Obj. risk below LB (share in %)	-	-	80.49		-	-	95.24		-	-	11.76	
Obj. risk below LB; mean distance betw. obj. risk & LB	-	-	43.98		-	-	42.95		-	-	15.16	
Observations	80	41	41	41	102	21	21	21	68	51	51	51

Notes: The table describes the relationship between initial and post-probe response in T1 and T2. Panel A describes the distribution of reported values, by treatment group and initial/post-probe question. LB denotes the lower bound and UB the upper bound for individuals that stated an interval in the post-probe question. Panel B presents the accuracy in risk prediction, by treatment group and initial/post-probe question.

Table A.10: Treatment Group 3 – Initial Response and Post-Probe Response

	Initial Response				Post-Probe Response		
	Min. w/o obesity	Max. w/o obesity	Min. with obesity	Max. with obesity	Exact Number	Imprecise Number	
						LB	UB
Panel A: Description of reported values - Response Distribution							
0 (share in %)	18.18	0	0	0	0	31.37	0
50 (share in %)	0	0	0	0	0	0	0
100 (share in %)	0	0	0	0	0	0	5.88
SCORE	13	74	45.5	0	26.9	25.5	25.5
SCORE +/- 5pp. (share in %)	87	100	94.8	53.2	76.9	74.5	71
Panel B: Accuracy in Risk Prediction							
Obj. risk within a 5 points distance	-	-	-	-	88.46	-	
Obj. risk within a 10 points distance	-	-	-	-	92.31	-	
Obj. risk inside interval (share in %)	-	-	-	-	-	66.76	
Obj. risk above UB (share in %)	-	-	-	-	-	15.68	
Obj. risk above LB; mean distance betw. obj. risk & UB	-	-	-	-	-	1.12	
Obj. risk below LB (share in %)	-	-	-	-	-	17.65	
Obj. risk below LB; mean distance betw. obj. risk & LB	-	-	-	-	-	2.15	
Observations	77	77	77	77	26	51	51

Notes: The table describes the relationship between initial and post-probe response in T3. Panel A describes the distribution of reported values, by treatment group and initial/post-probe question. LB denotes the lower bound and UB the upper bound for individuals that stated an interval in the post-probe question. Panel B presents the accuracy in risk prediction, by treatment group and initial/post-probe question.

Chapter 2

Early Child Care and Influenza Epidemics: Evidence from High Frequency Data

Abstract

Infectious diseases are a major burden on human health. Data from contact tracing show that child care facilities are among the most common locations where airborne diseases are spread. Yet empirical evidence on a causal impact of child care on the spread of infectious diseases is scarce. This paper provides evidence on the transmission of influenza as a consequence of the expansion of early child care in Germany. In a difference-in-differences analysis I show that a 1 percentage point rise in the child care coverage rate leads to a significant increase in the incidence rates of children by about 3%. To uncover the impact of the roll-out on age-specific transmission patterns, I develop a semi-structural model of disease diffusion based on epidemiological work. I find that disease transmission between children aged 0 to 2 and 3 to 6 significantly increases in response to the expansion of child care. A counterfactual analysis shows that policy interventions, such as mandatory vaccination before entry into child care and the closure of child care during local outbreaks, would significantly reduce infection rates by up to 11%. Besides, mandatory vaccination policies would be cost-effective decreasing net annual costs by about 20%.

2.1 Introduction

Infectious diseases are a major burden on human health. Seasonal influenza, for instance, infects more than 9% of the world's population each year and imposes an annual economic burden of approximately $16 billion in the United States and €800 million in Germany (Lambert and Fauci 2010; Molinari et al. 2007; Scholz et al. 2019). Exposure to infectious diseases is, therefore, costly for society through multiple channels. These include increased health utilization, loss in hours of schooling and work, long-lasting morbidity, and premature deaths (Almond and Mazumder 2005; Currie and Schwandt 2013; Kelly 2011; Schwandt 2017). Hence, a better understanding of the locations where people acquire infections is crucial in informing policy to enact targeted containment measures and thus, reduce costs on society.

Recent data from contact tracing show that the place number one where people get infected with airborne-transmitted diseases are private households, followed by educational institutions such as universities, schools, and child care facilities (Robert Koch Institute 2012, 2013).[1] Simultaneously, many OECD countries experienced a significant increase in the number of children attending public child care in the last two decades (OECD 2020). If child care facilities indeed spur the spread of infectious diseases, the expansion of public child care may impose unintended costs on society. Work in medical literature shows that formal child care is related to increased infection rates (e.g Ball et al. 2000; Ball et al. 2002; Côté et al. 2010). However, little is known whether child care facilities and their organizational structure, such as group sizes and the age structure of groups, causally affect the spread of infectious diseases.

This paper attempts to close that gap by providing causal evidence on the transmission of influenza as a consequence of the staggered roll-out of early child care – care offered to children under the age of 3 – in Germany. The study addresses two main questions: (i) What is the effect of the expansion of early child care on the spread of influenza? (ii) What are the economic benefits of policy interventions that aim to limit the spread of influenza within child care facilities? To do so, I consider two potential policy interventions: the closure of child care centers during outbreaks and mandatory vaccination before entry into public care.

To answer these questions, I use detailed high-frequency data recording daily cases of the incidence of influenza in Germany, covering the period from 2005 to 2016. According to German law, the diagnosis of influenza has to be reported to the national public health agency (Robert Koch Institute) in order to monitor the spread of diseases within Germany. The data include information on the day of diagnosis, county of residence, and date of birth of infected individuals. Importantly, the data allow me to distinguish between age groups and thus, to study whether and to which extent the expansion of formal child care has heterogeneous effects on age-specific incidence rates.

[1] Appendix Figure B.1 shows the locations of infection with the flu that could be pinned down by the German health authorities between 2005 and 2016 within Germany.

Besides, the data describe the spatial and temporal diffusion of disease incidence and thus, allow to uncover geographical differences in the impact of the roll-out on transmission patterns.

As a source of exogenous variation in child care, I exploit the staggered expansion of public child care centers targeted at children below the age of 3 in West Germany[2] between 2005 and 2016. During that time period Germany heavily invested into the expansion of formal child care. As a consequence, child care coverage rates increased on average by 20 percentage points (pp) per county over the time period. I argue that – conditional on time and region fixed effects – the timing of the roll-out is unlikely to vary with county characteristics that relate to disease incidence. This quasi-random variation results from varying and limited capacity of staff, infrastructure and financing at the municipality level.[3] My identification strategy is predicated on two empirical observations: First, the timing of the expansion is unrelated to the demographic and socio-economic regional characteristics prior to the roll-out. This has been established, for example, by Bauernschuster et al. (2016). Second, I find no empirical evidence that parental incentives to consult a physician, when a child falls ill, change as a co-movement of the roll-out. Hence, I can rule out that my results are driven by selective reporting behavior.

I first run a staggered difference-in-differences (DiD) analysis, investigating whether and to which extent yearly incidence rates change as a consequence of the expansion of child care. I find an economically and statistically significant effect of early child care on the incidence rates of children aged 0 to 2. More precisely, a 1pp rise in the child care coverage rate leads to an increase in the incidence rates by about 3%. The treatment effect is driven by urban areas, while I find no evidence for a significant impact of child care on infection rates in rural areas.

The DiD analysis provides reduced-form evidence on the impact of child care expansion on the spread of infectious diseases. Next, I develop a dynamic, semi-structural model of disease diffusion that builds on the Susceptible-Infected-Resistant Model (SIR) used in epidemiological work. The model extends the DiD analysis in at least two ways: First, it allows me to estimate spatial and temporal disease transmission patterns for various age groups and thus, to uncover policy-relevant margins of heterogeneity. To do so, I estimate age-specific transmission rates and analyze how these estimates respond to changes in child care coverage rates. Second, it provides me with a framework to study the effect of policy interventions on disease incidence and also to evaluate the economic benefits of these measures. Due to the dynamic structure of the model, a non-classical measurement error in incidence rates may lead to biased estimates (Adda 2016).[4] I address these endogeneity concerns using an instrumental variables (IV) approach. To this end, I create a novel

[2]The analysis only includes regions in former West Germany (without Berlin), as the child care system is still today substantially different in regions in the former German Democratic Republic compared to ones in former West Germany.

[3]The expansion primarily took place at the municipality level, while the analysis is conducted at the county level.

[4]A non-classical measurement error may arise, since incidence rates are more likely measured with an error, once infection rates are very high, and patients have to queue to see a physician. While a classical measurement error would lead to attenuation bias, Adda (2016) shows in his work, that due to the dynamic structure of SIR model a non-classical measurement error in the measure of infected and of susceptible individuals results in a complex error term, with serial dependence.

instrument based on lagged temperature. Temperature is a natural candidate to instrument incidence rates of influenza, as influenza viruses are known to survive longer under cold and dry conditions and weather conditions also determine socialization patterns (Deyle et al. 2016; Lowen and Steel 2014).

In line with prior epidemiological work, I find that disease transmission primarily occurs within the same age group. Children aged 0 to 2, however, are likely to spread influenza to children aged 3 to 6 and vice versa. The results show that transmission rates between these two age groups significantly increase in response to a rise in child care coverage rates. This effect may be driven by the institutional structure of child care facilities in Germany that often include mixed-age groups with children up to the age of 6. The expansion also leads to an increase in infection rates of adults providing evidence for spillover effects to the working population. In line with the findings of the DiD analysis, I document significant differences in the impact of child care provision between urban and rural areas on the spread of influenza. The findings suggest that there may exist systematic differences in the organizational structure of child care facilities between urban and rural areas that determine distinct transmission patterns.

In a subsequent step, I evaluate two counterfactual policy interventions using the framework of the SIR model that aim at limiting the spread of infectious diseases within child care facilities: mandatory vaccination before entry into child care and the closure of child care facilities during local outbreaks. I find that both policy measures significantly reduce infection rates by up to 11%. A cost-benefit analysis of mandatory vaccination shows that the measure would be cost-effective decreasing net annual costs by about 20% relative to the status quo.

This paper contributes to several strands of literature. First, it adds to a growing strand in economics that investigates the impact of human behavior on the spread of infectious diseases. These studies predominantly focus on the effect of economic activity and human mobility on disease transmission patterns (e.g. Adda 2016; Fang et al. 2020; Glaeser et al. 2020; Hufnagel et al. 2004; Oster 2005, 2012). While they study the incidence of different viruses such as HIV, influenza, gastritis and SARS-CoV-19 virus in various political and social environments (e.g. Sub-Saharan countries, Europe), they jointly show that travel intensity significantly increases infection rates. However, less attention has been paid to the contribution of educational institutions such as schools or public child care to the transmission of infectious diseases. Some studies show that school holidays not only reduce incidence rates of children, but also of adults (Adda 2016; Cauchemez et al. 2008). This finding underscores the importance to consider spillover effects to other age groups than children when evaluating the impact of educational institutions on disease transmission patterns. However, these studies use relatively broad age groups[5] and thus, can not draw conclusions on the effect of child care on infectious diseases. A few studies have analyzed the effect of child care reforms on infectious diseases (Baker et al. 2008, 2015; van den Berg and Siflinger 2020).

[5]For instance, Adda (2016) classifies the population into three age groups: children aged 0 to 18, adults aged 19 to 60, and the elderly aged above 60.

While these studies show that child care significantly increases the diagnosis of infectious diseases, they leave open a number of relevant questions addressed in my study, such as spillover effects on age groups other than children, geographical differences, and the impact of policy interventions that are likely to mitigate disease transmission within child care centers. Methodologically, my study expands the approach proposed by Adda (2016) by estimating age-specific transmission patterns.

Second, the study is closely related to literature in the field of epidemiology. There is a large strand in medical literature investigating the link between child care attendance and the diagnosis of infectious diseases, but these studies do often not consider the endogeneity of child care attendance (e.g. Ball et al. 2000; Ball et al. 2002; Côté et al. 2010).[6] My study identifies a causal effect of child care on disease transmission patterns by leveraging the staggered roll-out of child care provision in Germany.

More broadly, this paper adds to the literature evaluating the benefits of vaccination policies. A growing strand of literature documents that health interventions in childhood, including vaccination programs, have positive long-term effects on educational achievements, labor market performance, and cognitive skills (e.g. Baird et al. 2016; Bleakley 2007; Bütikofer and Salvanes 2018). However, large-scale health interventions often require major financial commitments. Systematic comparisons of the benefits and costs of each intervention may support policy makers to allocate scarce resources efficiently (Bärnighausen et al. 2014). Therefore, economic evaluations of health interventions play a major role in informing health policy. This study provides empirical evidence on the economic benefits of vaccination by evaluating the costs and benefits of mandatory vaccination against seasonal influenza before entry into child care.

The paper proceeds as follows. Section 2.2 describes the institutional background and the data. Section 2.3 summarizes the DiD strategy and presents reduced-form evidence. Section 2.4 outlines the model of epidemics and presents the results. In Section 2.5, I evaluate two counterfactual policy interventions. Section 2.6 discusses the empirical findings. Section 2.7 concludes.

[6] Closely related is the question whether children attending formal care contract more infectious diseases than those cared for at home, but acquire immunity from the increased number of infections that protects them later in life (e.g. Ball et al. 2000; Ball et al. 2002; van den Berg and Siflinger 2020). In the case of influenza, it is unlikely that acquired immunity gained in one year offers significant protection during the subsequent seasons, as influenza viruses are fast-mutating which prevents long-term immunization (Hay et al. 2001). Hence, even though it is unarguably an interesting question whether children attending formal child care obtain immunity earlier in life compared to children cared for at home and thus, are protected later in their life, it is not relevant in the case of influenza and will not be considered in this paper.

2.2 Institutional Background and Data

2.2.1 The Child Care System in Germany

Early Child Care System Public child care is provided at two levels in Germany. Early child care is available for children below the age of 3 and kindergarten is offered to children aged 3 to 6 (Felfe and Lalive 2018). The analysis focuses on early child care which is provided either by care centers or by extra-familial child minders (Strunz 2013). Roughly 85% of all children enrolled in early child care attend care centers. Child care facilities are typically operated by subsidized non-profit organizations such as municipalities, welfare organizations or the church (Felfe and Zierow 2018).[7] Public child care is heavily subsidized by municipalities, counties, and states. Parental fees cover only between 4% and 17% of the total operating costs (Bertelsmann Stiftung 2021b). There is, however, substantial variation in the fees which range from 0 to 400 Euro/month. Fees depend on state-specific regulations, family income, the number of siblings enrolled in child care, and the provided daily hours of child care (i.e. full-day slot vs. half-day slot) (Bertelsmann Stiftung 2018; Geis 2018). Demand for child care slots exceeds by far supply (Bien et al. 2006; Bundesministerium für Familie, Senioren, Frauen und Jugend 2020). In 2016, for instance, 28% of all children aged 0 to 2 were enrolled in child care, whereas 43% of all parents had expressed a desire to place their child in early child care in West Germany (Bertelsmann Stiftung 2021a).

In the context of infectious diseases, regulations on group sizes and the age composition within groups are particularly relevant, as they determine both the number of contacts with extra-familial members and transmission patterns across age groups. In Germany, children aged 0 to 2 can be either enrolled in groups comprising only children in this age cohort or in mixed-age groups with children up to the age of 6 (Bundesministerium für Familie, Senioren, Frauen und Jugend 2012).[8] Hence, the expansion of early child care slots is presumably not only affecting contact patterns among 0 to 2 year old children, but also socialization patterns between children aged 0 to 2 and children aged 3 to 6. The average group size depends on the age structure of groups. In 2013, for example, groups exclusively for children below the age of 3 consisted on average of 10 children, whereas groups with children between age 2 and age 6 comprised on average 21 children per group (Strunz 2013). It is further to point out that average group sizes remained rather constant over time suggesting that the increase in child care slots is predominantly achieved by creating new groups rather than increasing the number of children per group (Strunz 2013).

Children aged 0 to 2 years, who are not enrolled in early child care, are mainly cared for by their parents. Other care modes, such as paid informal care, care provision by grandparents or extended

[7] Only 3% of all child care centers are run by private, for-profit organizations (Bertelsmann Stiftung 2021c).

[8] There are 5 different types of group compositions: (i) groups targeted exclusively at children aged below 3 (31% in 2010) (ii) groups comprising children from age 2 to age 6 (21% in 2010) (iii) groups of children aged 0 to 6 (19% in 2010) (iv) groups of children aged 0 to 4 (23% in 2010) (v) care centers without any group structure (7% in 2010). Child care facilities without group structure can be either exclusively for children below the age of 3 or including children between 0 and 6 years.

family members, play only a minor role in Germany (Felfe and Lalive 2018).[9] Hence, the increase in the child care coverage rate can be interpreted as a shift from home-based care to group-based care.

Finally, it is to note that the child care system in Germany is characterized by large differences between East and West Germany that are historically constituted in the German partition after the Second World War. The former German Democratic Republic established a comprehensive child care system from the age of 0 onwards, while in West Germany child care was primarily a private matter (Bauernschuster et al. 2016; Bauernschuster and Schlotter 2015; Felfe and Lalive 2018). The differences in the child care systems between East and West Germany still continue to exist today, with substantially higher coverage rates in East than in West Germany (see Appendix Figure B.4). In the analysis, I focus on the staggered increase in child care provision in the 324 counties of former West Germany, excluding Berlin.

Expansion of Early Child Care In the early 2000s around 90% of all 3 to 6 year old children attended public child care, whereas coverage rates for 0 to 2 year old children were, particularly in West Germany very low (Camehl and Frauke 2017).

In 2005, the German government started to invest heavily into the expansion of early child care. A number of laws paved the way for an increase in the provision of subsidized child care. In 2005, the federal government decided to create 230,000 additional child care slots by 2010 (*Tagesbetreuungsausbaugesetz*). In April 2007, the government announced a target coverage rate of 35% nationwide by 2013. In 2008, the law on support for children (*Kinderförderungsgesetz*) reinforced the announcement by declaring that every child from age 1 year onwards would have a legal claim to an early child care slot by August 2013 (Felfe and Lalive 2018). As a consequence, Germany has witnessed a strong increase in the number of children aged 0 to 2 attending public child care institutions. In the period from 2005 to 2016, the average child care coverage rate increased from about 8% to 28% in West Germany (see Figure 2.1).

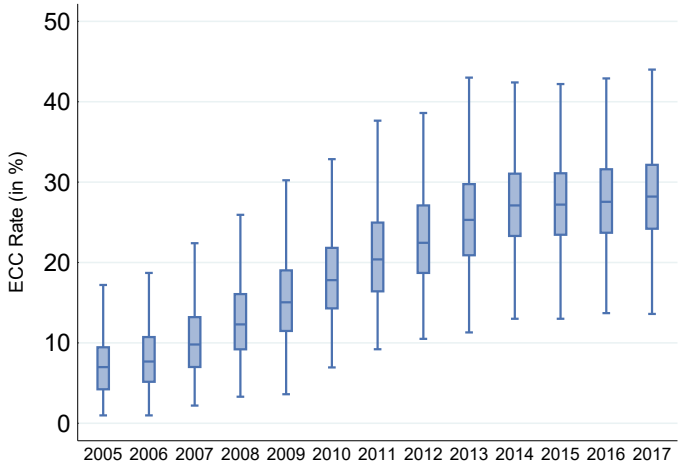
The expansion of child care facilities is financed by the federal government, the state governments, and local entities, i.e. municipalities and non-profit organizations (Diekmann and Thöne 2011).[10] Federal funds are allocated to states in line with the number of 0 to 2 year old children living in each state (*Gesetz über Finanzhilfen des Bundes zum Ausbau der Tagesbetreuung für Kinder*). To obtain funding, municipalities must submit expansion plans to the respective state government demonstrating both demand and a feasible implementation (Diekmann and Thöne 2011). A lack of infrastructure and staff present the main obstacles to submitting a persuading expansion plan. Strict regulations on the infrastructure limited the set of appropriate properties and prolonged the

[9] Felfe and Lalive (2018) report that children aged 2 to 3 years, who were not attending early child care, were cared for by informal modes of paid care for 3.1h per week, by their grandparents for 4.1h per week, and by other members of the extended family for 0.8h in 2011. The remaining hours per week are taken over by their parents.

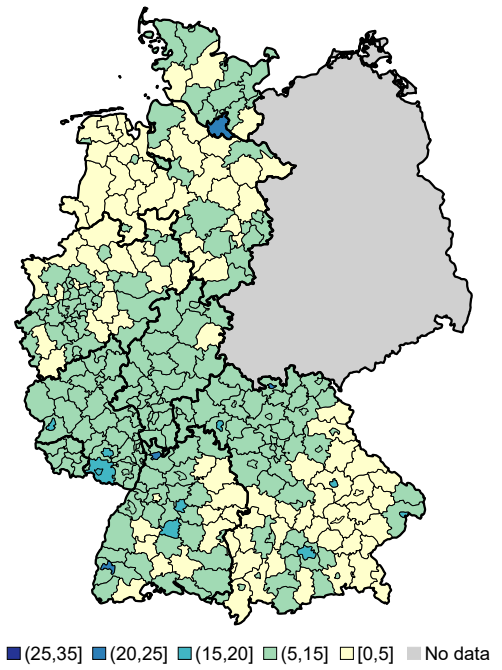
[10] The federal government enacted three investment programs, i.e. from 2008 to 2013 (2.15 billion Euro), from 2013 to 2014 (580 million Euro), and from 2015 to 2018 (550 million Euro).

Figure 2.1: Early Child Care Coverage Rate, West Germany

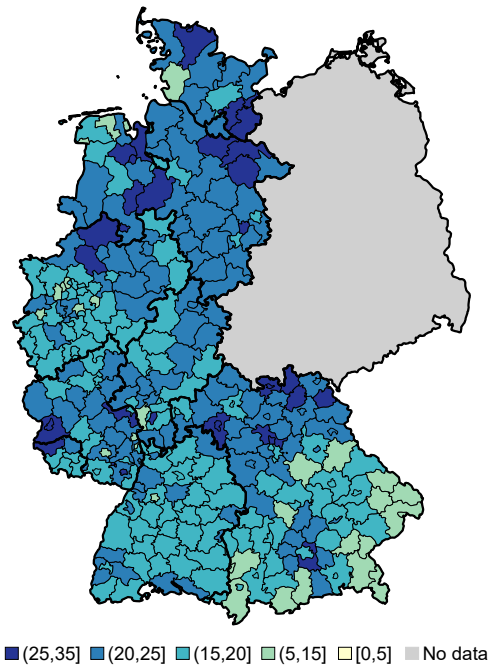
(a) ECC Rate from 2005 to 2016



(b) ECC Rate in 2005



(c) Change in ECC Rate from 2005 to 2016



Notes: The figures illustrate the ECC rate in West Germany between 2005 and 2016. Panel A displays the median, the 25th and 75th percentile of the annual child coverage rate. Panel B illustrates the ECC rate in percent across German counties in 2005 and Panel C shows the relative change in the ECC rates between 2005 and 2016 (in pp).

time until a municipality could submit a feasible expansion plan (Felfe and Lalive 2018). Further, Germany experienced a lack in early child care workers that hampered expansion.

As a consequence, expansion of early care facilities varied substantially across counties and over time. Figure 2.1 illustrates the temporal and spatial variation in the expansion of the child care provision. While the mean child care coverage rate increased by about 20pp in West Germany between 2005 and 2016, the child care coverage rate substantially differed across counties. In 2005, for instance, 21% of the children aged 0 to 2 were enrolled in public care in Hamburg, but only 4.5% of the children of the same age cohort attended formal child care in the neighboring county of Stade (see Figure 2.1 Panel B). Further, each county within West Germany is affected by the expansion, experiencing an increase in early child care slots of at least 8pp between 2005 and 2016 (see Figure 2.1 Panel C).

Administrative Data on Child Care Coverage Rate I use administrative data on child care provision from the Statistical Office Germany.[11] The data include information on the early child care coverage rate (ECC rate) that is defined as the number of children aged 0 to 2 attending formal child care per 100 children. Information on the ECC rate is available by year and county.[12]

2.2.2 Characteristics of Influenza and Incidence Data

Characteristics of Influenza Influenza is one of the major viral diseases worldwide. It is an airborne disease, infecting 5%-20% of the entire population per year. Among children the incidence is substantially higher with infection rates of 20%-35% (Amboss 2021). Symptoms typically include a sudden onset of fever, coughing, muscle pain, and headache. While most individuals recover from influenza within a few days, children and the elderly population are at particularly high risk of developing complications (e.g. acute otitis media, pneumonia) requiring hospitalization (Amboss 2021; Scholz et al. 2019).

Table 2.1 provides a brief overview of the characteristics of influenza, with information on incubation time, length of symptoms, and the period when individuals are contagious. Immunity can be acquired either by infection or by vaccination. As influenza viruses mutate very quickly, immunity typically lasts only for one influenza season. Hence, to maintain immunization, individuals have to be re-vaccinated each autumn. In Germany, as in most European countries, vaccination is targeted towards people above the age of 60, pregnant women, health care workers, and a small minority of people at risk (Robert Koch Institute 2020a). Coverage is relatively low. Despite of the target of achieving a nationwide coverage rate of 75% among people above the age of 60,

[11] See https://www.regionalstatistik.de/genesis/online/data and https:/www.inkar.de/

[12] Information on child care coverage rates is elicited each year in March. As children typically enter child care in August/September (Meiner et al. 2015), I construct yearly aggregates that start in calendar week 27 and end in calendar week 26 of the subsequent year. For instance, data collected in March 2007 is merged to weekly data on disease incidence from calendar week 27 in 2006 to calendar week 26 in 2007.

only about 30% of this age cohort is actually vaccinated (Reuss et al. 2010). Health authorities do not typically advise children to obtain a vaccination (European Centre for Disease Prevention and Control 2017; Robert Koch Institute 2020b). However, the World Health Organization's recommendation to vaccinate children has recently sparked a debate on this matter (World Health Organization 2020b).[13]

Table 2.1: Characteristics of Seasonal Influenza

Symptoms	fever, headache, coughing, muscle pain
Length of symptoms	5 to 7 days
Incubation time	1 to 2 days
Contagious phase	1 day before to 5 days after the onset of symptoms
Acquired immunity due to infection\vaccination	for one season
Infection fatality rate	0.1% - 0.2%
Annual incidence rates per 100,000, total pop. (in 2016)	50.04
Annual incidence rates per 100,000, < age 3 (in 2016)	155.64
Vaccine exists	yes
Vaccine recommended	age 60+ & chronic diseases
Vaccination rate, age 60+ (in 2016)	29.8%
Vaccination rate, < age 3 (in 2016)	no systematic vaccination
Compulsory to report lab confirmed diagnosis since	2001

Notes: The table summarizes the characteristics of influenza. Information on disease incidence and vaccination refers to the situation in Germany. Source: Amboss (2021) and Robert Koch Institute (2020b). The annual incidence rates are based on own calculations.

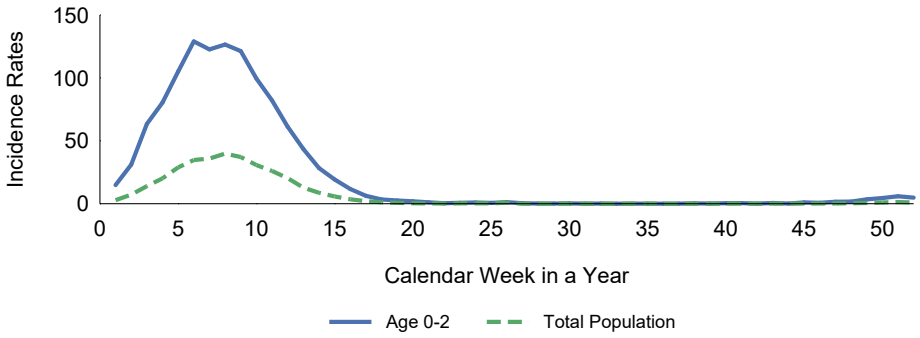
Influenza has a typical temporal pattern with annual recurrent peaks during the winter months and low incidence rates from May to September. Panel A in Figure 2.2 shows the seasonal pattern of influenza activity in Germany within a year, by calendar week, from the first week of January to the last week of December. The reported mean incidence rates for children below age 3 (blue line) are substantially higher than the average incidence rates of the entire population (green line), indicating that children aged 0 to 2 are profoundly more likely to be affected by influenza than people in older age cohorts.[14] Panel B in Figure 2.2 shows the time series patterns of incidence rates for West Germany on a monthly basis. The yearly incidence of influenza varies widely. There is ample variation across years with different viral strains giving rise to strong or weak seasons.

[13]The World Health Organization (2020b) recommends to vaccinate each child aged between 6 and 59 months against influenza. Some European countries, such as Austria, Finland or Poland, as well as, the USA, took over the advise by the World Health Organization and included vaccination against the flu for children from the age of 5 months onwards into the national recommended vaccination schedule.

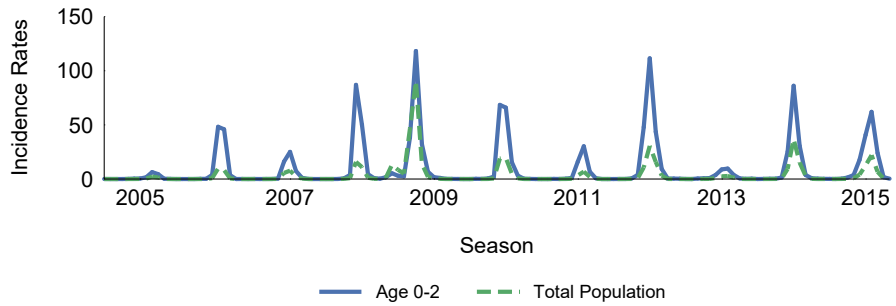
[14]The incidence rates presented here are based on the reported number of lab-diagnosed cases. Hence, the figure only gives suggestive evidence for higher incidence rates among children compared to the individuals in other age groups, as the testing rate might depend on the age of the infected.

Figure 2.2: Influenza - Seasonality and Time Trend

(a) Seasonality



(b) Time Trend



Notes: Panel A displays average weekly incidence rates (cases per 100,000 individuals) in West Germany averaged over the years between 2005 and 2016. Panel B illustrates the monthly incidence rates for the flu seasons 2005/06 to 2015/2016. The solid, blue line displays incidence rates of children aged 0 to 2 years, the dashed, green line refers to average incidence rates of the total population.

High Frequency Data on Infectious Diseases I use detailed, high-frequency data on the daily incidence of influenza in Germany, covering the period from summer 2005 to summer 2016. The data include laboratory confirmed cases of influenza (Robert Koch Institute 2016b). For each case reported, the data provide information on the day of diagnosis, county of residence, potential location of exposure,[15] hospitalization status, birth date, and gender. Besides, the data include information on child care attendance for a subset of children. This detailed data set allows me to analyze the spatial, temporal and age-specific evolution of disease transmission.

The data are provided by the German national public health agency, the Robert Koch Institute. According to the Protection Against Infection Act enacted in 2001 (*Infektionsschutzgesetz lfSG*),

[15] If the location of disease transmission is traceable, information on geographic and institutional location of exposure is included in the data set.

the diagnosis of a number of infectious diseases such as influenza, norovirus-gastritis or measles, has to be reported to the local public health departments in order to monitor the spread of diseases within Germany. The local health departments deliver the information on positive cases within 24 hours to the health authorities of the respective federal state. The state health department gathers the information and passes it on to the Robert Koch Institute (see Appendix Figure B.2 for a graphical illustration of the reporting system). To enhance comparability of information across regions, reported cases have to fulfill well-defined criteria to be included in the database. For instance, influenza cases are only added to the data set if the clinical diagnosis is confirmed by lab-diagnostics or alternatively, if a patient has flu-specific clinical symptoms and has had contact to a person who tested positive for influenza. Hence, the data clearly understate the scale of the actual cases, since asymptomatic cases and, symptomatic cases without a positive test or proven exposure to a positively tested individual are not recorded. This is, however, not a major concern for the analysis, as long as the reporting behavior is not systematically linked to the expansion of early child care (see discussion in Section 2.3.1). I supplement data on disease incidence with information on vaccination coverage rates of all publicly insured people aged 60 or above from administrative medical claims.[16]

2.2.3 Additional Data

I obtain information on daily temperature and precipitation from the German Meteorological Service[17] for the period from summer 2005 to summer 2016. I construct weekly aggregates to match the time dimension in the disease incidence data (see Appendix Section B1 for additional information on weather data). Further, to control for regional time-varying characteristics I use yearly data on demographic and socio-economic characteristics measured at the county level provided by the Statistical Office Germany. County characteristics include labor market and economic characteristics (e.g. female labor participation rate, unemployment rate, GDP per capita) and demographic characteristics, such as population density and population size. Appendix Table B.1 presents the summary statistics.

[16]Roughly 90% of the German population is covered by the social health insurance. For the seasons 2007/2008 to 2013/2014, the data is provided at the county level by the Zentralinstitut für die Kassenärztliche Versorgung in Deutschland (ZI) (Zentralinstitut für die Kassenärztliche Versorgung in Deutschland 2021a, 2021b). For the remaining seasons, I use data on vaccination rates published at the state level by the Robert Koch Institute (Reuss et al. 2010; Rieck et al. 2017, 2018; Robert Koch Institute 2016a).

[17]See https://opendata.dwd.de/climate_environment/CDC/observations_germany/climate/daily/kl/historical/. Accessed on 7th March 2021.

2.3 Staggered Difference-in-Differences Strategy

This section provides reduced-form evidence on the effect of the expansion of early child care on the spread of influenza using a DiD strategy. I first outline the identification strategy of the DiD analysis and present the results thereafter.

2.3.1 Empirical Strategy

Model Specification To examine the effect of early child care expansion on annual infection rates, I specify the following panel data regression:

$$I_{c,t,a} = \alpha + \beta \text{ECC rate}_{c,t} + X_{c,t}\gamma + \delta_c + \omega_t + \epsilon_{c,t} \tag{2.1}$$

where $I_{\text{t,c,a}}$ denotes the incidence rates (i.e. cases per 100,000 inhabitants of a particular age group) for individuals residing in county c, in year t and in age group a. ECC denotes the early child care coverage rate in county c and year t. $X_{\text{c,t}}$ represents a set of county-specific demographic and socio-economic characteristics. The model includes county fixed effects (δ_c) and year fixed effects (ω_t). The county indicators (δ_c) control for unobservable determinants of incidence rates that are time-invarying at the county level, while year fixed effects (ω_t) absorb common time shocks. The coefficient β captures the effect of a change in ECC rate on incidence rates. Error terms are clustered at the county level.

Identification Strategy As a source of exogenous variation, I exploit the staggered expansion of child care centers throughout West Germany. Because of limited capacity (i.e. staff/ infrastructure/ financing), the expansion progressed at different rates across counties. Identification is achieved by within county variation in early child care coverage rates over time. The identification strategy is based on the assumption that - conditional on county and year fixed effects - the timing of the roll-out is unlikely to vary with county characteristics that correlate with disease incidence.

A key threat to the identification of β is that the timing of the expansion is correlated with unobserved determinants of disease transmission. Potential regional characteristics that might violate the identifying assumption are counties' economic standing or female labor market participation, as they may impact both the timing of the expansion of early care facilities and disease transmission patterns.

Hence, to assess the identification strategy, I run a number of robustness tests: First, I investigate whether baseline county characteristics prior to the reform relate to the timing of the expansion of child care slots (see Appendix B2). Overall, the results show that the timing of the expansion does not correlate with baseline economic standing (i.e. unemployment rate, average household

income, GDP per capita) and demographic features (i.e. population size, population density). I find, however, that female labor participation and child care coverage in 2002 are related to the expansion of child care between 2006 and 2007. From 2007 onward, there appears to be no systematic relationship between the timing of the roll-out and these variables. The results align with previous studies, for example Bauernschuster et al. (2016), showing that the timing of the expansion is unrelated to demographic and socio-economic regional characteristics prior to the roll-out. Second, I regress child care coverage rates on county and year fixed effects, as well as on time-varying covariates. I find that 95% of the variation in the adjusted R^2 of the child care coverage rate can be linked to common time shocks and time-invariant county characteristics, whereas only 1% of the variation in the adjusted R^2 can be attributed to a large set of covariates.[18] Third, there might be co-movements in variables over the period of observation that confound with the number of reported cases of influenza. A major concern is that an increase in female labor market participation induced by the expansion of child care facilities may affect reporting behavior, as parental choice to consult a physician, if their child falls ill, may depend on their employment status. German sick leave regulations give parents the right to stay absent from work to care for their sick child, but require that the disease is confirmed by a physician.[19] Hence, it is conceivable that a change in reported cases is attributable to a change in reporting behavior. To tackle the concern that child care attendance is systematically related to parental choice of consulting a physician, I use data from a representative survey on child health (KiGGs study[20]). In the study, parents report the annual number of medical consultations, the annual number of influenza-like diseases, as well as child care attendance. The annual number of medical consultations for children conditional on the number of reported flu-like diseases per year is not significantly different for children attending formal child care than for children cared for at home (p-value = 0.36). The results suggest that the likelihood of visiting a physician if a child falls ill, does not depend on child care attendance (see Appendix Table B.2).

2.3.2 Results

Table 2.2 reports the impact of an increase in the ECC rate on annual incidence rates of children aged 0 to 2 and of the entire population using the regression model specified in Equation 2.1. The

[18] The set of time-varying control variables includes demographic characteristics (population density, population size), socio-economic features (i.e. average household income, female labor market participation, unemployment rate).

[19] According to German sick leave regulations, parents have the right to stay at home to care for their child if the child is below the age of 12 and no other adult can care for the sick child. If parents are publicly insured, each parent is entitled to stay at home for 10 days per calendar year and per child, for single parents it is 20 days. If parents have more than 2 children, the total number of days on sick leave are 25 days per parent. However, to be eligible for sick leave, it is necessary that the disease is confirmed by a physician (§45 SGB V). Work contracts may, however, prelude the possibility of paid leave (§616 BGB). If parents are not entitled to paid exemptions, but are publicly health insured, they obtain sick pay provided by the health insurance, that is commonly 70% of the regular income. In this case, sick leave can be accompanied with a loss in income. Hence, it is ex ante not clear whether the likelihood to see a physician increases, decreases or remains unchanged in the context of an increase in child care attendance.

[20] The KiGGs survey was conducted by the Robert Koch Institute and includes information on child health and socio-economic characteristics gathered between 2003 and 2006 and between 2014 and 2016.

reported estimates of β present the odds ratios derived by Poisson pseudo maximum likelihood regression. Thus, the coefficients can be interpreted as factors.

I find a significant effect of early child care on the incidence rates of influenza of children aged 0 to 2. More specifically, a 1pp rise in the ECC rate leads to an increase in the incidence rates by about 3% (Columns 1 and 2). Importantly, the impact of the ECC rate on infection rates of children remains significant at the 5% level when controlling for potential confounding mechanisms such as female labor participation or employment rates. There is no evidence for a significant effect of child care on incidence rates of the entire population in the full sample (Columns 5 and 6).

Table 2.2: DiD Estimates

	Age 0 to 2				**Total Population**			
	Full Sample		**Rural Area**	**Urban Area**	**Full Sample**		**Rural Area**	**Urban Area**
ECC Rate	1.047**	1.033**	1.001	1.050***	1.019	1.015	0.974	1.048***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Observations	3564	3564	1562	2002	3564	3564	1562	2002
Pseudo R^2	0.634	0.639	0.606	0.662	0.751	0.754	0.763	0.755
Season FE	yes	yes	yes	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes	yes	yes	yes
Control Variables	no	yes	yes	yes	no	yes	yes	yes

Notes: The table displays DiD estimates based on Poisson pseudo maximum likelihood regression using Equation 2.1. The dependent variables are the incidence rates of a particular age group. Coefficients are presented as odds ratios. Regressions are weighted by population size in a particular age group. The set of control variables includes temperature, rain, absolute humidity, GDP per capita, employment rate, female employment rate, and population density. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Next, I investigate whether the impact of child care on infection rates differs between urban and rural areas.[21] I find that the treatment effect is driven by urban areas. Urban areas experience a rise in the incidence rates by about 5% per percentage point increase in the ECC rate, while there is no evidence for a significant impact of child care on infection rates in rural areas (Columns 3 and 4). Further, I find that the expansion of child care in urban areas does not only significantly increase infection rates of children, but also of the entire population (Columns 7 and 8. See Section 2.6 for a discussion on different transmission patterns in rural and urban counties).[22] In summary, the DiD analysis suggests a significant impact of early child care on infection rates. The treatment effect seems to be driven by urban areas.

[21] I refer to the classification into urban and rural areas proposed by the Federal Institute for Research on Building, Urban Affairs and Spatial Development (*Bundesinstitut für Bau-, Stadt- und Raumforschung*) (BBSR 2019). "Urban Areas" denote *Kreisfreie Großstädte* with more than 100,000 inhabitants and *Städtische Kreise* with a population density of at least 150 inhabitants/km^2 and at least 50% of the population living in a city. "Rural Areas" denote the remaining areas. See Appendix Figure B.3 for a graphical illustration of urban and rural areas within West Germany.

[22] Appendix Table B.3 and Appendix B.4 present results on DiD estimates for different age groups, stratified by urban and rural areas.

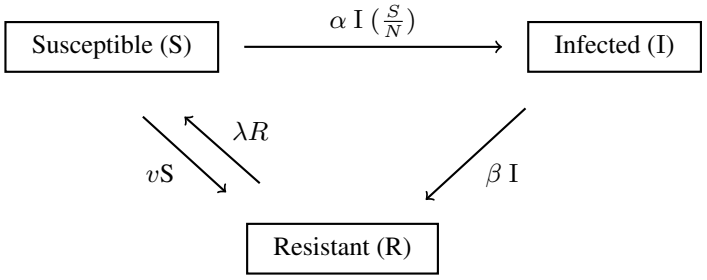
2.4 Model of Epidemics

The DiD analysis provides reduced-form evidence on a significant impact of the expansion of child care on infection rates. To identify the effect of child care on age-specific transmission patterns, I develop a semi-parametric model based on the Susceptible-Infectious-Resistant (SIR) model used in epidemiological work. This section presents the model of epidemics, the identification strategy, and empirical findings.

2.4.1 Description of the SIR Model

The SIR model dates back to Kermack and McKendrick (1927) and describes the dynamics of epidemics. In the SIR model the population is divided into three mutually exclusive groups: (i) A class of individuals that contracted the disease and is now infectious, called infected individuals I, (ii) A class of individuals who are healthy, but can contract the disease, the susceptibles S, and (iii) A class of individuals that acquired immunity either by vaccination or by infection, the resistant R. Figure 2.3 describes how individuals move from one group to another over time.

Figure 2.3: Flow Chart - SIR Model



Notes: The figure displays how infected, susceptible and resistant individuals move from one group to another per unit of time. The parameter α denotes the transmission rate, β the recovery rate, v the vaccination rate and λ the fraction of individuals that lose immunity per unit of time.

I make the following two assumptions: First, I abstract from births and deaths and second, I assume that the society is closed in a given year ($N = I + S + R$). The stock of new infected individuals depends on the number of currently infected individuals I, the transmission rate α, and the fraction of susceptible individuals in the population ($\frac{S}{N}$), i.e. the probability that a randomly selected contact is a susceptible individual. Individuals recover from the disease at rate β. Let v denote the vaccination rate and λ the fraction of individuals that lose immunity. For influenza the likelihood of getting infected twice within the same influenza season[23] is rather low. Hence, I assume that immunity lasts for one influenza season, until a new epidemic starts.

[23] I define an influenza season to last one year, starting each year in calendar week 27 when the infection rates are at the lowest level.

I apply a spatial, discrete-time version of the SIR model (Finkenstädt and Grenfell 2000) and construct weekly aggregates of disease incidence. Due to disease-specific characteristics of influenza[24], it is plausible to assume that infected individuals recover – in the sense of not being infectious anymore – within one week (i.e. β=1). Hence, I obtain the following specification that describes disease transmission patterns across weeks:

$$\underbrace{I_{t+1}}_{\substack{\text{number of}\\ \text{individuals}\\ \text{that get infected}}} = \underbrace{\underbrace{cr_t}_{\substack{\text{contact}\\ \text{rate}}} \times \underbrace{p}_{\substack{\text{prob. that contact}\\ \text{results in a transmission}}}}_{\textbf{transmission rate } \alpha} \times \underbrace{I_t}_{\substack{\text{number}\\ \text{of infectious individuals}}} \times \underbrace{(\frac{S_t}{N_t})}_{\substack{\text{prob. that}\\ \text{contact is a susceptible}}} \tag{2.2}$$

Equation 2.2 describes the stock of new infected individuals in week $t+1$ conditional on the transmission rate α, the stock of infected in week t, and the fraction of susceptibles in week t. The transmission rate α can further be divided into two components, namely the contact rate cr (i.e. the number of contacts per unit of time t) and the probability p that a contact results in a transmission. The contact rate cr is a behavioral parameter, as infected individuals choose whether and to which extent they meet susceptibles (and vice versa). The increase in child care slots may plausibly affect contact rates among children and thus, disease transmission patterns, as children attending formal child care may have contact to a larger number of individuals in addition to the family compared to children cared for at home. The parameter p is predominately a biological, disease-specific factor indicating how contagious a disease is. It can, however, also be affected by human-behavior, e.g. hand-washing and other hygienic behavior. Hence, the roll-out of child care centers might also impact the parameter p by inducing a change in child hygienic behavior. Therefore, it is plausible to assume that both contact rate cr and the transmission parameter p are a function of the regional child care coverage rate.

Econometric Model: Age-Stratified Spatial-Time Spread of Diseases Next, I introduce a spatial dimension to Equation 2.2. Social mixing behavior within counties is likely to be different from the one between counties. Hence, I estimate separate transmission rates for within (α^{WR}) and between regional (α^{BR}) spread. An individual can get infected by individuals residing in the same county, by individuals residing in any other county within West Germany or by an endemic component that captures, e.g. the risk of imported cases from outside the study region or seasonal effects. I follow epidemiological work by using a classic competing risks framework in which the forces of infection – within regions, across regions, and the endemic component – are additive (e.g. Bauer and Wakefield 2018).

[24]Note that the infectious period lasts less than a week and infection fatality rates (IFR) are negligible small for influenza ($\sim$ 0.1% to 0.3%).

Equation 2.3 describes disease transmission patterns between infectives in age group a' and susceptibles in age group a.[25] To account for differences in the population size across counties and age groups, I take the proportion of infected ($\frac{I}{N}$) and the proportion of susceptible individuals ($\frac{S}{N}$) rather than the stock of infected and susceptibles. Normalizing the susceptibles by the population allows me to interpret the presented coefficients as the marginal effects of a change in the infection rate on future infection rates, when the entire population is susceptible to the disease (i.e. basic reproduction rate R_0[26]). I further assume that the contact rate is proportional to the ratio of the size of the population containing infected individuals to the population size containing susceptible individuals ($\frac{N_{a'}}{N_a}$). Intuitively, this means that the likelihood that a susceptible meets an infected is higher, if the population size containing infected individuals is relatively large compared to the population size of susceptibles.

$$\begin{aligned}\frac{I_{t,c,a}}{N_{t,c,a}} = & \underbrace{\sum_{a' \in A} \alpha^{\text{WR}}_{a,a'} \quad \frac{I_{t-\tau,c,a'}}{N_{t-\tau,c,a'}} \quad \frac{S_{t-\tau,c,a}}{N_{t-\tau,c,a}} \quad \frac{N_{t-\tau,c,a'}}{N_{t-\tau,c,a}}}_{\text{Within regional spread}} \\ & + \underbrace{\alpha^{\text{BR}}_{a,a'=total} \quad \frac{\sum_{r \in R \setminus c} \frac{1}{d_{r,c}} \frac{I_{t-\tau,r,a'=total}}{N_{t-\tau,c,a'=total}}}{\sum_{r \in R \setminus c} \frac{1}{d_{r,c}}} \quad \frac{S_{t-\tau,c,a}}{N_{t-\tau,c,a}} \frac{N_{t-\tau,c,a'=total}}{N_{t-\tau,c,a}}}_{\text{Between regional spread}} \\ & + \underbrace{\delta_c + \phi_t}_{\substack{\text{Endemic component} \\ \text{Region + Time FE}}} + \epsilon_{c,t}\end{aligned} \tag{2.3}$$

The parameter α^{WR} captures age-specific disease transmission patterns between infectives in age group a' and susceptibles in age group a, both living in county c. The parameter α^{BR} measures disease transmission between individuals living in different counties. In the case of between-county spread, I do not differentiate between the infectives' age. More precisely, I estimate the effect that infectives of any age living in county $r \in R \setminus c$ spread influenza to susceptibles in age group a living in county c. I normalize the incidence of infectives living outside county c by the population in county c.[27] This can be interpreted as the likelihood that susceptibles living in county

[25] I stratify the population in five age groups: (i) age 0 to 2 (ii) age 3 to 6 (iii) age 7 to 26 (iv) age 27 to 60, and (v) above age 60.

[26] The basic reproduction rate is defined as the average number of secondary cases generated per infected in a completely susceptible population.

[27] To this end, I divide the sum of infected in all counties $r \in R \setminus c$ by the population in county c.

c meet infectives from other counties $r \in R \setminus c$ in county c.[28] I further account for differences in the distance ($d_{\text{r,c}}$) between each county $r \in R \setminus c$ and county c.[29] In other words, I assume that individuals living in counties close to one another meet more frequently than individuals living further apart.[30] The parameter τ describes the incubation time and is set to τ=1, as the incubation time for influenza lasts less than one week. I further include week and county fixed effects that capture the endemic component.

Effect of Child Care on the Transmission Rate To tackle the question whether and to which extent an increase in the child care coverage rate affects transmission rates within a county, I decompose the transmission rate using the following additive functional form:

$$\alpha^{\text{WR}}_{a',a} = \alpha^{0}_{a',a} + \alpha^{1}_{a',a}\text{ECC Rate}_{t-1,c} + \sum_{k=2}^{K} \alpha^{\text{k}}_{a',a} X_{t-1,c} \tag{2.4}$$

where $\alpha^{1}_{a',a}$ captures the effect of the child care coverage rate (ECC rate) on the age-specific transmission rate of influenza within a county. I denote a set of $K-1$ county-specific variables $X_{\text{t-1,c}}$ that may affect the spread of the flu. These variables include meteorological variables (temperature, precipitation, and absolute humidity), population density and measures for economic activity (i.e. female labor market participation, GDP per capita). In addition, I control for season-specific transmission rates and calendar month-specific transmission rates.

Computation of the Stock of Susceptible Individuals To estimate Equation 2.3, I compute the stock of the susceptible population S for each county c, week t, and age group a. To this end, I impose the following assumptions: First, immunity against influenza lasts one season, until the next epidemic starts. Second, the effectiveness of vaccination amounts to 60%.[31] Third, I assume that vaccination rates for individuals below the age of 60 are zero (see data description in

[28] The spread of diseases *between* counties can be modelled in two different ways: First, one can assume that infections occur *in county c in that the susceptibles live* (option 1). In this case, it is assumed that one infective living in county r travels to county c and spreads the disease to susceptibles residing in county c. Second, vice versa, one can assume that infections between susceptibles and infectives occur *in county r in that the infectives live* (option 2). The key difference is that the sum of infected outside county c is either divided by the population in county c (option 1) or by the population in county r (option 2). I specify Equation 2.3 in line with option 1, assuming that infections between individuals living in different counties occur in county c. However, the model could be easily rewritten such that it aligns with option 2. Results using option 1 and option 2 are similar.

[29] The sum of infected living outside c is weighted by the inverse of the distance between county r and c. More precisely, I use the inverse distance-weighting matrix that is defined by $w_{c,r} = (\sum_{r \in R \setminus c} \frac{\frac{I_r}{N_c}}{\text{distance}_{c,r}}) / (\sum_{r \in R \setminus c} \frac{1}{\text{distance}_{c,r}})$.

[30] Prior work in epidemiology shows that most contacts between individuals indeed occur within a kilometer of the individuals' home. They document that the distribution of travel distance decays as power law, indicating that it is plausible to use the inverse distance weighting matrix as a proxy for social interaction between individuals living in different regions (Brockmann et al. 2006; Read et al. 2012).

[31] The effectiveness of vaccination varies annually, on average it amounts between 40% to 60% (Center fors Disease Control and Prevention 2020).

Section 2.2.2). Hence, the stock of susceptible individuals in a given week is described by taking the entire population minus those who are vaccinated and those who have been infected in the respective influenza season.

2.4.2 Identification Strategy

In addition to potential threats to the identification strategy discussed in Section 2.3.1, OLS coefficients estimated using Equation 2.3 may be biased due to a non-classical measurement error. As mentioned above, the number of reported cases certainly understates the scale of influenza infections, i.e. not every infected person sees a physician and not every physician delivers a report. Besides, incidence rates are more likely measured with an error when infection rates are very high, as patients may have to queue to see a physician (Adda 2016). This may result in a non-classical measurement error. While a classical measurement error would lead to attenuation bias, Adda (2016) shows, that due to the dynamic structure of the SIR model, a non-classical measurement error in the measure of infected and susceptible individuals results in a complex error term with serial correlation. The overall structure of the error term makes it impossible to assess the direction of the bias, as the exact form of the error term will vary across diseases and age groups. Hence, to obtain consistent estimates of transmission parameters, I use lagged temperature as instrument. More specifically, I instrument the interaction terms denoted in Equation 2.3 as (i) within regional spread and (ii) between regional spread.

Lagged Temperature as Instrument I construct two instruments based on lagged temperature. One instruments for disease transmission within a region (i.e. IV^{WR}, Equation 2.5a), the other one instruments for disease transmission across regions (i.e. IV^{BR}, Equation 2.5b).

$$IV^{WR}_{c,t-1} = \underbrace{\Big(\sum_{i=1}^{i=t-3} \text{median temperature}_{\,c} - \text{temperature}_{i,c}\Big)}_{\substack{\text{Part 1}\\ \text{Epidemic curve in county c}}} \times \underbrace{(\eta + \text{temperature}_{t-2,c})}_{\substack{\text{Part 2}\\ \text{Recent temperature in c}}} \tag{2.5a}$$

$$IV^{BR}_{c,t-1} = \underbrace{\Big(\sum_{i=1}^{i=t-3} \text{median temperature}_{\text{r}\in\text{R}\setminus\text{c}} - \sum_{r\in R\setminus c} \frac{\frac{1}{d_{r,c}} \times \text{temperature}_{i,r}}{\frac{1}{d_{r,c}}}\Big)}_{\substack{\text{Part 1}\\ \text{Epidemic curve in all other counties except c}}} \times \underbrace{(\eta + \text{temperature}_{t-2,c})}_{\substack{\text{Part 2}\\ \text{Recent temperature in c}}} \tag{2.5b}$$

Each instrument consists of two parts: The first part sums over the difference between the median temperature during an epidemic outbreak and the temperature in week i of the outbreak. It presents a humped-shaped curve that grows, as long as the temperature is below the median temperature and declines once the temperature exceeds the median temperature, reflecting the typical dynamic pattern of seasonal influenza outbreaks (see Appendix Figure B.7). The onset of an epidemic

outbreak is defined as the first week of at least three consecutive weeks in which the incidence rate is larger than 0.5 per 100,000 inhabitants. Analogously, the end of an epidemic outbreak is defined as the first of at least five subsequent weeks in that the incidence rate is lower than 0.5 per 100,000 inhabitants. The sum is set to zero if (i) no outbreak takes place or (ii) the sum is below zero. The second part allows to shift the curve up and down depending on temperature in $t-2$. To instrument between regional spread, I interact the first part with the temperature in $t-2$ in county c based on the assumption that infections between individuals living in different counties always occur in county c, in which the susceptibles live.

The instruments must fulfill two conditions: First, they must be strong predictors for the interaction between infectives and susceptibles in $t-1$ (i.e. relevance condition). Temperature appears to be a natural candidate to instrument disease transmission patterns of influenza for biological and behavioral reasons: It is well documented that influenza viruses are more stable under cold and dry conditions than in warm temperature, which explains the seasonal pattern of this diseases (Deyle et al. 2016; Lowen and Steel 2014).[32] Weather is also likely to affect socializing patterns and thus, transmission rates. The results from the first stage regressions show that the relevance condition is satisfied with p-values of the Sanderson-Windmeijer F-Statistic well below 1% (see Appendix Table B.6). Second, the instruments need to be orthogonal to unobserved components of infectives in t conditional on county and week fixed effects (i.e. exclusion restriction). Exogenous week-to-week variation in weather conditions are plausibly unrelated to infection rate in t, except through disease transmission pattern in $t-1$.

2.4.3 Results

This section summarizes the empirical results based on the SIR model. I proceed in two steps: First, I use a simplified version of the model described in Equation 2.3 in which I differentiate only between the susceptibles' age, but not between the infectives' age (see Econometric Model 1). Second, I run an extended model of disease diffusion investigating transmission patterns within and across age groups of the infective and the susceptible population (see Econometric Model 2).

Econometric Model 1: Spatial-Time Spread of Diseases Table 2.3 presents the results considering disease transmission between infectives of any age and susceptibles in a particular age group.

[32] Various mechanisms are discussed in the medical literature, such as increased virus half-life at lower temperatures and more efficient transmission paths via aerosols or respiratory droplets under dry and cold conditions.

Table 2.3: Econometric Model 1: Spatial-Time Spread of Diseases

	OLS Results						IV Results					
	Age<3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**	**Age<3**	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
Panel A: Transmission Rates												
Within Regional Spread (α^{WR})	0.0299***	0.0724***	0.2474***	0.1777***	0.0574***	0.6059***	0.0806***	0.1704***	0.2378***	0.3114***	0.1449***	0.9617***
	(0.003)	(0.006)	(0.015)	(0.012)	(0.006)	(0.030)	(0.006)	(0.011)	(0.010)	(0.014)	(0.014)	(0.013)
Between Regional Spread (α^{BR})	0.0008	0.0084*	0.0661***	0.0431***	0.0120***	0.1368***	-0.0169	0.0172	0.0043	0.1215***	0.0574**	0.1813***
	(0.003)	(0.004)	(0.012)	(0.008)	(0.004)	(0.021)	(0.011)	(0.021)	(0.018)	(0.027)	(0.023)	(0.038)
Observations	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004
Adj.R2.	0.323	0.381	0.577	0.598	0.325	0.666						
Endogeneity Test							92.023	64.960	11.008	53.942	55.091	52.473
p-value - Endog. Test							0.000	0.000	0.004	0.000	0.000	0.000
Panel B: Effect of ECC Rate on Transmission Rate												
α^{WR} x ECC rate	0.00058**	0.00052	0.0011	0.0028***	0.00072	0.0060***	0.0012**	0.0011	-0.00065	-0.00026	-0.0013	0.00028
	(0.0003)	(0.0005)	(0.0012)	(0.0007)	(0.0006)	(0.0022)	(0.0006)	(0.0007)	(0.0008)	(0.0011)	(0.0014)	(0.0014)
Observations	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004
Adj.R2.	0.391	0.448	0.621	0.664	0.470	0.700						
Panel C: Urban Areas - Effect of ECC Rate on Transmission Rate												
α^{WR} x ECC rate	0.00080**	0.0014**	0.0011	0.0034***	0.00064	0.0074**	0.0017**	0.00089	-0.00023	0.000023	-0.00023	-0.000020
	(0.0004)	(0.0006)	(0.0015)	(0.0010)	(0.0009)	(0.0031)	(0.0008)	(0.0008)	(0.0002)	(0.0001)	(0.0002)	(0.0001)
Observations	103922	103922	103922	103922	103922	103922	103922	103922	103922	103922	103922	103922
Adj.R2.	0.446	0.514	0.617	0.677	0.483	0.698						
Panel D: Rural Areas - Effect of ECC Rate on Transmission Rate												
α^{WR} x ECC rate	0.00035	-0.00036	0.0012	0.0023**	0.00069	0.0049	0.00020	0.0000088	-0.00082	-0.00071	0.000087	-0.00098
	(0.0002)	(0.0006)	(0.0020)	(0.0010)	(0.0006)	(0.0030)	(0.0005)	(0.0009)	(0.0013)	(0.0015)	(0.0018)	(0.0020)
Observations	81082	81082	81082	81082	81082	81082	81082	81082	81082	81082	81082	81082
Adj.R2.	0.329	0.390	0.639	0.660	0.465	0.710						

Notes: The table displays estimates of Equation 2.3 in that one infected individual of any age spreads influenza to susceptibles in a particular age group. Columns 1-6 are based on OLS estimation, Columns 7-12 on IV regressions. Each column presents estimates from a different regression. The dependent variables are the incidence rates of a particular age group. Panel A displays estimated transmission rates within a county (α^{WR}) and between counties (α^{BR}). Panel B to D present the change in transmission rates in response to an increase in the ECC rate in the total sample (Panel B), in urban areas (Panel C) and in rural areas (Panel D). All regressions include county fixed effects and week-in-a-year fixed effects. In addition, regressions presented in Panel B to D control for season dummies, month dummies, climate, GDP per capita and distance to an airport interacted with transmission rates within a county. The complete set of regression coefficients is presented in Appendix Table B.7 to Appendix Table B.12. Regressions are weighted by population size in a particular age group. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Panel A summarizes the estimated transmission rates within a region (α^{WR}) and between regions (α^{BR}) using OLS regression (Columns 1 to 6) and an instrumental variable approach (Columns 7 to 12). The OLS estimates show that 100 infected individuals transmit influenza to 74 susceptibles (Column 6). This corresponds to an estimate of the basic reproduction rate R=0.74 with R defined as the average number of secondary cases generated per infected in a completely susceptible population. While some epidemiological studies document a basic reproduction rate for seasonal influenza below 1, the median estimate of the basic reproduction rate for seasonal influenza is 1.28 (Biggerstaff et al. 2014). One reason for the discrepancy between the estimated reproduction rate in this study and the median value documented in the literature is that OLS coefficients may be biased due to a non-classical measurement error. The Hausman tests displayed at the bottom of Panel A show that the exogeneity of the regressors is indeed rejected for all age groups. Compared with the OLS results, transmission rates estimated based on the IV strategy tend to have a larger effect size. In the IV framework, the estimated R-value is 1.14 indicating that infections spread exponentially (Column 12). The results further illustrate that individuals of all age groups get infected at a significant rate from people living in the same county ($p<0.01$).[33] The risk of infection from people living outside the county seems to be negligibly small for younger age cohorts, while it plays an important role for older age cohorts.[34]

Next, I study how the estimated transmission rates change in response to an increase in the ECC rate. To this end, I interact weekly transmission rates (α^{WR}) with annual ECC rates. Panel B summarizes the results including all counties in West Germany. The coefficients can be interpreted as the additional number of individuals that get infected by one infected individual due to a 1pp increase in the ECC rate. I find that transmission rates of children aged 0 to 2 significantly increase in response to an increase in the ECC rate. The results suggest that 100 infected individuals transmit influenza to 2.4 additional children due the increase of the average child care coverage rate by 20pp between 2005 and 2016 (Column 7).[35] Besides, the effect size of the expansion on infection rates of children aged 3 to 6 is comparable to the one of children aged 0 to 2, but not significant.

Panel C and D represent the effect of the expansion of child care on disease transmission, separately for urban and rural areas. In line with the results of the DiD analysis, I observe substantial differences in transmission patterns between urban and rural regions. While the expansion in child care significantly increases the spread of influenza in urban areas, I find no evidence for an impact of the roll-out on infection rates in rural areas.

[33] The IV estimates of within regional transmission rates show that 100 individuals infected with influenza spread the disease to 8 children aged 0 to 2, 17 children aged 3 to 6, 24 individuals aged 7 to 26, 31 individuals aged 27 to 60 and 15 individuals above 60 living in the same county as the infected.

[34] The IV estimates of between regional transmission rates present that increase in the average number of infected living outside a county by a 100 infections results in 12 additional infections of individuals aged 27 to 60 and 5 additional infections of individuals above 60 within a county of interest.

[35] The IV estimates show that 100 infected individuals spread influenza to additional 0.12 children in response to 1pp increase in the early child care coverage rate.

Econometric Model 2: Age-Stratified Spatial-Time Spread of Diseases I now consider an extended model of disease diffusion which differentiates between age groups of the infective and the susceptible population. I first estimate transmission rates between infectives and susceptibles within and across age groups and then, investigate how these estimates change due to an increase in early child care coverage rates. Estimates are based on OLS regressions, as the instruments can not be applied to different age groups of infectives. The results of the extended model are consistent with the findings in the IV framework based on econometric model 1. Both model specifications illustrate that mainly infection rates of children aged 0 to 2 and children aged 3 to 6 increase in response to the roll-out.[36]

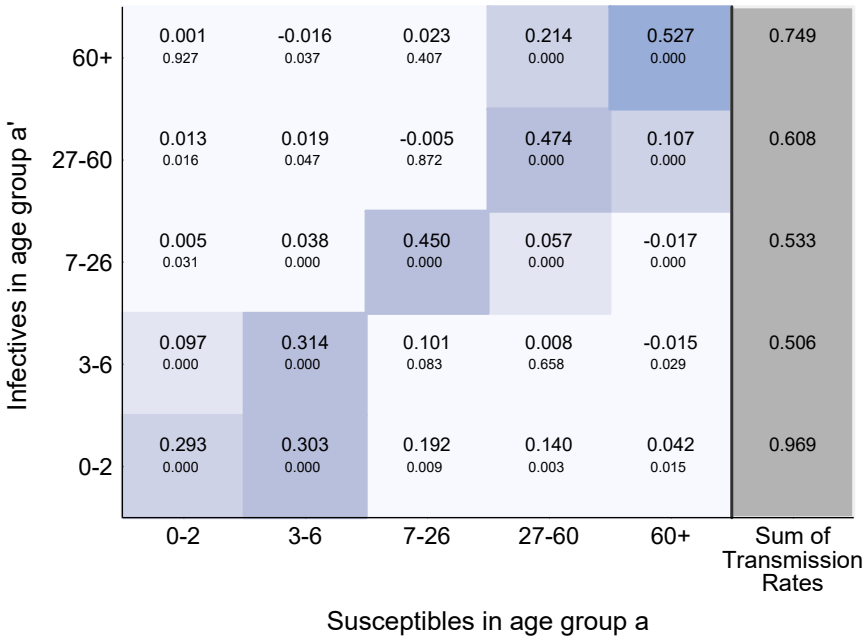
Figure 2.4 and Appendix Table B.15 summarize the results. Panel A in Figure 2.4 displays the estimated age-specific transmission rates per week within a county ($\alpha^{\text{WR}}_{a'a}$) and the respective p-values. The y-axis presents the infectives' age group a', while the x-axis denotes the susceptibles' age group a. For instance, 100 infected children aged 0 to 2 transmit influenza to approximately 30 children aged 3 to 6 years living in the same county within a week. Panel A provides three key insights: First, the strong diagonal pattern in the matrix shows that disease transmission occurs predominantly within the same age group. This aligns with previous studies in epidemiology showing that people mainly socialize with people of the same age which results in higher transmission rates within age groups than across age groups (e.g. Meyer and Held 2017; Mossong et al. 2008). Second, there are two salient clusters describing disease transmission across age groups: Children aged 0 to 2 and children aged 3 to 6 infect each other at particularly high rates. The same holds for adults and the elderly population. Third, the estimates suggest that an infected child below the age of 3 transmits influenza to a higher number of individuals per week than infected individuals in any other age group. This corresponds to findings in medical studies showing that children shed greater quantities of influenza viruses for longer periods of time compared to adults (King et al. 2005). Hence, the results point towards the role of children as main disseminators of influenza within a community discussed in the epidemiological literature (Gonzalez et al. 2000; Heikkinen et al. 2004).

In a subsequent step, I analyze whether and to which extent an increase in the early child care coverage rate affects the estimated age-specific transmission rates. To this end, I interact age-specific transmission rates ($\alpha^{WR}_{a'a}$) with annual and county-specific early child care coverage rates. Panel B in Figure 2.4 summarizes the estimated effects of a 1pp increase in the child care coverage rate on age-specific transmission rates and the respective p-values. The coefficients can be interpreted as the additional number of infections in age group a caused by an infected individual in age group a' due to a 1pp increase in the child care coverage rate. I find that the expansion in child care significantly increases disease transmission between 0 to 2 year old children and 3 to 6 year old children. The expansion, however, appears not to affect transmission patterns among 0 to 2 year

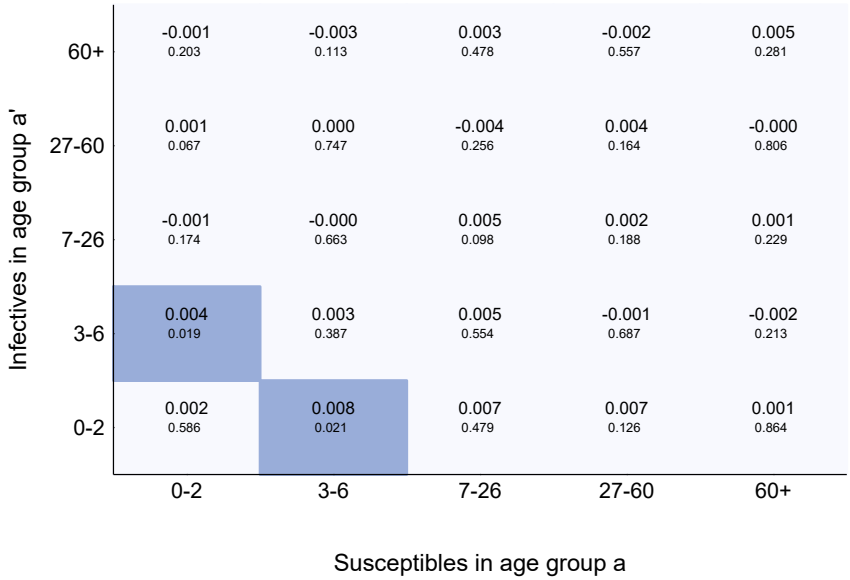
[36] Besides, the IV estimates of within regional transmission rates presented in Table 2.3 Panel A are consistently larger than OLS estimates. Hence, estimates on transmission rates based on OLS regression can be regarded as lower bounds.

Figure 2.4: Econometric Model 2: Age-Stratified Spatial-Time Spread of Diseases

(a) Age-specific Transmission Rates



(b) Effect of Child Care on Age-specific Transmission Rates



Notes: Panel (a) presents the estimated age-specific transmission rates within counties ($\alpha^{WR}_{a'a}$) based on Equation 2.3 using OLS regressions and the respective p-values. Cells with p-values below 0.001 and an effect size above 0.05 are highlighted in blue. Panel (b) presents the change in transmission rates in response to an increase in the ECC rate. Cells with p-values below 0.01 are highlighted in blue. Each column presents estimates from a different regression. The dependent variables are the incidence rates of susceptibles in a particular age group. All regressions include county and week-in-a-year fixed effects. In addition, regressions presented in Panel (b) control for season dummies, month dummies, climate, GDP per capita and distance to an airport interacted with transmission rates within a county. The complete set of regression coefficients is presented in Appendix Table B.15. Regressions are weighted by population size in a particular age group.

old children. The results suggest that contact rates of 0 to 2 year old children with 3 to 6 year old children increase in response to the expansion in child care, whereas contact patterns within the age cohort of 0 to 2 year old children remain constant. One possible explanation for this finding is that children in the same age group have contact to each other at the same rate with and without attending child care. Children not enrolled in formal child care may meet peers in privately organized parent-child play groups or at playgrounds instead of in child care facilities. However, children in different age cohorts may not meet each other at the same rate outside child care centers and thus, the expansion in early child care increases contact rate between these age cohorts. This could be explained by age-mixed groups in child care centers (see Section 2.2.1 for more details).

As in the previous analyses, I find that the treatment effect is driven by urban areas (see Appendix Figure B.8, Appendix Table B.16, and Appendix Table B.17).

2.5 Policy Interventions: Mandatory Vaccination and Closure of Child Care Centers

In this section, I use the dynamic model of disease diffusion specified in Equation 2.3 as a framework to study the impact of two counterfactual policy interventions: mandatory vaccination before entry into formal child care and the closure of child care facilities during local outbreaks. I first explore the effect of these policies on reducing the spread of influenza and then, evaluate economic benefit and costs induced by these policy measures. In the counterfactual analysis, I assume that the setup is unchanged except for the introduction of the respective policy. Thus, I rule out by assumption that parental choice of home care versus formal child care is affected by the policy.[37]

2.5.1 Effect of Policies on Disease Incidence

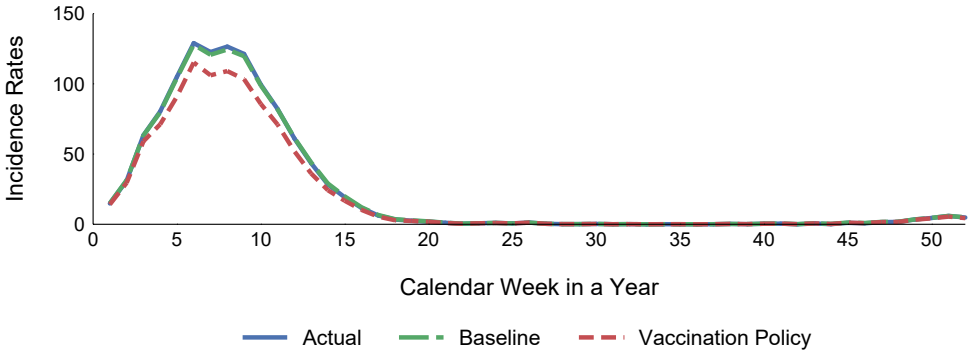
Mandatory Vaccination before Entry in Child Care Mandatory vaccination before entry into child care is one strategy adopted in many European countries to contain the spread of infectious diseases (e.g. France, Italy, Slovenia). In Germany, the government only recently enforced mandatory vaccination against measles before entry into child care (*Masernschutzgesetz*, March 2020). To study the impact of mandatory vaccination against influenza before entry into child care, I estimate dynamically how incidence rates would have evolved, if each child attending formal care was vaccinated.[38] Figure 2.5 summarizes the average weekly incidence rates of children aged 0 to 2 (blue line), estimated incidence rates using the dynamic model without imposing any policy (green line), and with enforcing mandatory vaccination (red line). First of all, the figure illustrates that the dynamic model of disease diffusion describes the actual incidence rates reported over the

[37] In the counterfactual analysis, I leave out the 2009/10 influenza season (known as swine flu) due to occurrence of a new virus mutant and slightly different disease features. Results, however, do not change significantly, when considering the 2009/10 influenza season in the analysis.

[38] I assume that (i) vaccination rate is identical to the child care coverage rate in a specific year and county and (ii) the efficiency of the vaccine is 60%.

year well (i.e. blue and green line are nearly the same). Further, I find that mandatory vaccination reduces annual mean incidence rates of children below the age of 3 by about 11% and of children aged 3 to 6 by 3% ($p<0.01$, see Table 2.4). The data show that mandatory vaccination benefits mostly children, who see a significant reduction in the incidence of influenza.

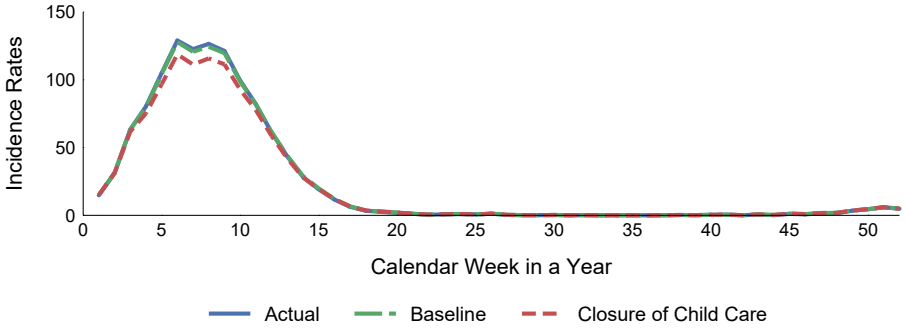
Figure 2.5: Counterfactual Analysis: Mandatory Vaccination before Entry into Child Care



Notes: The figure displays the average weekly reported incidence rates of children aged 0 to 2 (blue line), estimated incidence rates based on the dynamic model of disease diffusion without imposing any policy (green line), and with enforcing mandatory vaccination (red line). The data are aggregated over counties and averaged over the years between 2005 and 2016, excluding the 2009/10 influenza season.

Closure of Child Care Centers during Local Outbreaks Another feasible policy intervention to limit the spread of infectious diseases within child care facilities is to close care facilities during local outbreaks. To evaluate the impact of closure of child care centers during local outbreaks on the incidence rates of influenza, I estimate dynamically how infection rates would have evolved, if all child care centers in a county were closed for a two-week spell, once incidence rates in a county exceed 4.75 (lab-diagnosed) cases per 100,000 inhabitants for at least three subsequent weeks (i.e. the incidence rate is higher than 95% of the week-county pairs; definition 1). Alternatively, an outbreak is defined as incidence rates above 1.85 cases per 100,000 inhabitants for at least three subsequent weeks (i.e. the incidence rate is higher than 90% of the week-county pairs; definition 2). Figure 2.6 and Table 2.4 illustrate that a closure of child care facilities for two weeks during local outbreaks significantly reduces incidence rates by about 2 to 4%. Importantly, closure of child care facilities not only causes a significant decrease in the incidence rates of children, but also of people aged 6 to 60. Infection rates of the elderly population are not affected by the policy intervention.

Figure 2.6: Counterfactual Analysis: Closure of Child Care Centers



Notes: The figure displays the average weekly reported incidence rates of children aged 0 to 2 (blue line), estimated incidence rates based on the dynamic model of disease diffusion without imposing any policy (green line), and with closure of child care centers using definition 1 for local outbreaks (red line). The data are aggregated over counties and averaged over the years between 2005 and 2016, excluding the 2009/10 influenza season.

Table 2.4: Effect of Policy Interventions on Disease Incidence

	Age: <3	**Age: 3-6**	**Age: 7-26**	**Age: 27-60**	**Age: 60+**
Vaccination Policy	-0.108***	-0.030***	-0.015**	-0.008*	-0.003
	(0.040)	(0.018)	(0.012)	(0.011)	(0.011)
Closure of Child Care Centers (Def. 1)	-0.024**	-0.020**	-0.018**	-0.020**	-0.011
	(0.022)	(0.018)	(0.016)	(0.019)	(0.015)
Closure of Child Care Centers (Def. 2)	-0.041***	-0.035***	-0.031***	-0.033**	-0.017
	(0.028)	(0.025)	(0.021)	(0.024)	(0.019)

Notes: The table reports the mean relative effect of each policy intervention on annual disease incidence rates, e.g. mandatory vaccination reduces annual incidence rates of children aged 0 to 2 by 10.8%. The standard deviation is presented in brackets. The stars denote the p-value of a two-sided t-test displaying whether a policy significantly reduces annual incidence rates compared to mean annual incidence rates without policy intervention. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

2.5.2 Cost and Benefit Analysis

I now evaluate the economic benefits of mandatory vaccination policies and the closure of child care facilities during outbreaks, measured by the costs saved due to avoided infections. To this end, I draw on epidemiological literature that calculates the costs per case of influenza based on health insurance claims (Scholz et al. 2019). The data consider direct medical costs, such as outpatient care, hospitalization and medication, and also indirect costs caused by sick leave.[39] Total costs per case are estimated separately for different age groups. A limitation of this data set is that indirect costs are only calculated for adults, but not for children (i.e. a loss in parents' productivity caused

[39]Indirect costs are calculated as a function of the average number of working days lost due to an infection, average income and the percentage of individuals that participate actively in the labor market. See Scholz et al. (2019) for more details.

by a sick child, is not taken into account). Hence, the estimated costs per sick child represent a lower bound of the total costs. I supplement the data from Scholz et al. (2019) with information on costs of premature death that are estimated based on the concept of the value of a statistical life (VSL).[40] Appendix Table B.18 summarizes the costs considered in the analysis. Medical costs are highest for children and the elderly. For adults, direct medical costs are relatively low. However, total costs per case are highest for adults due to indirect costs caused by sick leave.[41]

The economic gains of the policy interventions are given by multiplying costs per case with the number of cases avoided. Recall from Section 2.2.2, that the data on disease incidence provided by the system of notifiable diseases includes mainly lab-diagnosed cases. Hence, using this data set to calculate the number of cases avoided likely underestimates the actual reduction in the number of cases. For this reason, I use medical claims data that document the annual share of people clinically diagnosed with influenza-like diseases (ICD-10 codes: J09 to J11) to calculate the number of avoided cases (Scholz et al. 2019). In doing so, I impose the assumption that the estimated relative effect size of the policy interventions is the same for clinically and lab-diagnosed cases. This procedure allows me to estimate the benefits of the two policies in monetary terms for clinically diagnosed cases.[42]

Table 2.5 summarizes the estimated annual economic benefits of the two policy interventions across all age groups. The results provide suggestive evidence that economic benefits due to closure of child care facilities are larger than those resulting from mandatory vaccination policies.

The implementation of policy interventions also imposes costs on society, which makes them not per se desirable as a policy measure to curb influenza. Therefore, it is important to compare economic benefits to the costs imposed by a policy intervention. For this reason, I evaluate net benefits of introducing mandatory vaccination policies.[43] To calculate the costs imposed by mandatory vaccination policies, I use information on costs per flu shot from medical billing data.[44] To vaccinate each child attending child care between 2005 and 2016, would have imposed annual average costs of €2.3 million on the German health care system (Table 2.5, Row 2).

[40]The VSL measures the society's willingness to pay for a small change in the probability of a fatality (Ashenfelter 2006; Kniesner and Viscusi 2019; Viscusi and Aldy 2003). There is substantial variation in the calculated VSL (Viscusi and Aldy 2003). In this study, I use a range between €1.3 and €6 million, which corresponds to the range found in the literature (Adda 2016).

[41]In Germany, adults stay absent from work for an average of 6.71 days due to an influenza caused sickness spell (Scholz et al. 2019).

[42]More specifically, I take the share of people infected with influenza-like diseases (ILD) to calculate the annual average number of infected with ILD in West Germany (Appendix Table B.19). I further estimate the number of cases with ILD avoided due to the policy interventions (Appendix Table B.20 Column 3) by multiplying the relative effect size of the respective policy (Appendix Table B.20 Column 1) with the annual number of infected without any policy intervention (Appendix Table B.19, Column 4).

[43]I abstract from calculating the costs of closing child care facilities for a two-week spell, as there is no comprehensive database on the costs imposed by the policy.

[44]Costs per flu shot are about €7 and comprise the costs of the vaccine and of consulting a physician.The costs per flu shot vary between privately and publicly insured patients. For publicly insured individuals, the costs were about €7 in 2012 (Medical billing: Symbolnummer: 89111, 89112) and varied slightly across regions in Germany. For privately insured individuals, the costs were about €20 per shot (Medical billing: GOÄ: 1 and 375; factor 2.3).

Table 2.5: Estimated Benefits and Costs, by Policy Intervention

	Mandatory Vaccination	**Closure of Child Care Centers**	
		Outbreak Def. 1	Outbreak Def. 2
Annual Benefits (in €)	2,981,324	5,488,629	9,068,098
Annual Costs (in €)	2,306,590	-	-
Annual Net Benefits (in €)	+ 674,734	-	-

Notes: The table reports the estimated annual economic benefits of mandatory vaccination and closure of child care centers based on health insurance claims (Row 1). The data on health insurance claims include clinically diagnosed cases of influenza-like diseases between 2012 and 2014 (Scholz et al. 2019). The economic benefits are calculated by multiplying the number of avoided cases per age group (Appendix Table B.20) with the lower bound of the calculated total costs per case in a particular age (Appendix Table B.18). The estimate reported presents the aggregate over all age cohorts. Row 2 presents the annual costs of introducing mandatory vaccination policy and is approximated by taking the costs per flu shot times the average number of children attending child care. Row 3 presents the annual net benefit of mandatory vaccination policies.

Comparing costs and benefits of enforced vaccination policies shows that mandatory vaccination policies would reduce annual costs by 23% compared to the status quo (Table 2.5, Row 3).

2.6 Discussion

This paper provides causal evidence on the impact of early child care on the spread of influenza. I find that the roll-out of child care increases disease incidence of children aged 0 to 2, children aged 3 to 6 and, in some model specifications, adults. The effect is driven by urban areas, while there is no significant impact of the expansion in child care on infection rates in rural areas.

Several mechanisms might explain the observed differences in transmission patterns between rural and urban counties: First, there might be differences in the baseline incidence rates prior to the reforms between rural and urban areas. If child care centers have a multiplier effect on disease incidence, higher baseline incidence rates in urban areas than in rural areas may explain the findings. The data on disease incidence, however, show that mean baseline incidence rates do not differ significantly between urban and rural areas, ruling out that explanation. Second, the effect of child care expansion on incidence rates might be non-linear. At the start of the expansion, supply of child care was significantly lower in rural areas compared to urban areas ($p< 0.001$). Third, differences in the expansion strategy and the organizational structure of child child care between urban and rural areas may drive the results. The expansion of child care can occur in at least two ways: By expanding existing facilities or by creating new facilities. Expanding existing child care centers requires that facilities exist prior to the reforms and is likely to result in an increased num-

ber of children per center compared to newly built centers. Children enrolled in large facilities interact with a higher number of children at common play areas compared to children attending small centers. As rural areas started at very low levels of child care coverage rates in 2005, expansion likely occurred primarily by creating new facilities. In contrast, child care provision in urban areas started at higher levels and thus, expansion is likely to be implemented by enlarging existing facilities and creating new ones. Additionally, group structures, for instance in terms of the age composition, may significantly vary across rural and urban areas. An alternative explanation for heterogeneous treatment effects between urban and rural areas might be differences in the transportation mode to the care center. Children living in urban areas are likely to be taken to the child care center by public transport and thus, are arguably exposed to a greater risk of infection than children in rural areas that are brought to child care centers by foot or car. Finally, this heterogeneity does not appear to be specific to influenza, as I find a similar pattern for norovirus gastritis, a highly contagious viral disease infecting particularly children (see Appendix Table B.5).

The roll-out of early child care also changes female labor force participation (Müller and Wrohlich 2019). Hence, some of the increase in infection could be coming from the mother's workplace. To investigate whether and to which extent the effect is driven by an increase in the mothers' labor supply, I run the analysis with and without controlling for county-specific female labor participation. I find that the effect of the expansion of child care on infection rates is nearly the same in both specifications. This indicates that the location of increased infections is indeed child care and not mothers' workplace (see Appendix Table B.7, Appendix Table B.8, Appendix Table B.13, and Appendix Table B.14).

The contribution of child care facilities to the propagation of infectious diseases is highly debated in the context of the ongoing Covid-19 pandemic (Gilliam et al. 2021; Lopez et al. 2020). Findings from this study can, however, only be partly transferred to other diseases such as Covid-19, as transmission patterns strongly depend on disease-specific parameters such as incubation time, infectivity, and transmission mode. Influenza and Covid-19 share some characteristics – both viruses are transmitted by contact, droplets, and fomites and cause respiratory symptoms (World Health Organization 2021). Hence, it is plausible to assume that a change in contact patterns in response to increased child care attendance might affect the spread of Covid-19 or more generally, the spread of any airborne-transmitted infectious disease. However, the effect size of child care on disease transmission may vary depending on other disease-specific characteristics such as viral load in patients of different age groups. In the case of Covid-19, age-specific disease transmission patterns appear to strongly differ from those of influenza.[45] Summing up, this study provides evidence that child care facilities spur the spread of airborne-transmitted infectious diseases. To

[45]Children appear to be less susceptible to infections of Covid-19, have lower rates of hospitalisation and when infected, less often lead to onward transmission than adolescents and adults (European Centre for Disease Prevention and Control 2020; Parri et al. 2020; Viner et al. 2020). In contrast, children are more often infected with influenza, have higher rates of hospitalization and when infected, more often lead to onward transmission compared to adolescents and adults below 60 (Amboss 2021; Gonzalez et al. 2000; Heikkinen et al. 2004).

assess, however, the exact effect size of child care attendance on the transmission of other viruses than influenza, additional studies are required.

2.7 Conclusion

This paper contributes to a better understanding of disease transmission patterns by exploring the impact of early child care on the spread of influenza. Using a difference-in-differences approach, I find that early child care significantly increases infection rates of influenza. To identify age-specific transmission patterns, I extend a semi-parametric model of disease diffusion that builds on epidemiological work. The results show that transmission patterns between children aged 0 to 2 and children aged 3 to 6 are profoundly increased in response to a rise in early child care provision. I further document significant differences in the impact of child care on the spread of influenza between urban and rural areas. However, the geographic differences do not appear to be specific to influenza, as I find a similar pattern for norovirus gastritis, another highly contagious viral disease. The heterogeneous treatment effects between urban and rural areas indicate that there exist systematic differences in the organizational structure of child care facilities between urban and rural areas that need to be further explored in future research. In a counterfactual analysis, I find that mandatory vaccination before entry into child care and the closure of child care facilities significantly reduce disease incidence by up to 11%. These policy interventions benefit mostly children, who see a significant reduction in the incidence of influenza. In addition, a cost-benefit analysis of mandatory vaccination policy shows that the policy would be cost-effective decreasing net annual costs by about 20% relative to the status quo.

The emergence and rapid transmission of Covid-19 pandemic has propelled the importance of infectious diseases for society and the economy into the public spotlight. Less salient, but still substantial are the health and economic hazards of established viruses such as seasonal influenza (Girard et al. 2005; Lambert and Fauci 2010; Schwandt 2017). Identifying the "hotspots" in society that are particularly relevant for the spread of infectious may help individuals to protect themselves, and also policy makers to enact targeted containment measures and thus, minimize costs to society.

Appendix B

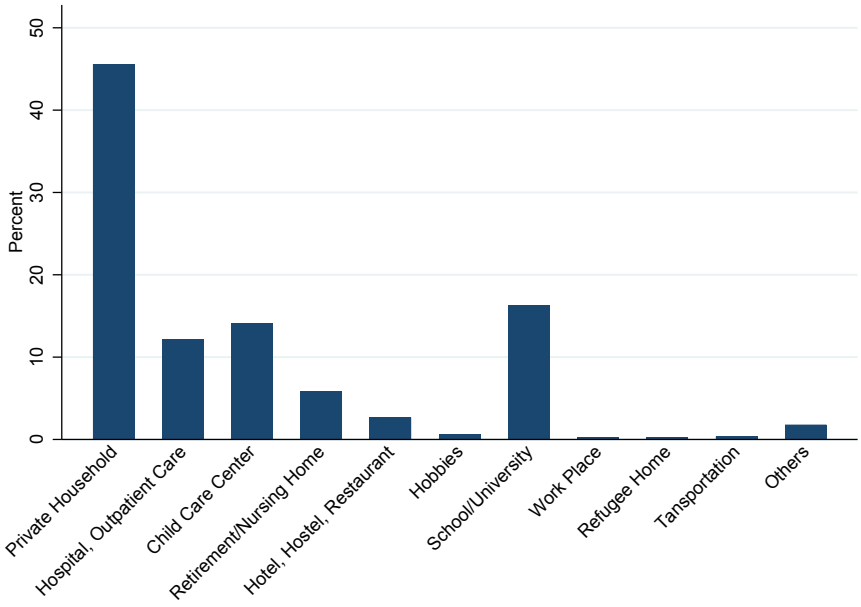
B1 Data Description

Table B.1: Summary Statistics

	Mean	S.D.
Child Care Coverage Rate		
ECC rate in 2005 (in %)	7.32	3.83
ECC rate in 2010 (in %)	19.57	5.69
ECC rate in 2016 (in %)	27.13	5.91
Kindergarten coverage rate in 2005 (in %)	86.53	6.90
Kindergarten coverage rate in 2010 (in %)	92.70	3.92
Kindergarten coverage rate in 2016 (in %)	93.56	3.35
Incidence of influenza		
Annual incidence rate, total population	50.68	78.79
Annual incidence rate, aged 0 to 2	131.25	207.72
influenza-vaccination rate (age 60+)	35.30	6.97
Socio-Economic Characteristics		
Female labor market participation (in %)	48.41	5.00
GDP per capita (in €)	33.14	14.50
Unemployment rate (in %)	6.14	2.78
Regional Characteristics		
Population density (population per km^2)	566.48	694.93
Population size, total population	20,1813.14	17,7974.66
Population size, aged 0 to 2	5,186.96	5,060.32
Distance to airport (mean time in minutes by car)	48.25	21.35
ID 1 - urban/rural areas (dummy for rural=1)	0.44	0.50
ID 2 - urban/rural areas (= large cities)	0.18	0.38
ID 2 - urban/rural areas (= cities)	0.39	0.49
ID 2 - urban/rural areas (= rural with small cities)	0.24	0.43
ID 2 - urban/rural areas (= sparsely populated areas)	0.20	0.40
Weather		
Temperature (in celsius)	9.53	1.41
Precipitation (in mmHg)	2.19	0.55
Absolute humidity (in g/m^3)	8.19	0.55

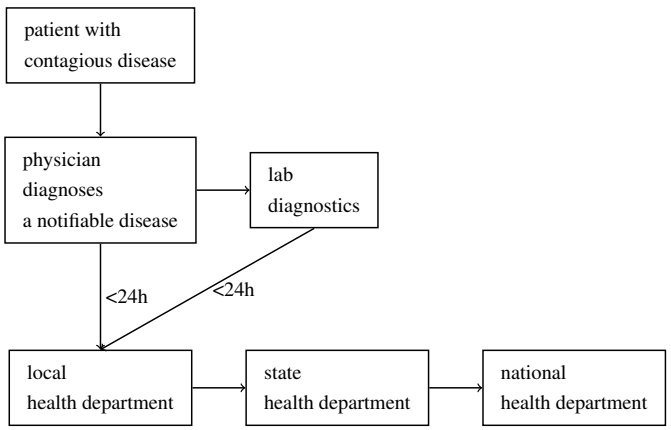
Notes: The table reports the mean and standard deviation of the set of variables included in the analysis, by county. The data is pooled over the period of observation (2005 to 2016). I use two different classification schemes for rural and urban areas proposed by BBSR (2019): First, a binary variable (ID 1): "Urban Areas" denote *Kreisfreie Großstädte* with more than 100,000 inhabitants and *Städtische Kreise* with a population density of at least 150 inhabitants/km^2 and at least 50% of the population living in a city. "Rural Areas" denote the remaining counties. Second, a variable that groups German counties into four categories (ID 2): "Large Cities" denote *Kreisfreie Großstädte* with more than 100,000 inhabitants. "Cities" present *Städtische Kreise* with a population density of at least 150 inhabitants/km^2 and at least 50% of the population living in a city. "Rural Areas with Small Cities" include (i) counties with a population density less than 150 inhabitants/km^2 and at least 50% of the population living in a city and (ii) counties with a population density of at least 100 inhabitants/km^2 and less 50% of the population living in a city. "Sparsely Populated Areas" denote counties with a population density smaller than 100 inhabitants/km^2 and less than 50% of the population living in a city.

Figure B.1: Locations of Influenza Infections in Germany



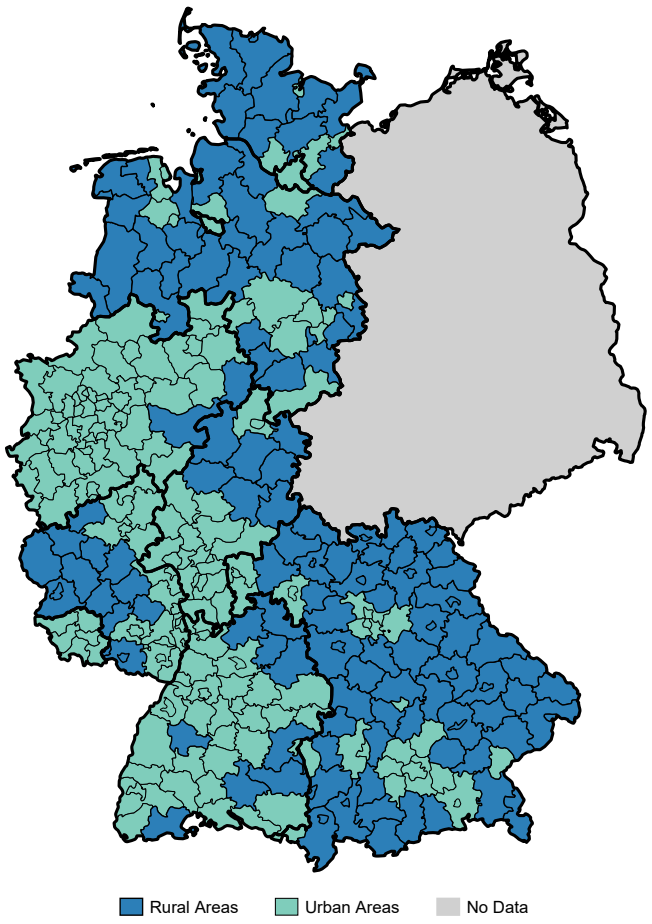
Notes: The figure summarizes the locations of infection with influenza derived from contact tracing by German health authorities (Robert Koch Institute). The data is aggregated over the years between 2005 and 2016.

Figure B.2: Reporting System of Notifiable Diseases in Germany



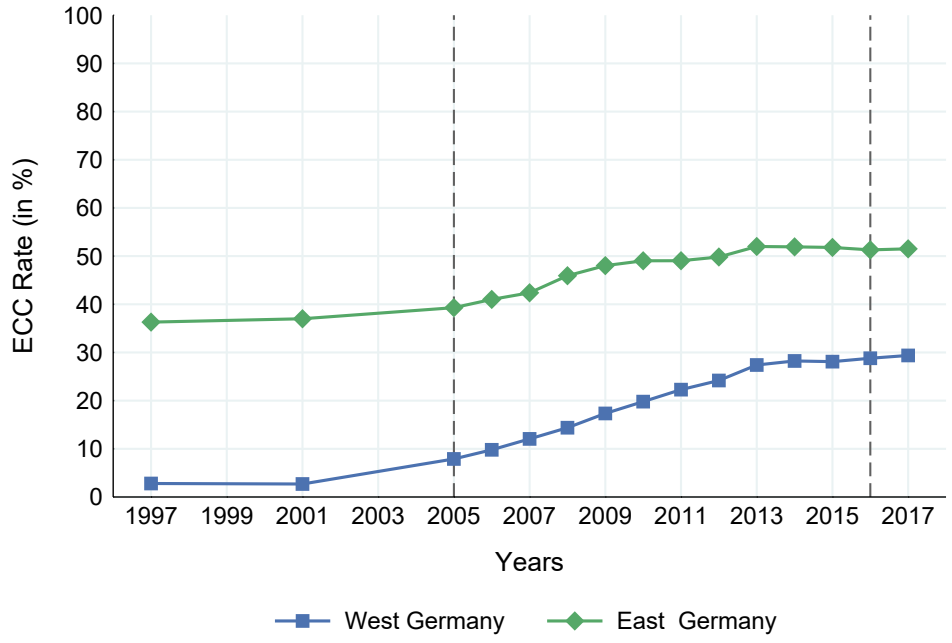
Notes: The figure illustrates the reporting system of notifiable diseases in Germany. The diagnosis of a notifiable disease has to be reported by a physician or a lab to the local health department that passes the information on to the respective federal health department. The state health department in turn transmits the information to the national department for public health (Robert Koch Institute).

Figure B.3: Classification of Counties into Rural and Urban Areas



Notes: The figure illustrates the classification of the 324 counties in West-Germany into rural (green) and urban areas (blue). The classification is based on the definition of the Federal Institute for Research on Building, Urban Affairs and Spatial Development (*Bundesinstitut für Bau-, Stadt- und Raumforschung*) (BBSR 2019).

Figure B.4: Trend in the Early Child Care Coverage Rate, East and West Germany



Notes: The figure illustrates the annual early child care coverage rates for children aged between 0 to 2 between 1997 and 2017. The blue line presents the early child care coverage rates in West Germany and the green line presents the annual child care rate in East Germany. The vertical dashed, grey lines indicate the start and end point of the observation period.

Supplementary Information on Weather Data The data is obtained from 439 active weather stations allocated within Germany. Based on geographic information, I link each municipality to its nearest weather station. The mean distance between the weather station and a municipality is 10km (S.D.=3.8km). In case a county consists of more than one municipality, I take the average temperature and precipitation per week within a county. In doing so, I obtain weekly information on temperature, precipitation and absolute humidity per county.

B2 DiD Strategy

Assessing the Identification Strategy

I study whether baseline county characteristics prior to the reforms, that induced the roll-out of child care slots (see Section 2.2.1), are related to the timing of the expansion. In doing so, I follow Akerman et al. (2015) and use Equation B.1:

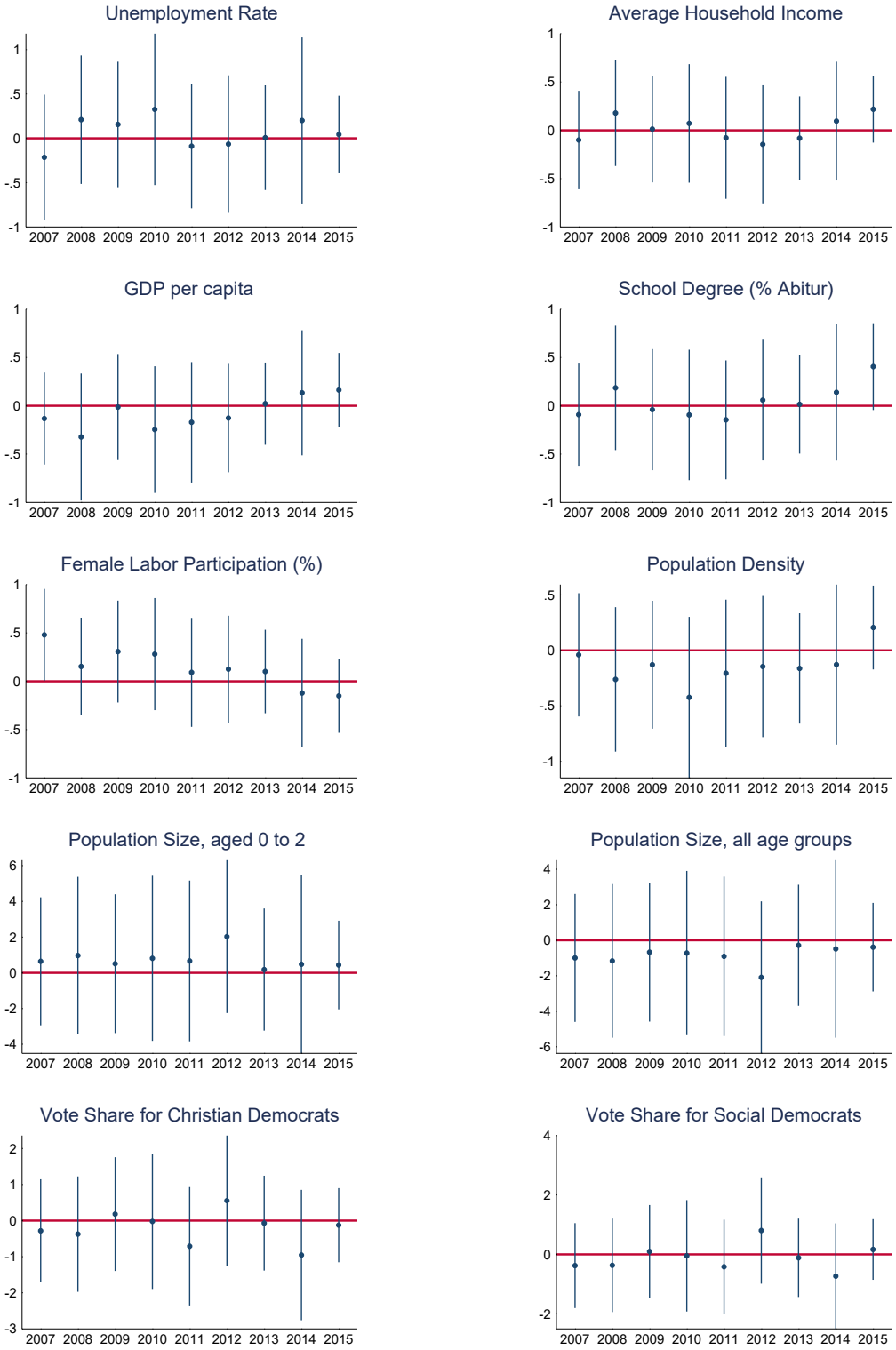
$$\Delta E_{c,t} = \delta_c + (\omega_t \times X_c)'\eta_t + \epsilon_{c,t} \tag{B.1}$$

where $\Delta E_{c,t}$ denotes the change in early child care coverage rate between two subsequent years, i.e. $\Delta E_{c,t} = E_{c,t} - E_{c,t-1}$. The parameters δ_c and ω_t represent county and year fixed effects, respectively. X_c includes baseline county characteristics prior to the expansion in early child care. More specifically, I study whether counties' economic performance (i.e. unemployment rate, average household income, female labor market participation and education), regional characteristics (i.e. population density, number of 0 to 2 year old children), and political attitudes are related to the timing of the roll-out.[46] In addition, I investigate whether average incidence rates for influenza in the pre-expansion period correlate with the timing of child care expansion. For each variable included in X_c, I take the latest available observation between 2002 and 2004.

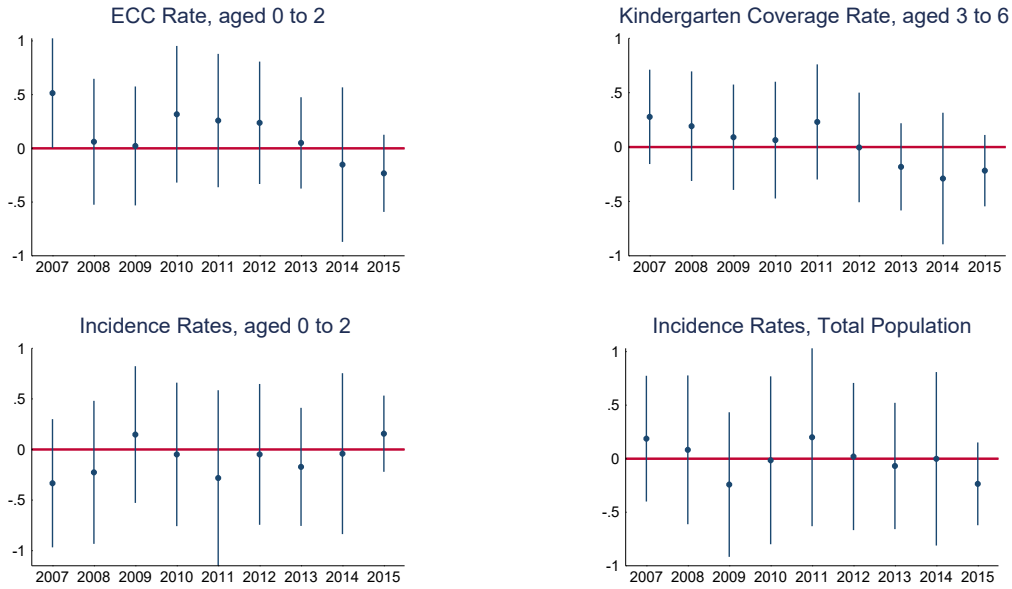
[46] Prior literature shows that a higher vote share for Social Democrats is related to expanding child care from half to full day slots (Felfe and Zierow 2018).

Figure B.5: Timing of Child Care Expansion and Baseline Covariates

(a) Socio-Economic Characteristics



(b) Child Care Coverage Rates and Disease Incidence



Notes: The figures report estimates of the vector η_t for each year t and the associated 95% confidence intervals based on Equation (B.1). All variables included in X_c are standardized.

Selective Reporting - Likelihood of GP Visit Conditional on Symptoms and Child Care Attendance

Table B.2: Number of GP Visits and Child Care Attendance

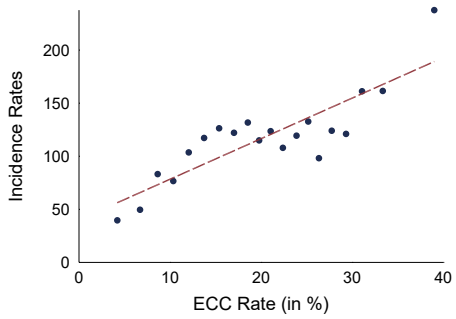
	Annual No. of Medical Consultation		
	Home Care	Center Care	p-value
	β/(SE)	β/(SE)	$\beta^{\text{Home Care}}=\beta^{\text{Care Center}}$
No. of Flu-like Diseases in the Last 12 Month	0.69***	0.49***	0.36
	(0.17)	(0.12)	
No. of Observations	1353	232	

Notes: The table displays the estimated effect of the number of flu-like diseases on the number of medical consultation for a sample of children aged 0 to 2 that is cared for at home (Column 1) and for a sample of children aged 0 to 2 that is attending public child care (Column 2). Column 3 reports the p-value of the differences in means between children cared for at home and children attending public child care. All models control for child age and gender, survey year fixed effects, and general health conditions. Robust standard errors are shown in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Data source: KiGGS survey data gathered between 2003-2006 and 2014-2016 by the Robert Koch Institute.

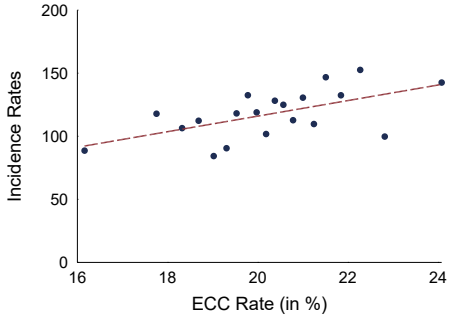
Additional Results

Figure B.6: Early Child Care Rates and Incidence Rates of Influenza

(a) W/o Controls and W/o FE



(b) With Controls and FE



Notes: The binned scatter plots display the link between incidence rates of children aged 0 to 2 (i.e. cases per 100,000 citizens aged 0 to 2) and the ECC rate per year and county. Panel (a) includes the raw data. Panel (b) controls for county and time fixed effects and a set of control variables (temperature, rain, absolute humidity, GDP per capita, employment rate, female employment rate, and population density).

Table B.3: DiD Estimates, by Age Groups

	Age 0 to 2		**Age 3 to 6**		**Age 7 to 26**		**Age 27 to 60**		**Age 60+**		**Total Population**	
ECC rate 0 to 2	1.047**	1.033**	1.033	1.023	1.009	1.009	1.018	1.016	1.014	1.009	1.019	1.015
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Observations	3564	3564	3564	3564	3564	3564	3564	3564	3553	3553	3564	3564
Pseudo R^2	0.634	0.639	0.646	0.649	0.845	0.847	0.725	0.729	0.724	0.728	0.751	0.754
Control variables	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
Season FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes

Notes: The table displays DiD estimates based on Poisson pseudo maximum likelihood regression using Equation (2.1). The dependent variables are the incidence rates in a particular age group. Coefficients are presented as odds ratios. The set of control variables includes temperature, rain, absolute humidity, GDP per capita, employment rate, female employment rate, and population density. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: DiD Estimates, by Age Groups, Urban and Rural Areas

	Age 0 to 2		**Age 3 to 6**		**Age 7 to 26**		**Age 27 to 60**		**Age 60+**		**Total Population**	
	Rural	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**	**Rural**	**Urban**
ECC rate 0 to 2	1.001	1.050***	0.988	1.038*	0.966*	1.045**	0.979	1.047**	1.016	1.014	0.974	1.048***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	1562	2002	1562	2002	1562	2002	1562	2002	1551	2002	1562	2002
Pseudo R^2	0.606	0.662	0.644	0.660	0.843	0.852	0.736	0.732	0.736	0.729	0.763	0.755
Season FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Control variables	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes

Notes: The table displays DiD estimates based on Poisson pseudo maximum likelihood regression using Equation 2.1, shown separately for urban and rural areas. The dependent variables are the incidence rates in a particular age group. Coefficients are presented as odds ratios. The set of control variables includes temperature, rain, absolute humidity, GDP per capita, employment rate, female employment rate, and population density. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.5: Sensitivity Analysis: DiD Estimates, Norovirus Gastritis by Age Group

	Age 0 to 2				**Total Population**			
	Full Sample		**Rural Area**	**Urban Area**	**Full Sample**		**Rural Area**	**Urban Area**
ECC rate 0 to 2	1.029**	1.025**	1.012	1.035**	1.012**	1.008	0.994	1.016**
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	3564	3564	1562	2002	3564	3564	1562	2002
Pseudo R^2	0.711	0.713	0.688	0.729	0.669	0.672	0.653	0.683
Season FE	yes	yes	yes	yes	yes	yes	yes	yes
County FE	yes	yes	yes	yes	yes	yes	yes	yes
Control variables	no	yes	yes	yes	no	yes	yes	yes

Notes: The table displays DiD estimates based on Poisson pseudo maximum likelihood regression using Equation 2.1, shown separately for urban and rural areas. The dependent variables are the incidence rates of norovirus gastritis in a particular age group. Coefficients are presented as odds ratios. The set of control variables includes temperature, rain, absolute humidity, GDP per capita, employment rate, female employment rate, and population density. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

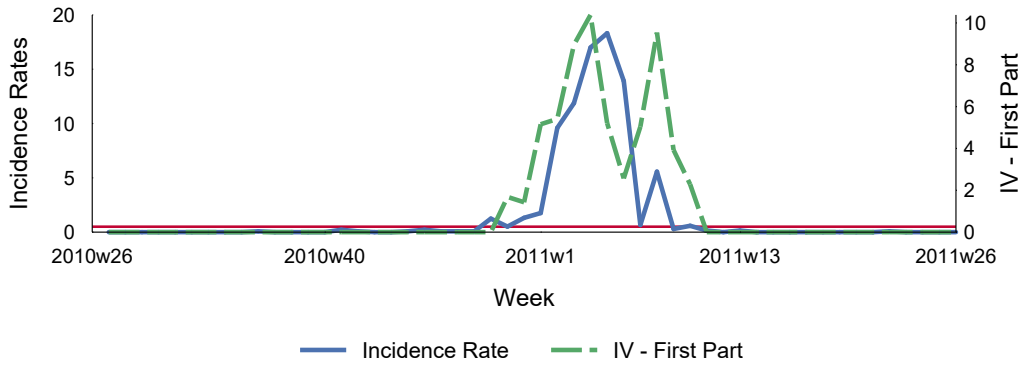
B3 IV Strategy

Table B.6: IV Strategy, First Stage

	Age: <3		Age: 3 to 6		Age: 7 to 26		Age: 27 to 60		Age: 60+		Total	
	WR	**BR**	**WR**	**BR**	**WR**	**BR**	**WR**	**BR**	**WR**	**BR**	**WR**	**BR**
IV^{WR} (Part 1 & 2)	-0.3397***	-0.0924***	-0.2568***	-0.0692***	-0.0419***	-0.0111***	-0.0186***	-0.0051***	-0.0318***	-0.0098***	-0.0087***	-0.0024***
	(0.107)	(0.022)	(0.073)	(0.016)	(0.011)	(0.003)	(0.005)	(0.001)	(0.008)	(0.002)	(0.002)	(0.001)
IV^{WR} (Part 1)	5.2003***	0.9337***	3.9191***	0.7105***	0.6376***	0.1158***	0.2857***	0.0526***	0.4747***	0.1025***	0.1321***	0.0254***
	(0.892)	(0.234)	(0.608)	(0.173)	(0.095)	(0.027)	(0.046)	(0.012)	(0.070)	(0.021)	(0.020)	(0.006)
IV^{BR} (Part 1 & 2)	-0.2719***	-0.5114***	-0.1818***	-0.3639***	-0.0276***	-0.0574***	-0.0135***	-0.0266***	-0.0200***	-0.0408***	-0.0059***	-0.0119***
	(0.074)	(0.060)	(0.054)	(0.043)	(0.009)	(0.006)	(0.004)	(0.003)	(0.006)	(0.004)	(0.002)	(0.001)
IV^{BR} (Part 1)	1.6914***	3.7094***	1.1238***	2.6422***	0.1688**	0.4175***	0.0824***	0.1929***	0.1190**	0.2931***	0.0360**	0.0864***
	(0.598)	(0.427)	(0.433)	(0.297)	(0.069)	(0.044)	(0.031)	(0.021)	(0.050)	(0.030)	(0.014)	(0.009)
Observations	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004	185004
F-Stat.	70.044	30.684	69.544	33.000	68.904	37.209	73.726	34.891	78.610	39.009	74.236	36.542
p-value F-Stat.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Sanderson-Windmeijer F-Stat.	57.272	36.629	51.569	34.393	47.550	34.135	52.299	35.019	53.227	38.485	51.503	35.782
p-value Sanderson-Windmeijer F-Stat.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Notes: The table displays the estimates of the first stage regressions for disease transmission within regions (WR) and between regions (BR). Coefficients are scaled by dividing by 100,000. Each column represents a separate regression. IV^{WR} and IV^{BR} denote the functional form described in Equation (2.5a) and (2.5b). All regressions include county and week-in-a-year fixed effects. Standard errors are clustered at the county level.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure B.7: Graphical Illustration of the Instrument



Notes: The figure shows one example of how the first part of the instrument mimics an outbreak in Munich in the influenza season 2010/2011. The green line displays the sum over the difference between the median temperature during an outbreak and the temperature in week i of the same outbreak. It presents a humped-shaped curve that grows as long as the temperature is below the median temperature and declines ones the temperature exceeds the median temperature. The curve overlaps with reported incidence rates (cases per 100,000 inhabitants) of influenza (blue line). The horizontal red line at an incidence rate of 0.5 presents the threshold value for the definition of an outbreak. The onset of an outbreak is defined as the first week of at least three consecutive weeks in that the incidence rate is larger than 0.5, the end of a local outbreak is defined as the first of at least five subsequent weeks in that the incidence rate is lower than 0.5. The sum over the difference between the median temperature of an outbreak and the temperature in week i of the epidemic outbreak is set to zero if (i) no outbreak takes place or (ii) the sum is below zero.

B4 Econometric Model 1: Time-Space Model

Table B.7: Econometric Model 1: Time-Space Model, OLS Regressions

	OLS Results					
	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
α^{WR}	0.12***	0.28***	0.25**	0.27***	0.086	1.01***
	(0.0223)	(0.0578)	(0.0979)	(0.0695)	(0.0552)	(0.1980)
α^{BR}	0.0050***	0.013***	0.049***	0.036***	0.0081**	0.12***
	(0.0015)	(0.0035)	(0.0137)	(0.0076)	(0.0037)	(0.0245)
α^{WR} × ECC rate	0.00058**	0.00052	0.0011	0.0028***	0.00072	0.0060***
	(0.0003)	(0.0005)	(0.0012)	(0.0007)	(0.0006)	(0.0022)
α^{WR} × gdp per capita	-0.000025	0.000087	-0.00027	0.00086**	0.00026	0.00085
	(0.0001)	(0.0002)	(0.0005)	(0.0003)	(0.0003)	(0.0009)
α^{WR} × airport	-0.000070	-0.00010	-0.000028	0.00016	-0.000087	-0.000055
	(0.0001)	(0.0002)	(0.0005)	(0.0002)	(0.0002)	(0.0008)
α^{WR} × population density	0.0000051*	0.0000042	-0.000046**	-0.000025**	-0.0000048	-0.000070**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
α^{WR}× cities	0.0039	0.016	0.0019	-0.030	-0.0054	-0.0083
	(0.0054)	(0.0099)	(0.0371)	(0.0256)	(0.0170)	(0.0617)
α^{WR}× rural areas with small cities	0.0046	0.016	-0.027	-0.024	-0.0098	-0.040
	(0.0050)	(0.0118)	(0.0422)	(0.0305)	(0.0196)	(0.0699)
α^{WR}× sparsely populated areas	0.0022	0.0077	-0.025	-0.024	-0.00065	-0.038
	(0.0050)	(0.0109)	(0.0434)	(0.0298)	(0.0198)	(0.0708)
α^{WR} × rain	0.00049	0.0024	0.015**	0.0042*	-0.0029***	0.020*
	(0.0006)	(0.0021)	(0.0066)	(0.0025)	(0.0011)	(0.0102)
α^{WR} × temperature	-0.0017**	-0.0028*	-0.015***	-0.0072**	-0.0033	-0.032***
	(0.0007)	(0.0015)	(0.0052)	(0.0029)	(0.0020)	(0.0087)
α^{WR} × absolute humidity	0.0024	-0.0025	0.040**	-0.0015	0.0034	0.048*
	(0.0018)	(0.0050)	(0.0159)	(0.0090)	(0.0054)	(0.0247)
Observations	185004	185004	185004	185004	185004	185004
Adj.R2.	0.391	0.448	0.621	0.664	0.470	0.700

Notes: The table displays estimates of Equation 2.3, in that one infected individual of any age spreads influenza to susceptibles in a particular age group. Estimates are based on OLS regressions. Each column presents estimates from a different regression. The dependent variables are the incidence rates of a particular age group. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table B.8: Econometric Model 1: Time-Space Model, IV Regressions

	IV-Results					
	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
α^{WR}	0.10***	0.35***	0.66***	0.39***	0.24***	1.78***
	(0.0324)	(0.0708)	(0.0802)	(0.0980)	(0.0809)	(0.1906)
α^{BR}	-0.023*	-0.041**	-0.057*	0.028	0.000020	-0.090
	(0.0118)	(0.0209)	(0.0301)	(0.0302)	(0.0228)	(0.0666)
α^{WR} × ECC rate	0.0012**	0.0011	-0.00065	-0.00026	-0.0013	0.00028
	(0.0006)	(0.0007)	(0.0008)	(0.0011)	(0.0014)	(0.0014)
α^{WR} × gdp per capita	-0.00021	-0.00019	0.00011	-0.000038	0.000038	-0.00032
	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0005)	(0.0005)
α^{WR} × airport	-0.000036	-0.00044	0.00066*	0.00032	-0.00019	0.00036
	(0.0002)	(0.0005)	(0.0003)	(0.0004)	(0.0004)	(0.0005)
α^{WR} × population density	0.0000040	-0.0000023	-0.000017	0.0000044	0.00000011	-0.000013
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
α^{WR}× cities	0.0010	0.0039	-0.0063	-0.0042	-0.019	-0.023
	(0.0143)	(0.0234)	(0.0224)	(0.0352)	(0.0321)	(0.0386)
α^{WR}× rural areas with small cities	0.0080	0.044	0.0070	-0.032	-0.047	-0.018
	(0.0153)	(0.0298)	(0.0258)	(0.0403)	(0.0361)	(0.0470)
α^{WR}× sparsely populated areas	0.010	0.035	0.0046	-0.020	-0.029	0.0026
	(0.0169)	(0.0267)	(0.0257)	(0.0440)	(0.0379)	(0.0475)
α^{WR} × rain	0.00059	0.0014	0.00074	-0.00034	-0.0018	0.0012
	(0.0013)	(0.0027)	(0.0031)	(0.0039)	(0.0025)	(0.0083)
α^{WR} × temperature	-0.0023	-0.0086***	-0.0076**	-0.0056	-0.0011	-0.026***
	(0.0014)	(0.0029)	(0.0037)	(0.0038)	(0.0028)	(0.0094)
α^{WR} × absolute humidity	0.012***	0.016**	0.0052	-0.0045	-0.0072	0.025
	(0.0044)	(0.0072)	(0.0115)	(0.0121)	(0.0106)	(0.0277)
Observations	185004	185004	185004	185004	185004	185004

Notes: The table displays estimates of Equation 2.3, in that one infected individual of any age spreads influenza to susceptibles in a particular age group. Estimates are based on IV regressions. Each column presents estimates from a different regression. The dependent variables are the incidence rates of a particular age group. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table B.9: Econometric Model 1: Time-Space Model, OLS Regressions, Urban Regions

	OLS Results					
	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
$\alpha^{WR} \times$ ECC rate	0.00080**	0.0014**	0.0011	0.0034***	0.00064	0.0074**
	(0.0004)	(0.0006)	(0.0015)	(0.0010)	(0.0009)	(0.0031)
α^{WR}	0.15***	0.41***	0.19	0.19**	0.067	1.01***
	(0.0296)	(0.0919)	(0.1508)	(0.0758)	(0.0609)	(0.2814)
α^{BR}	0.0067***	0.012***	0.053***	0.035***	0.0059	0.12***
	(0.0023)	(0.0042)	(0.0201)	(0.0106)	(0.0043)	(0.0359)
$\alpha^{WR} \times$ gdp per capita	-0.000029	-0.000076	-0.0000034	0.00088***	0.00025	0.00099
	(0.0001)	(0.0003)	(0.0006)	(0.0003)	(0.0004)	(0.0011)
$\alpha^{WR} \times$ population density	0.0000051*	0.0000028	-0.000043*	-0.000033***	-0.0000035	-0.000077*
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha^{WR} \times$ airport	-0.00016*	-0.00041	-0.00029	-0.000018	0.000089	-0.00080
	(0.0001)	(0.0003)	(0.0007)	(0.0003)	(0.0002)	(0.0012)
$\alpha^{WR} \times$ rain	0.00063	0.0016	0.023**	0.0085**	-0.0012	0.034**
	(0.0010)	(0.0033)	(0.0094)	(0.0036)	(0.0019)	(0.0154)
$\alpha^{WR} \times$ cities	0.0053	0.015	0.015	-0.040*	-0.0056	-0.0071
	(0.0053)	(0.0093)	(0.0419)	(0.0217)	(0.0193)	(0.0687)
$\alpha^{WR} \times$ temperature	-0.0024**	-0.0054**	-0.022***	-0.010***	0.00052	-0.043***
	(0.0011)	(0.0024)	(0.0073)	(0.0033)	(0.0023)	(0.0117)
$\alpha^{WR} \times$ absolute humidity	0.0047	-0.0014	0.059**	0.013	-0.0062	0.079**
	(0.0033)	(0.0080)	(0.0241)	(0.0107)	(0.0067)	(0.0367)
Observations	103922	103922	103922	103922	103922	103922
Adj.R2.	0.446	0.514	0.617	0.677	0.483	0.698

Notes: The table displays estimates of Equation 2.3, in that one infected individual of any age spreads influenza to susceptibles in a particular age group, shown separately for urban counties. Estimates are based on OLS regressions. Each column presents estimates from a different regression. The dependent variables are the incidence rates of a particular age group. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table B.10: Econometric Model 1: Time-Space Model, IV Regressions, Urban Regions

	IV-Results					
	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
α^{WR} × ECC rate	0.0017**	0.00089	-0.00023	0.000023	-0.00023	-0.000020
	(0.0008)	(0.0008)	(0.0002)	(0.0001)	(0.0002)	(0.0001)
α^{WR}	0.087*	0.29***	0.073***	0.0093	0.035***	0.041***
	(0.0462)	(0.0841)	(0.0171)	(0.0066)	(0.0105)	(0.0070)
α^{BR}	-0.043*	-0.090**	-0.0081	0.0037	0.0028	-0.0035
	(0.0222)	(0.0415)	(0.0076)	(0.0031)	(0.0042)	(0.0040)
α^{WR} × gdp per capita	-0.00048**	-0.00031	0.000036	0.000043	0.000076	0.000030*
	(0.0002)	(0.0003)	(0.0000)	(0.0000)	(0.0001)	(0.0000)
α^{WR} × population density	-0.0000028	-0.0000051	-0.0000032	0.00000097	0.00000043	-0.00000035
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
α^{WR} × airport	0.000074	0.00025	0.000032	-0.0000083	-0.000023	0.0000016
	(0.0003)	(0.0004)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
α^{WR} × rain	-0.00086	0.0013	-0.00015	-0.000040	0.00015	0.000013
	(0.0020)	(0.0022)	(0.0005)	(0.0004)	(0.0004)	(0.0003)
α^{WR}× cities	-0.014	-0.023	-0.0038	0.0012	-0.0028	-0.0018
	(0.0154)	(0.0237)	(0.0041)	(0.0023)	(0.0037)	(0.0017)
α^{WR} × temperature	-0.0044**	-0.013***	-0.0019***	-0.00062**	0.00070*	-0.0011***
	(0.0022)	(0.0041)	(0.0007)	(0.0003)	(0.0004)	(0.0004)
α^{WR} × absolute humidity	0.015**	0.011	0.0020	0.00093	-0.0032**	0.00100
	(0.0063)	(0.0076)	(0.0021)	(0.0009)	(0.0013)	(0.0009)
Observations	103922	103922	103922	103922	103922	103922

Notes: The table displays estimates of Equation 2.3, in that one infected individual of any age spreads influenza to susceptibles in a particular age group, shown separately for urban counties. Estimates are based on IV regressions. Each column presents estimates from a different regression. The dependent variables are the incidence rates of a particular age group. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.11: Econometric Model 1: Time-Space Model, OLS Regressions, Rural Areas

	OLS Results					
	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
α^{WR}	0.082***	0.17***	0.35***	0.41***	0.066	1.12***
	(0.0310)	(0.0461)	(0.1063)	(0.0843)	(0.0504)	(0.1880)
α^{BR}	0.0044**	0.016***	0.047**	0.040***	0.016***	0.12***
	(0.0019)	(0.0049)	(0.0186)	(0.0109)	(0.0049)	(0.0330)
α^{WR} × ECC rate	0.00035	-0.00036	0.0012	0.0023**	0.00069	0.0049
	(0.0002)	(0.0006)	(0.0020)	(0.0010)	(0.0006)	(0.0030)
α^{WR} × gdp per capita	0.00024	0.00065	-0.0011	-0.00022	0.00030	-0.00022
	(0.0002)	(0.0005)	(0.0019)	(0.0009)	(0.0004)	(0.0028)
α^{WR} × population density	-0.0000087	-0.000024	-0.000046	0.000027	-0.000015	-0.000061
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)
α^{WR} × airport	0.000027	0.00016	0.00029	0.00018	-0.00021	0.00057
	(0.0001)	(0.0002)	(0.0006)	(0.0003)	(0.0002)	(0.0009)
α^{WR} × sparsely populated areas	-0.0040	-0.014**	-0.0017	0.00044	0.011**	-0.0077
	(0.0026)	(0.0068)	(0.0247)	(0.0128)	(0.0049)	(0.0362)
α^{WR} × rain	0.00017	0.0035**	0.0087	-0.00018	-0.0046***	0.0067
	(0.0006)	(0.0018)	(0.0088)	(0.0032)	(0.0013)	(0.0114)
α^{WR} × temperature	-0.0018*	-0.0012	-0.0041	-0.0000047	-0.0064**	-0.013
	(0.0010)	(0.0018)	(0.0063)	(0.0047)	(0.0029)	(0.0115)
α^{WR} × absolute humidity	0.0026	0.00048	0.014	-0.026*	0.012*	0.00038
	(0.0022)	(0.0056)	(0.0219)	(0.0135)	(0.0069)	(0.0338)
Observations	81082	81082	81082	81082	81082	81082
Adj.R2.	0.329	0.390	0.639	0.660	0.465	0.710

Notes: The table displays estimates of Equation 2.3, in that one infected individual of any age spreads influenza to susceptibles in a particular age group, shown separately for rural counties. Estimates are based on OLS regressions. Each column presents estimates from a different regression. The dependent variables are the incidence rates of a particular age group. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table B.12: Econometric Model 1: Time-Space Model, IV Regressions, Rural Areas

	IV-Results					
	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
α^{WR}	0.15***	0.42***	0.72***	0.54***	0.046	1.96***
	(0.0367)	(0.0995)	(0.0963)	(0.1310)	(0.1010)	(0.2554)
α^{BR}	0.0019	-0.011	-0.010	0.037	0.027	0.039
	(0.0145)	(0.0313)	(0.0653)	(0.0364)	(0.0370)	(0.0976)
α^{WR} × ECC rate	0.00020	0.0000088	-0.00082	-0.00071	0.000087	-0.00098
	(0.0005)	(0.0009)	(0.0013)	(0.0015)	(0.0018)	(0.0020)
α^{WR} × gdp per capita	-0.000077	-0.00072	-0.00045	0.000076	0.000016	-0.0013
	(0.0004)	(0.0008)	(0.0008)	(0.0009)	(0.0010)	(0.0013)
α^{WR} × population density	0.000016	0.000017	0.000062	-0.000070	-0.000033	0.0000069
	(0.0000)	(0.0000)	(0.0001)	(0.0001)	(0.0000)	(0.0001)
α^{WR} × airport	-0.000025	-0.00055	0.00085*	0.00020	-0.00032	0.00033
	(0.0002)	(0.0009)	(0.0005)	(0.0005)	(0.0004)	(0.0006)
α^{WR} × sparsely populated areas	-0.00021	-0.011	0.0028	0.011	0.019	0.022
	(0.0077)	(0.0171)	(0.0230)	(0.0193)	(0.0168)	(0.0260)
α^{WR} × rain	0.0016	-0.0027	0.00067	-0.0015	-0.0035	-0.0046
	(0.0017)	(0.0049)	(0.0055)	(0.0061)	(0.0039)	(0.0153)
α^{WR} × temperature	0.00056	-0.0032	-0.00074	-0.00099	-0.0086**	-0.012
	(0.0015)	(0.0036)	(0.0041)	(0.0059)	(0.0042)	(0.0129)
α^{WR} × absolute humidity	0.0063	0.014	-0.0026	-0.022	0.021	0.012
	(0.0053)	(0.0115)	(0.0138)	(0.0216)	(0.0174)	(0.0447)
Observations	81082	81082	81082	81082	81082	81082

Notes: The table displays estimates of Equation 2.3, in that one infected individual of any age spreads influenza to susceptibles in a particular age group, shown separately for rural counties. Estimates are based on IV regressions. Each column presents estimates from a different regression. The dependent variables are the incidence rates of a particular age group. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.13: Econometric Model 1: Time-Space Model, OLS Regressions, Female Labor Participation

	OLS Results					
	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
α^{WR}	0.095***	0.16**	0.31*	0.20**	0.20***	0.96***
	(0.0320)	(0.0628)	(0.1580)	(0.0842)	(0.0537)	(0.2531)
α^{BR}	0.0052***	0.013***	0.049***	0.036***	0.0079**	0.12***
	(0.0015)	(0.0035)	(0.0137)	(0.0077)	(0.0035)	(0.0246)
$\alpha^{WR}\times$ ECC rate	0.00052**	0.00020	0.0013	0.0026***	0.00100	0.0058**
	(0.0002)	(0.0005)	(0.0013)	(0.0008)	(0.0006)	(0.0023)
$\alpha^{WR}\times$ gdp per capita	-0.000026	0.000085	-0.00026	0.00086**	0.00026	0.00085
	(0.0001)	(0.0002)	(0.0005)	(0.0004)	(0.0003)	(0.0009)
$\alpha^{WR}\times$ female labor market participation (in %)	0.00040	0.0024**	-0.0011	0.0014	-0.0023*	0.0011
	(0.0004)	(0.0010)	(0.0026)	(0.0018)	(0.0013)	(0.0046)
$\alpha^{WR}\times$ population density	0.0000054**	0.0000058	-0.000047**	-0.000024**	-0.0000058	-0.000070**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha^{WR}\times$ airport	-0.000055	-0.000012	-0.000066	0.00021	-0.00018	-0.000013
	(0.0001)	(0.0002)	(0.0005)	(0.0002)	(0.0001)	(0.0008)
$\alpha^{WR}\times$ cities	0.0032	0.012	0.0040	-0.032	-0.00057	-0.011
	(0.0057)	(0.0101)	(0.0374)	(0.0280)	(0.0172)	(0.0639)
$\alpha^{WR}\times$ rural areas with small cities	0.0038	0.011	-0.025	-0.027	-0.0048	-0.043
	(0.0053)	(0.0127)	(0.0427)	(0.0336)	(0.0200)	(0.0733)
$\alpha^{WR}\times$ sparsely populated areas	0.0012	0.0013	-0.022	-0.028	0.0060	-0.041
	(0.0054)	(0.0120)	(0.0447)	(0.0330)	(0.0204)	(0.0756)
$\alpha^{WR}\times$ rain	0.00053	0.0026	0.015**	0.0044*	-0.0032***	0.020**
	(0.0006)	(0.0020)	(0.0066)	(0.0025)	(0.0010)	(0.0101)
$\alpha^{WR}\times$ temperature	-0.0018**	-0.0033**	-0.015***	-0.0075***	-0.0029	-0.032***
	(0.0007)	(0.0015)	(0.0052)	(0.0029)	(0.0020)	(0.0087)
$\alpha^{WR}\times$ absolute humidity	0.0031	0.0017	0.039**	0.00092	-0.00044	0.050**
	(0.0020)	(0.0044)	(0.0164)	(0.0090)	(0.0054)	(0.0251)
Observations	185004	185004	185004	185004	185004	185004
Adj.R2.	0.392	0.450	0.621	0.665	0.472	0.700

Notes: The table displays estimates of Equation 2.3, in that one infected individual of any age spreads influenza to susceptibles in a particular age group. Estimates are based on OLS regressions. The dependent variables are the incidence rates of a particular age group. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Regressions are weighted by population size in a particular age group. In addition to Table B.7, the regressions control for female labor participation. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.14: Econometric Model 1: Time-Space Model, IV Regressions, Female Labor Participation

	IV-Results					
	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
α^{WR}	0.16***	0.28***	0.75***	0.34***	0.44***	2.03***
	(0.0556)	(0.0931)	(0.1009)	(0.1150)	(0.1117)	(0.1984)
α^{BR}	-0.020*	-0.050**	-0.057*	0.025	0.012	-0.085
	(0.0112)	(0.0203)	(0.0302)	(0.0285)	(0.0230)	(0.0639)
$\alpha^{WR}\times$ ECC rate	0.0012**	0.0011	-0.00058	-0.00030	-0.0010	0.00045
	(0.0006)	(0.0007)	(0.0008)	(0.0012)	(0.0013)	(0.0014)
$\alpha^{WR}\times$ gdp per capita	-0.00019	-0.00015	0.00018	-0.000024	0.000019	-0.00016
	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0004)
$\alpha^{WR}\times$ female labor market participation (in %)	-0.0012	0.0017	-0.0016	0.0010	-0.0047**	-0.0046*
	(0.0010)	(0.0021)	(0.0016)	(0.0022)	(0.0022)	(0.0024)
$\alpha^{WR}\times$ population density	0.0000044	-0.0000069	-0.000018	0.0000027	0.0000059	-0.000014
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha^{WR}\times$ airport	-0.000093	-0.00036	0.00058*	0.00036	-0.00039	0.00013
	(0.0002)	(0.0005)	(0.0003)	(0.0004)	(0.0004)	(0.0005)
$\alpha^{WR}\times$ cities	0.0049	-0.0067	-0.0021	-0.0090	0.0025	-0.011
	(0.0158)	(0.0280)	(0.0252)	(0.0411)	(0.0358)	(0.0419)
$\alpha^{WR}\times$ rural areas with small cities	0.012	0.036	0.013	-0.035	-0.028	-0.0022
	(0.0164)	(0.0293)	(0.0278)	(0.0450)	(0.0397)	(0.0487)
$\alpha^{WR}\times$ sparsely populated areas	0.014	0.026	0.0099	-0.024	-0.0074	0.017
	(0.0182)	(0.0292)	(0.0282)	(0.0497)	(0.0417)	(0.0504)
$\alpha^{WR}\times$ rain	0.00046	0.0017	0.00062	-0.00016	-0.0026	0.00087
	(0.0013)	(0.0025)	(0.0031)	(0.0039)	(0.0024)	(0.0082)
$\alpha^{WR}\times$ temperature	-0.0022	-0.0088***	-0.0074**	-0.0057	-0.00064	-0.026***
	(0.0014)	(0.0029)	(0.0037)	(0.0038)	(0.0028)	(0.0093)
$\alpha^{WR}\times$ absolute humidity	0.011**	0.016**	0.0024	-0.0040	-0.010	0.017
	(0.0047)	(0.0075)	(0.0118)	(0.0120)	(0.0104)	(0.0276)
Observations	185004	185004	185004	185004	185004	185004

Notes: The table displays estimates of Equation 2.3, in that one infected individual of any age spreads influenza to susceptibles in a particular age group. Estimates are based on IV regressions. Each column presents estimates from a different regression. The dependent variables are the incidence rates of a particular age group. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Regressions are weighted by population size in a particular age group. In addition to Table B.8, the regressions control for female labor participation. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B5 Econometric Model 2: Age-Specific Time-Space Model

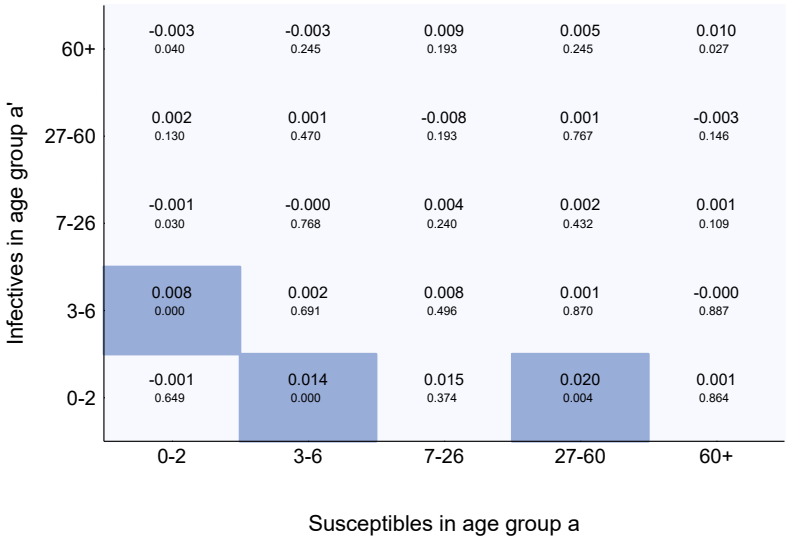
Table B.15: Econometric Model 2: Age-Specific Time-Space Model

	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
$\alpha^{WR}_{<3}$	0.21**	0.089	0.094	0.29***	0.13***	0.79*
	(0.0931)	(0.1037)	(0.2822)	(0.0954)	(0.0480)	(0.4168)
$\alpha^{WR}_{3\text{-}6}$	0.085	0.29**	0.12	0.29***	0.15***	0.95**
	(0.0615)	(0.1379)	(0.2081)	(0.0724)	(0.0325)	(0.4460)
$\alpha^{WR}_{7\text{-}26}$	0.081***	0.23***	0.31***	0.25***	0.097***	0.98***
	(0.0288)	(0.0367)	(0.1030)	(0.0662)	(0.0268)	(0.1777)
$\alpha^{WR}_{27\text{-}60}$	0.00016	0.071	0.13	0.44***	0.16***	0.78***
	(0.0254)	(0.0604)	(0.1298)	(0.0864)	(0.0379)	(0.2477)
α^{WR}_{60+}	0.035	0.059	-0.043	0.30***	0.40***	0.69***
	(0.0225)	(0.0444)	(0.1138)	(0.0669)	(0.0930)	(0.2183)
$\alpha^{WR}_{<3}\times$ ECC rate	0.0016	0.0078**	0.0070	0.0075	0.00061	0.026
	(0.0029)	(0.0034)	(0.0099)	(0.0049)	(0.0036)	(0.0162)
$\alpha^{WR}_{3\text{-}6}\times$ ECC rate	0.0037**	0.0029	0.0045	-0.0014	-0.0022	0.011
	(0.0016)	(0.0034)	(0.0076)	(0.0034)	(0.0018)	(0.0147)
$\alpha^{WR}_{7\text{-}26}\times$ ECC rate	-0.00054	-0.00037	0.0047*	0.0021	0.00068	0.0063
	(0.0004)	(0.0008)	(0.0028)	(0.0016)	(0.0006)	(0.0051)
$\alpha^{WR}_{27\text{-}60}\times$ ECC rate	0.0013*	0.00036	-0.0041	0.0042	-0.00046	0.0019
	(0.0007)	(0.0011)	(0.0036)	(0.0030)	(0.0019)	(0.0079)
$\alpha^{WR}_{60+}\times$ ECC rate	-0.0013	-0.0026	0.0031	-0.0023	0.0055	0.00092
	(0.0010)	(0.0016)	(0.0043)	(0.0040)	(0.0051)	(0.0122)
α^{BR}_{total}	0.0051***	0.013***	0.058***	0.036***	0.0071***	0.12***
	(0.0012)	(0.0028)	(0.0129)	(0.0067)	(0.0022)	(0.0241)
$\alpha^{WR}_{total}\times$ gdp per capita	-0.000055	0.000044	0.000018	0.00057***	0.000042	0.00058
	(0.0000)	(0.0001)	(0.0004)	(0.0002)	(0.0001)	(0.0007)
$\alpha^{WR}_{total}\times$ airport	-0.000061	-0.000094	0.000062	-0.000030	-0.00014*	-0.00014
	(0.0000)	(0.0001)	(0.0004)	(0.0002)	(0.0001)	(0.0007)
$\alpha^{WR}_{total}\times$ temperature	-0.00089**	-0.0031**	-0.0069	-0.0067***	-0.0021***	-0.020***
	(0.0004)	(0.0015)	(0.0043)	(0.0019)	(0.0005)	(0.0076)
$\alpha^{WR}_{total}\times$ rain	0.00060	0.0027	0.013**	0.0035	-0.0041***	0.017*
	(0.0005)	(0.0017)	(0.0061)	(0.0023)	(0.0009)	(0.0092)
$\alpha^{WR}_{total}\times$ absolute humidity	-0.0012	-0.0030	0.017	-0.013**	-0.0057***	-0.00067
	(0.0012)	(0.0024)	(0.0122)	(0.0054)	(0.0015)	(0.0204)
$\alpha^{WR}_{total}\times$ large cities	0.0028	0.0027	0.042	0.030	0.0066	0.088
	(0.0035)	(0.0086)	(0.0405)	(0.0213)	(0.0074)	(0.0632)
$\alpha^{WR}_{total}\times$ cities	0.0013	0.010	0.022	-0.00052	-0.0045	0.031
	(0.0023)	(0.0062)	(0.0216)	(0.0098)	(0.0042)	(0.0341)
$\alpha^{WR}_{total}\times$ rural areas with small cities	0.0013	0.0068	0.0015	-0.0012	-0.0085**	-0.00066
	(0.0021)	(0.0055)	(0.0198)	(0.0105)	(0.0040)	(0.0332)
$\alpha^{WR}_{total}\times$ population density	0.0000017	-0.00000051	-0.000049**	-0.000021**	-0.0000021	-0.000076**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	185004	185004	185004	185004	185004	185004
Adj.R2.	0.437	0.479	0.629	0.686	0.530	0.704

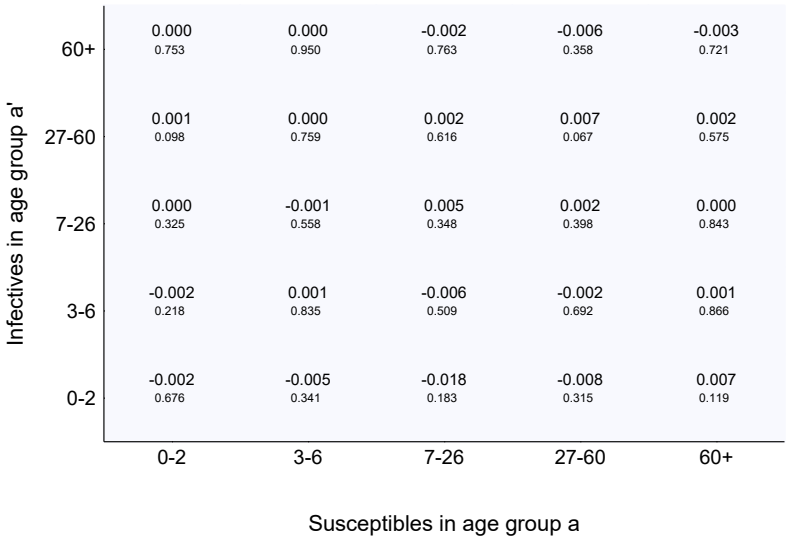
Notes: The table displays estimates of Equation 2.3. Results are based on OLS regressions. All regressions include county fixed effects, week-in-a-year fixed effects, season fixed effects, and month dummies interacted with transmission rates. Each column presents estimates from a different regression with incidence rates as the dependent variables. Regressions are weighted by population size in a particular age group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure B.8: Econometric Model 2: Age-Stratified Spatial-Time Spread of Diseases, Urban and Rural Areas

(a) Effect of Child Care on Age-specific Transmission Rates - Urban Areas



(b) Effect of Child Care on Age-specific Transmission Rates - Rural Areas



Notes: The figures present the change in the estimated transmission rates within a county ($\alpha^{WR}_{a'a}$) in response to an increase in the ECC rate in urban areas (Panel A) and in rural areas (Panel B). Cells with p-values below 0.01 are highlighted in blue. Each column presents estimates from a different regression. The dependent variables are the incidence rates of susceptibles in a particular age group. All regressions include county and week-in-a-year fixed effects. In addition, regressions control for season dummies, month dummies, climate, GDP per capita, and distance to an airport interacted with transmission rates within a county. The complete set of regression coefficients is presented in Appendix Table B.16 and Appendix Table B.17. Regressions are weighted by population size in a particular age group.

Table B.16: Econometric Model 2: Age-Specific Time-Space Model, Urban Areas

	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
$\alpha^{WR}_{<3}$	0.10	0.011	-0.013	0.28***	0.097**	0.42
	(0.1315)	(0.1239)	(0.4022)	(0.1010)	(0.0469)	(0.5782)
$\alpha^{WR}_{3\text{-}6}$	0.19**	0.64***	0.37	0.34***	0.14***	1.71***
	(0.0967)	(0.1492)	(0.2356)	(0.0854)	(0.0345)	(0.5209)
$\alpha^{WR}_{7\text{-}26}$	0.078***	0.32***	0.52***	0.30***	0.10***	1.36***
	(0.0203)	(0.0544)	(0.1213)	(0.0634)	(0.0234)	(0.1921)
$\alpha^{WR}_{27\text{-}60}$	0.039	0.14***	0.24	0.54***	0.13***	1.08***
	(0.0288)	(0.0482)	(0.1487)	(0.1272)	(0.0318)	(0.2972)
α^{WR}_{60+}	0.18**	0.17***	-0.10	0.35*	0.038	0.62*
	(0.0770)	(0.0474)	(0.1713)	(0.2009)	(0.0853)	(0.3430)
$\alpha^{WR}_{<3}\times$ ECC rate	-0.0014	0.014***	0.015	0.020***	0.00078	0.051*
	(0.0031)	(0.0039)	(0.0168)	(0.0068)	(0.0045)	(0.0265)
$\alpha^{WR}_{3\text{-}6}\times$ ECC rate	0.0078***	0.0017	0.0076	0.00084	-0.00038	0.020
	(0.0021)	(0.0043)	(0.0111)	(0.0051)	(0.0026)	(0.0202)
$\alpha^{WR}_{7\text{-}26}\times$ ECC rate	-0.00097**	-0.00029	0.0042	0.0016	0.00083	0.0049
	(0.0004)	(0.0010)	(0.0035)	(0.0021)	(0.0005)	(0.0065)
$\alpha^{WR}_{27\text{-}60}\times$ ECC rate	0.0017	0.0012	-0.0081	0.0011	-0.0026	-0.0056
	(0.0011)	(0.0017)	(0.0062)	(0.0038)	(0.0018)	(0.0117)
$\alpha^{WR}_{27\text{-}60}\times$ ECC rate	-0.0031**	-0.0030	0.0093	0.0049	0.010**	0.016
	(0.0015)	(0.0026)	(0.0071)	(0.0042)	(0.0046)	(0.0133)
α^{BR}_{total}	0.0058***	0.011***	0.059***	0.032***	0.0045*	0.12***
	(0.0018)	(0.0036)	(0.0196)	(0.0095)	(0.0026)	(0.0358)
$\alpha^{WR}_{total}\times$ GDP per capita	-0.00011**	-0.000091	0.00026	0.00054**	0.000044	0.00065
	(0.0000)	(0.0002)	(0.0005)	(0.0002)	(0.0001)	(0.0008)
$\alpha^{WR}_{total}\times$ airport	-0.00013**	-0.00030*	-0.00022	-0.00022	-0.00015	-0.00092
	(0.0001)	(0.0002)	(0.0006)	(0.0003)	(0.0001)	(0.0010)
$\alpha^{WR}_{total}\times$ temperature	-0.0011*	-0.0048**	-0.0069	-0.0062**	-0.0020***	-0.022*
	(0.0006)	(0.0024)	(0.0060)	(0.0027)	(0.0007)	(0.0110)
$\alpha^{WR}_{total}\times$ rain	0.00078	0.0016	0.022***	0.0076**	-0.0032**	0.031***
	(0.0007)	(0.0026)	(0.0073)	(0.0031)	(0.0013)	(0.0118)
$\alpha^{WR}_{total}\times$ absolute humidity	-0.0011	-0.0077**	-0.0046	-0.018**	-0.0058***	-0.034
	(0.0023)	(0.0038)	(0.0187)	(0.0087)	(0.0018)	(0.0316)
$\alpha^{WR}_{total}\times$ population density	0.00000055	-0.00000086	-0.000052**	-0.000028***	-0.00000091	-0.000089**
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
$\alpha^{WR}_{total}\times$ cities	0.0028	0.012**	0.0092	0.00064	-0.0085*	0.016
	(0.0020)	(0.0055)	(0.0216)	(0.0107)	(0.0044)	(0.0349)
Observations	103922	103922	103922	103922	103922	103922
Adj.R2.	0.503	0.550	0.627	0.698	0.582	0.705

Notes: The table displays estimates of Equation 2.3 for urban areas. Results are based on OLS regressions. The regressions control for county and week-in-a-year fixed effects in levels and season fixed effects, and month dummies interacted with transmission rates. Each column presents estimates from a different regression with incidence rates as the dependent variables. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Table B.17: Econometric Model 2: Age-Specific Time-Space Model, Rural Areas

	Age: <3	**Age: 3 to 6**	**Age: 7 to 26**	**Age: 27 to 60**	**Age: 60+**	**Total**
$\alpha^{WR}_{<3}$	0.31***	0.12	0.23	0.31*	0.050	1.06*
	(0.1053)	(0.1522)	(0.3773)	(0.1724)	(0.0654)	(0.5864)
$\alpha^{WR}_{3\text{-}6}$	0.020	-0.045	-0.16	0.17*	0.12***	0.082
	(0.0412)	(0.1054)	(0.2844)	(0.0952)	(0.0419)	(0.4681)
$\alpha^{WR}_{7\text{-}26}$	0.076	0.15***	0.076	0.22**	0.070*	0.58**
	(0.0474)	(0.0454)	(0.1484)	(0.1012)	(0.0376)	(0.2548)
$\alpha^{WR}_{27\text{-}60}$	-0.039	0.050	0.062	0.38***	0.18***	0.63*
	(0.0503)	(0.0804)	(0.1984)	(0.1105)	(0.0560)	(0.3451)
α^{WR}_{60+}	-0.0040	-0.052	-0.17	0.26***	0.46***	0.41
	(0.0196)	(0.0553)	(0.1477)	(0.0876)	(0.0680)	(0.2740)
$\alpha^{WR}_{<3}\times$ ECC rate	-0.0015	-0.0050	-0.018	-0.0075	0.0073	-0.028
	(0.0036)	(0.0053)	(0.0134)	(0.0075)	(0.0046)	(0.0239)
$\alpha^{WR}_{3\text{-}6}\times$ ECC rate	-0.0020	0.00066	-0.0065	-0.0018	0.00051	-0.0076
	(0.0016)	(0.0032)	(0.0098)	(0.0045)	(0.0030)	(0.0174)
$\alpha^{WR}_{7\text{-}26}\times$ ECC rate	0.00048	-0.00069	0.0046	0.0017	0.00018	0.0068
	(0.0005)	(0.0012)	(0.0049)	(0.0020)	(0.0009)	(0.0077)
$\alpha^{WR}_{27\text{-}60}\times$ ECC rate	0.0012*	0.00037	0.0021	0.0073*	0.0016	0.013
	(0.0007)	(0.0012)	(0.0043)	(0.0040)	(0.0028)	(0.0097)
$\alpha^{WR}_{27\text{-}60}\times$ ECC rate	0.00043	0.00010	-0.0015	-0.0058	-0.0029	-0.0088
	(0.0014)	(0.0017)	(0.0050)	(0.0063)	(0.0080)	(0.0179)
α^{BR}_{total}	0.0046***	0.016***	0.054***	0.039***	0.015***	0.13***
	(0.0016)	(0.0044)	(0.0176)	(0.0093)	(0.0043)	(0.0327)
$\alpha^{WR}_{total}\times$ GDP per capita	0.00021	0.00055	-0.00055	0.000078	0.00018	0.00040
	(0.0001)	(0.0003)	(0.0016)	(0.0008)	(0.0003)	(0.0025)
$\alpha^{WR}_{total}\times$ airport	-0.0000099	0.000088	0.00048	0.00011	-0.00014	0.00065
	(0.0001)	(0.0001)	(0.0005)	(0.0003)	(0.0001)	(0.0009)
$\alpha^{WR}_{total}\times$ temperature	-0.00080	-0.0012	-0.0060	-0.0066***	-0.0020**	-0.017**
	(0.0006)	(0.0012)	(0.0052)	(0.0020)	(0.0009)	(0.0083)
$\alpha^{WR}_{total}\times$ rain	0.00052	0.0049***	0.010	0.00013	-0.0052***	0.010
	(0.0006)	(0.0017)	(0.0097)	(0.0031)	(0.0012)	(0.0122)
$\alpha^{WR}_{total}\times$ absolute humidity	-0.00061	0.0028	0.041***	-0.0081	-0.0043**	0.037
	(0.0012)	(0.0030)	(0.0139)	(0.0060)	(0.0022)	(0.0229)
$\alpha^{WR}_{total}\times$ population density	-0.0000065	-0.000022**	-0.000043	0.000013	-0.000014	-0.000067
	(0.0000)	(0.0000)	(0.0001)	(0.0000)	(0.0000)	(0.0001)
$\alpha^{WR}_{total}\times$ sparsely populated areas	0.0028	0.012**	0.0092	0.00064	-0.0085*	0.016
	(0.0020)	(0.0055)	(0.0216)	(0.0107)	(0.0044)	(0.0349)
Observations	81082	81082	81082	81082	81082	81082
Adj.R2.	0.364	0.418	0.647	0.682	0.487	0.714

Notes: The table displays estimates of Equation 2.3 for rural areas. Results are based on OLS regressions. The regressions control for county and week-in-a-year fixed effects in levels and season fixed effects, and month dummies interacted with transmission rates. Each column presents estimates from a different regression with incidence rates as the dependent variables. Regressions are weighted by population size in a particular age group. Standard errors are clustered at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B6 Cost and Benefit Analysis

Table B.18: Costs of Influenza per Case

	Direct Costs (in €)			Indirect Costs (in €)	Costs of Death (in €)	Total (in €)
	Outpatient Costs	Inpatient Costs	Pharamceuticals		VSL × Risk of Death	
Age 0 to 6	47	73.4	7.76	-	1.3 – 6	129.48 - 134.18
Age 7 to 26	50	33	7	-	1.3 – 6	111.54 – 116.24
Age 27 to 60	53.85	14.39	6	338.76	6.5 – 42	419.5 – 455
Age 60+	65.71	58.48	7.40	83.91	22.1–102	237.6 – 317.2

Notes: Data on direct and indirect costs are taken from Scholz et al. (2019). The study calculates medical costs weighted by the probability of health care usage separately for different age groups. Note that Scholz et al. (2019) use slightly different age groups to those in this study. Hence, the numbers are to be regarded as approximations. The value of statistical life is assumed to be between €1.3-6 million. Mortality risk due to an infection with influenza varies across age groups. For people aged 0 to 26, it is 0.1 per 100,000 inhabitants, for people aged 27 to 60, it is 0.5 per 100,000 inhabitants and for individuals above age 60 1.7 per 100,000 inhabitants.[47]

Table B.19: Relative and Absolute Number of Influenza Cases

	System of Notifiable Diseases		Insurance Claims	
	Share of Infected	Number of Cases	Share of Infected	Number of Cases
Age 0 to 2	0.108%	1816.96	2%	33700.44
Age 3 to 6	0.144%	3276.99	2.31%	52981.04
Age 7 to 26	0.039%	5558.81	1.4%	198251.20
Age 27 to 60	0.024%	7318.24	1.17%	363989.44
Age 60+	0.015%	2438.90	0.57%	91964.04

Notes: The table summarizes the average annual number of individuals infected with influenza in absolute and relative terms. Columns 1 and 2 consider cases reported to the system of notifiable diseases between 2005 and 2016. Column 3 reports the share of clinically diagnosed cases with influenza-like diseases taken from Scholz et al. (2019). Scholz et al. (2019) refer to data from health insurance claims covering about 10% of the German population between 2012 and 2014. Column 4 is based on the author's calculations. The absolute number of cases with flu-like diseases is calculated by the share of infected, based on health claims times the age-specific population size in West Germany.

[47] See Gesundheitsberichterstattung des Bundes (2021).

Table B.20: Effect of Policy Interventions on Disease Incidence

	No. of Avoided Cases (in %)	Absolute No. of Avoided Cases	
		System of Notifiable Diseases	Insurance Claims
Policy - Mandatory Vaccination before Entry into Child Care			
Age 0 to 2	-10.8%	-196.23	-3639.56
Age 3 to 6	-3%	-98.36	-1590.32
Age 7 to 26	-1.5%	-84.16	-3001.41
Age 27 to 60	-0.8%	-60.26	-2997.08
Age 60+	-0.3%	-7.05	-266.02
Policy - Closure of Child Care Centers during Local Outbreaks (Def. 1)			
Age 0 to 2	-2.4%	-44.36	-822.86
Age 3 to 6	-2%	-64.99	-1050.77
Age 7 to 26	-1.8%	-100.26	-3575.79
Age 27 to 60	-2%	-148.31	-7376.64
Age 60+	-1.1%	-28.02	-1056.44
Policy - Closure of Child Care Centers during Local Outbreaks (Def. 2)			
Age 0 to 2	-4.1%	-74.98	-1390.73
Age 3 to 6	-3.5%	-113.19	-1829.98
Age 7 to 26	-3.1%	-174.98	-6240.52
Age 27 to 60	-3.3%	-243.40	-12106.25
Age 60+	-1.7%	-42.31	-1595.57

Notes: The table summarizes the impact of two policy interventions – mandatory vaccination before entry into child care and closure of child care centers during local outbreaks – on average annual incidence rates of influenza. Column 1 reports the estimated effect of the policy interventions on disease incidence in percent per age group. Column 2 summarizes the absolute average number of cases avoided due to the policy interventions based on lab-confirmed cases. Analogously, Column 3 reports the the number of cases reduced due to the containment measures based on health insurance claims data.

Table B.21: Estimated Benefits and Costs, by Policy Intervention

	Mandatory Vaccination		**Closure of Child Care Centers**			
			Outbreak Def. 1		Outbreak Def. 2	
	Notifiable System	Claims	Notifiable System	Claims	Notifiable System	Claims
Annual Benefits (in €)	87,126	2,981,324	122,798.	5,488,629	203,821.9	9,068,098
Annual Costs (in €)	2,306,590		-	-	-	-
Net Benefits (in €)	-2,219,464	+ 674,734	-	-	-	-

Notes: The table reports the calculated economic benefits of the two policy interventions, i.e. mandatory vaccination and the closure of child care centers based on the data from the system of notifiable diseases (mainly lab-diagnosed cases) and health insurance claims (clinically diagnosed cases). In addition, it presents the costs and net benefits linked to the introduction of mandatory vaccination policies.

Chapter 3

The Role of Information When Risk Levels Change: Evidence from COVID-19*

Abstract

This paper investigates the role of public policies that provide information on changes in health risks. We study this in the context of Covid-19, where behavioral responses have been important in mitigating the spread of the virus and, hence, the costs of the pandemic. To identify behavioral responses induced by public information versus other sources, we combine high-frequency data with facts about the incubation period of Covid-19 and reporting time in Germany. Using an event study design on local unexpected outbreaks, we find that mobility significantly decreases by about 2 to 3% in response to public information about the outbreak, while private knowledge about people falling sick does not appear to cause a change in behavior. There are important heterogeneities in the behavioral responses, where responses are stronger in counties with high population density, with more hotels per capita, and with a higher share of college educated. These findings are consistent with behavioral changes depending on the relative risk and costs of changing mobility.

*This chapter is based on joint work with Pavel Obraztcov, Gregory Veramendi, and Joachim Winter.

3.1 Introduction

Individuals often face health decisions under uncertainty (e.g. Banerjee et al. 2020; Dupas 2011; Kim et al. 2019; Oster 2012, 2017; Prina and Royer 2014). One of the primary roles of public health agencies is to provide information on health risks so that individuals can make better choices. Information on public health risks may be particularly important when risk levels change depending on the state of the world. In this case, individual responses can be crucial for mitigating the costs of these risks. An important public policy is providing quick and credible information about the state of the world, not only for policy-makers to make decisions, but also for individuals to be able to respond quickly to changes in risk levels. Yet it remains unclear to what extent individuals react to information on health risks. Some studies argue that people's behavior is sensitive to information on health risk (Chan et al. 2016; Oster 2017; Philipson 2000), while others find that individuals appear reluctant to undertake costly behaviors with health benefits in response to health information (Kim et al. 2019; Oster 2012, 2018; Prina and Royer 2014).

The Covid-19 pandemic is one setting where individual behavior is especially important for mitigating the spread and, hence, the costs of the virus. Covid-19[1] is an *overdispersed* pathogen, in that a small fraction of individuals is responsible for a large fraction of the transmission. In the case of Covid-19, studies have shown that only 10–20% of individuals are responsible for 80–90% of transmission clusters often called "super-spreader" events (see e.g. Baggett et al. 2020; Endo et al. 2020; Hamner et al. 2020; James et al. 2020; Lemieux et al. 2020; Majra et al. 2021; Riou and Althaus 2020).[2] This feature of the virus implies that the local risk level can change dramatically when a transmission cluster occurs. In this case, having quick, credible, and localized information about outbreaks can allow individuals to change their behavior with respect to the local state of the epidemic and be an important part of the mitigation of the outbreak.

This paper studies the role of publicly-provided information in mitigating the pandemic. We estimate the effect of information on positive cases using an event study approach that examines the changes in mobility patterns in a county when an unexpected outbreak occurs.[3] In other settings, it would be difficult to isolate the role of a specific source of information from other sources without experimental variation. In the case of Covid-19, we combine high-frequency data with known features of the incubation time and reporting time to isolate the role of public information on positive cases from other possible confounding explanations. Briefly, Covid-19 has a median incubation period of about five days (Guan et al. 2020; Lauer et al. 2020; Li et al. 2020) and there is a median time of six days from first symptoms to the reporting of a case

[1] The expressions Covid-19, coronavirus and SARS-CoV-2 are used interchangeably.

[2] Some commonly known "super-spreader" events or clusters are the example of a Korean woman infecting 1000+ others in a few days, a woman in Garmisch-Partenkirchen that attended several bars despite of symptoms and caused jump in local infection rates in September. Another famous example is a wedding in Hamm in the beginning of September.

[3] Mobility is measured using cell phone data on the number of trips taken within and across counties each day. See Section 3.3 for more information.

by the German national public health authority.[4] Hence, any behavioral changes in mobility due to private information of individuals getting sick should lead to changes in mobility in the seven to ten days before the outbreak is publicly observed. In contrast, behavioral responses caused by public information will occur after the start of the observed outbreak. The use of daily data on positive cases and mobility patterns allows us to distinguish mobility changes due to different sources of information.

We identify local unexpected outbreaks by comparing the number of cases in a seven-day period to the number of cases we would expect based on a parsimonious epidemiological model. We define a local unexpected outbreak as a county and seven-day period where the observed number of cases exceeds the expected number based on our model. This definition captures all of the well-known outbreaks in Germany, along with many others.[5] Our preferred specification identifies 259 outbreaks between February and November.[6]

We use the outbreaks to perform an event study of county-level mobility near the time of the outbreak controlling for state times day and county times day-of-the-week fixed effects. Controlling for non-pharmaceutical interventions (NPIs) makes little difference in our estimates which indicates that we are capturing a voluntary response in mobility. We find that public information about local outbreaks significantly reduces the number of trips taken inside the county and also reduces the number of trips between other counties and the outbreak county. The number of trips is reduced by between 1.8% to 2.5%. We do not, however, find any changes in the number of trips before the beginning of the observed outbreak, ruling out changes due to private knowledge of a risky event or private knowledge of individuals with symptoms.

We further investigate heterogeneity in response to local outbreaks along three distinct dimensions: Counties may differ in (i) costs of adjusting mobility, e.g. depending on the fraction of workers that can work from home, (ii) relative risks of infection, e.g. depending on transportation modes and structure of urbanization, and (iii) average beliefs about the risk of the virus. We find that counties with larger tourism sector and a larger share of workers that can work from home[7] react stronger to local outbreaks indicating that behavioral responses are indeed larger if costs of adjusting mobility are lower. We do not, however, find evidence that use of public transport or beliefs on the risk of Covid-19[8] affect people's reaction to local outbreaks. Interestingly, we find that the effect size

[4] There is a median time of four days between first symptoms and a positive test results. In addition, there is a 1 to 2 day lag until information on positive tests is compiled and made publicly available by the national pubic health agency (see Section 3.3.1).

[5] e.g. Heinsberg in February 2020, Gütersloh in June 2020, and Berchtesgaden in October 2020.

[6] Our preferred specification requires that the excess cases in a county is at the 98th percentile for all counties in Germany between February and November 2020. See Section 3.4.2 for more information.

[7] We take the share of people with a college-degree as measure for the likelihood of working from home. Previous literature shows that people with higher educational degree are more likely to do home office in response to the pandemic (Gaudecker et al. 2020).

[8] We use the vote share of the AfD in the last state elections as a proxy for beliefs. A number of studies in the US have shown that partisanship affects people's beliefs on the risks of Covid-19. Similarly in Germany, the AfD party was critical of the government handling of Covid-19 and one may expect that counties with high AfD vote shares would, likewise, respond less to an outbreak.

of the response to local outbreaks is larger during the 2nd wave of the pandemic in autumn 2020 than during the 1st wave in spring 2020. Hence, while people may generally fatigue about the pandemic and its restrictions, intrinsic responses to localized outbreaks appear not to be affected by this phenomenon. Finally, we provide evidence that our results are robust to a number of sensitivity checks.

This paper contributes to several strands of literature. First, the paper adds to a growing strand in economics studying the role of information in shaping people's health behavior and beliefs (e.g. Bollinger et al. 2011; Cawley et al. 2020; Dupas 2011; Oster 2017; Wisdom et al. 2010). This body of literature investigates whether people adjust their behavior to information on health risk. Yet evidence on that matter is inconclusive. A strand of literature shows that people appear reluctant to undertake costly behaviors with corresponding health benefits. For example, people are resistant to sexual behavior change in the face of HIV, to change diet in response to a diabetes diagnosis, and lack regular cancer screening (Caldwell et al. 1999; Cummings and Cooper 2011; Hut and Oster 2018; Kim et al. 2019; Oster 2012, 2018; Prina and Royer 2014). Another strand of literature, however, argues that people are sensitive to changes in health risk and demand for self-protection (Chan et al. 2016; Kremer 1996; Oster 2017; Philipson 2000). These studies document a prevalence-elasticity of private demand for prevention against disease. By investigating whether people adjust their behavior voluntarily beyond social distancing policies, we provide evidence on people's intrinsic willingness to change behavior in response to a change in health risk.

Second, in the light of the Covid-19 pandemic, a recent strand of literature investigates the effect of information about the risk of the virus on people's behavioral response and ultimately, on mitigating the spread of the disease. These studies examine various channels of information provision such as Twitter, TV shows, the word of political leaders, and also experimentally provided information in different social and political environments, e.g. India, Mexico, Brazil, and USA. They jointly show that providing information significantly determines people's health behavior (Ajzenman et al. 2020; Banerjee et al. 2020; Brzezinski et al. 2020; Grossman et al. 2020; Gutierrez et al. 2020). In the context of fast-changing risk levels it is crucial that people *quickly* obtain information on *local* risk and respond to a change in localized risk. This study adds to this question by investigating the timing of public information measured by the number of reported cases and people's behavioral response.

Finally, the paper relates to a strand of literature in economics and epidemiology investigating the effect of people's travel activities on the spread of infectious diseases (e.g. Adda 2016; Brockmann et al. 2006; Grenfell et al. 2001; Hufnagel et al. 2004; Oster 2005, 2012). Prior work documents that travel routes and intensity significantly impact transmission patterns of different viruses such as HIV, influenza, gastritis, and measles in various countries, e.g. Sub-Saharan countries, USA, and European countries. In the light of the COVID-19 pandemic, a fast-growing body of literature aims to identify the effect of mobility on the propagation of the coronavirus and discusses the contribution of NPIs to mitigating the spread of the virus by reducing physical contact between

people (Fang et al. 2020; Ferguson et al. 2020; Glaeser et al. 2020; Gupta et al. 2020). This paper studies the reverse relationship, namely the impact of unanticipated increases in infection rates on mobility uncovering intrinsic avoidance behavior due to a change in risk levels.

The paper proceeds as follows. Section 3.2 gives background information on the characteristics of Covid-19 and describes the experience of the pandemic in Germany including the political and policy responses. Section 3.3 presents the data and descriptive statistics. Section 3.4 outlines the empirical strategy. Section 3.5 presents the results and Section 3.6 concludes.

3.2 Background

In this section, we first give an overview of the characteristics of Covid-19 relevant to our analysis. Then we describe the experience of the pandemic in Germany including the political and policy responses.

3.2.1 Characteristics of Covid-19

The Covid-19 pandemic is a major threat to human health. According to the World Health Organization, there have been 112 million confirmed cases and 2.5 million deaths worldwide.[9] A SARS-CoV-2 infection primarily causes respiratory diseases with symptoms ranging from mild upper respiratory tract illness to severe pneumonia with acute respiratory distress syndrome and death (Chen et al. 2020; Huang et al. 2020). At the same time, many individuals infected with the virus never develop symptoms (Huang et al. 2020). Hence, the clinical spectrum appears to be wide, encompassing asymptomatic infections, mild diseases comparable to a common cold, as well as severe lower respiratory tract diseases with many patients being hospitalized, and death (Gandhi et. al 2020b). Potential long-term consequences of an infection with Covid-19, such as fatigue and dyspnoea, are currently studied (Huang et al. 2021; Zhao et al. 2020). Recent estimates suggest that the infection fatality rate (IFR) of Covid-19 is about 1%, which is substantially higher than, for example, the estimated IFR of 0.1% of influenza (Pritsch et al. 2020; Staerk et al. 2020; World Health Organization 2020a). Particularly, the elderly population and individuals with underlying medical conditions, such as cardiovascular diseases, diabetes or chronic lung illnesses, are at risk of developing a severe course of Covid-19 infections (Centers for Disease Control and Prevention 2021; Wu and McGoogan 2020; Zhou et al. 2020).

Covid-19 spreads rapidly – within a couple of months, the world turned from a few reported cases in the city of Wuhan in China to a state in which almost all countries were affected by the new coronavirus (Bloom et al. 2020). The transmission of Covid-19 occurs from human-to-human, primarily through droplets, aerosols, and close contact with infected individuals (Gandhi et. al

[9] See https://covid19.who.int/ accessed on 27th February 2021. For comparison, seasonal influenza causes worldwide approximately 300,000–500,000 deaths per year (Girard et al. 2005; Lambert and Fauci 2010).

2020b). The transmissibility of Covid-19 exhibits two disease-specific features relevant for our study: First, there is growing evidence in the medical literature that Covid-19 is an overdispersed virus, where a small fraction of individuals is responsible for a large fraction of the transmission (see e.g. Baggett et al. 2020; Endo et al. 2020; Hamner et al. 2020; James et al. 2020; Lemieux et al. 2020; Majra et al. 2021; Riou and Althaus 2020). Second, in contrast to other established viruses[10], presymptomatic and asymptomatic patients are infectious (Gandhi et. al 2020a). For symptomatic cases, the median incubation period is four to five days, with a range from 2 to up to 14 days (Gandhi et. al 2020b; Lauer et al. 2020).[11]

In the absence of a Covid-19 vaccine, prevention of infections was limited to reducing the physical proximity between individuals and to wearing face masks (Chu et al. 2020; Ferguson et al. 2020).[12] A number of public health measures (NPIs) were aimed at reducing contact rates in the population and thereby mitigating the spread of the virus. Examples of the policy interventions adopted include closing schools, restaurants and retail events, and contact banning.

3.2.2 Covid-19 Pandemic and Policy Response in Germany

At the end of January 2020, the World Health Organization declared the outbreak of the novel Coronavirus SARS-CoV-2 (i.e. Severe Acute Respiratory Syndrome Coronavirus 2) as a "public health emergency of international concern" (World Health Organization 2020b). At this time the epicenter of the outbreak was Wuhan city, the capital of Hubei province in China. About a month later, the virus had reached Europe and started to spread uncontrolled within Germany.[13] As a consequence, the German government introduced a number of regulations to limit social contacts and thus, the diffusion of the Covid-19 virus.[14] The political response can be classified into at least four stages (Figure 3.1): The first stage is characterized by the increasing political and social awareness due to a growing number of local outbreaks such as in the county of Heinsberg after a carnival party. To curb the spread of the virus, the government appealed to all citizens to avoid social contacts whenever possible (e.g. Angela Merkel held nationally-televised speeches on March 12 and 18) and started to gradually impose social distancing policies. Most schools, childcare facilities, and retail stores were closed starting on March 16th onwards. Travel restrictions, such as enhanced controls at the borders and a 30-day entry ban for non-EU inhabitants, were enforced. The second stage is described by a national contact ban. During that period it was prohibited to meet more than one person from outside one's household and required to keep a minimum distance

[10] e.g. SARS-CoV-1 causing a pandemic in 2003 (Gandhi et. al 2020a).

[11] Additional information is drawn from the European Center for Disease Prevention and Control (European Centre for Disease Prevention and Control 2021a) and from the Robert Koch Institute (Robert Koch Institute 2021c)

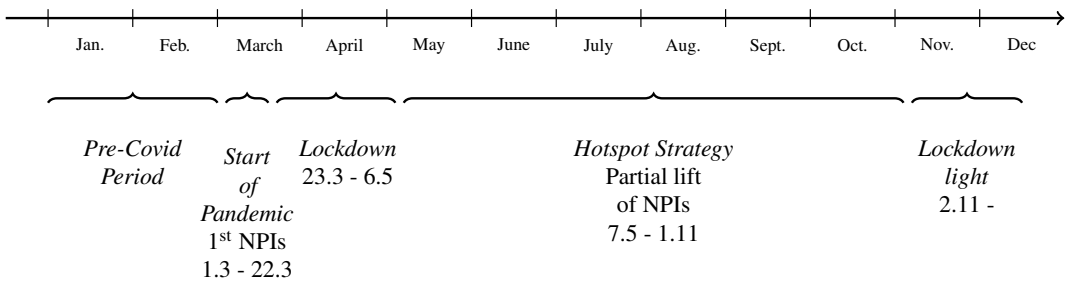
[12] Note that recently some vaccines for the protection against Covid-19 have been approved. However, during the period we consider in our study no vaccines against Covid-19 were available.

[13] In Germany, COVID-19 began to propagate uncontrolled after the detection of two cases in the end of February, 2020. An earlier outbreak at the end of January had been completely contained.

[14] Information provided in this subsection is drawn from the German ministry of health (Bundesministerium für Gesundheit 2021) (in German) and from the European Center for Disease Prevention and Control (European Centre for Disease Prevention and Control 2021b).

of 1.5 meters (Glogowsky et al. 2020). By mid-April incidence rates started to decrease, which allowed authorities to gradually relax social distancing policies. During the third stage (May to October), incidence rates were comparatively low on the national level. Local outbreaks, however, led to temporary rises in the disease rates at the county level. One prominent example of a cluster event is the outbreak of Covid-19 in a meat processing plant in Gütersloh. The government agreed to respond locally to outbreaks, once the seven-day incidence rate within a county exceeded 50 cases per 100,000 inhabitants. Travel restrictions within the EU were largely removed during the summer months. In October, the number of local outbreaks increased rapidly, so that local policy interventions were not sufficient any more. As a response, the German government announced enhanced regulations at the national level from November, 2nd onwards. The fourth stage is described as so-called "lockdown light" – schools, child care facilities, and retail shops remained open, while restaurants and bars were closed.

Figure 3.1: Periods of the Covid-Pandemic in Germany



Notes: The figure illustrates the different periods of political response to the Covid-19 pandemic in Germany.

3.3 Data and Descriptive Statistics

In this section, we describe the data we use to build a balanced panel for 401 districts (i.e. NUTS Level 3 regions) and 350 days spanning the period from January 1st to December 20th, 2020 (140,350 observations).[15]

[15]Due to technical reasons, the mobility data is missing for 4 days in December.

3.3.1 Covid-19 Incidence Data

We use daily data on laboratory-confirmed cases of Covid-19 within Germany. The data is provided by the national public health institution, the Robert Koch Institute (RKI).[16] According to the Protection Against Infection Act (*Infektionsschutzgesetz*), lab-diagnosed cases of Covid-19 are reported to the local public health departments in order to monitor the temporal and spatial diffusion of the virus within Germany.[17] The reporting system follows strict rules to improve comparability across regions: Physicians and laboratories are obliged to inform local health departments about a positive test result within at most 24 hours. The local health departments, in turn, deliver the information to the health authorities of the respective federal state with at most a one day delay. The state health department gathers the information and passes it to the Robert Koch Institute at the national level on the same or the next working day. Only reported cases that fulfill well-defined criteria are included in the data set (*Infektionsschutzgesetz, §11*). While the standardized procedure ensures high data quality, as well as comparability across time and regions, it creates a two day lag between the registration of positive test results at the local health department and publicly provided information on new Covid-19 cases by the Robert Koch Institute.[18]

The data set covers information on the day of reporting to the local health department, day of first symptoms, and the county of residence of the individual infected. Individuals are not included in the data set if they are not laboratory tested. Hence, the data likely understates the actual number of cases of Covid-19 as a study conducted in Munich shows (Pritsch et al. 2020).[19]

Figure 3.2 illustrates the daily number of cases reported to the Robert Koch Institute between the end of February 2020 and December 2020. Germany experienced two waves of Covid-19 infections: The first wave started at the end of February and continued to the end of April. The second wave started in October. We classify the pandemic into three phases – 1st wave, summer, and 2nd wave – to study heterogeneous response behavior over time. Appendix Table C.1 presents the summary statistics of the new infections as counts and incidence rates separately by phase of the pandemic and by federal states.

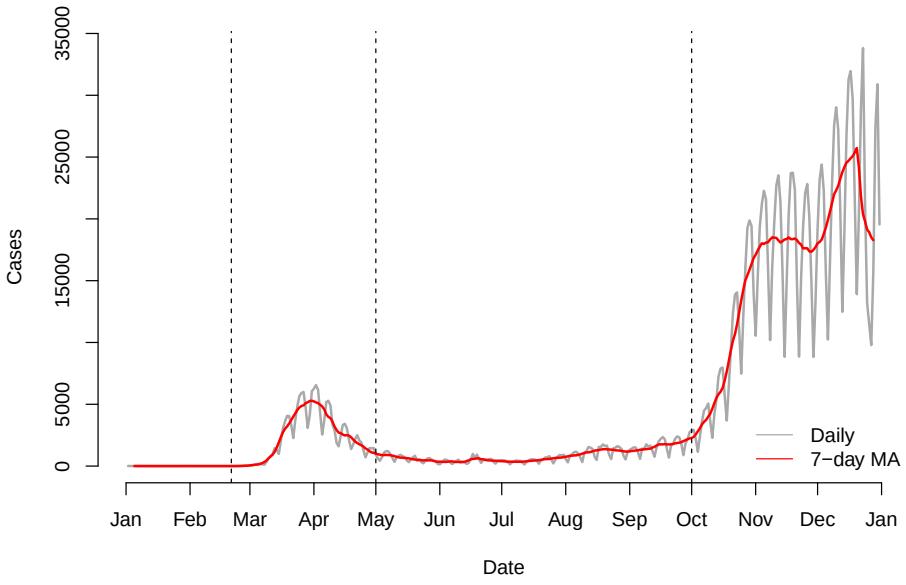
[16]The data is publicly available via COVID-19 Data Hub (COVID-19 Datenhub 2021).

[17]The information is taken from the official webpage of the Robert Koch Institute (in German) (Robert Koch Institute 2021b).

[18]Note that local health departments also publish information on positive test results which may create some discrepancy between information published by local health departments and the Robert Koch Institute. Nationwide newspapers such as "Bild", "Frankfurter Allgemeine Zeitung", and "Handelsblatt" use the RKI data as source for providing information on county-specific incidence rates (websites accessed on 3rd March 2021).

[19]The study shows that in a representative sample of 2,994 private households living in Munich 1.82% individuals are tested positive for SARS-CoV-2 specific antibodies indicating that these individuals are/were infected with Covid-19. During the same time period, however, only 0.46% of the citizens in Munich have reported a positive PCR-test result to the national health agency. Hence, the reported number of cases likely understates the "true" number of individuals infected with Covid-19.

Figure 3.2: Cases of Infection



Notes: The figure illustrates the daily number of cases (grey line) and the 7-day moving average of the cases (red line) reported by the Robert Koch Institute between January 1st, 2020 and December 31st, 2020. The vertical dotted grey lines present the different phases of the pandemic. The first phase describes the 1st wave of the pandemic between the end of February and the end of April. The second phase covers the summer months. The third phase refers to the 2nd wave of the pandemic starting in October.

3.3.2 Data on Aggregate Mobility

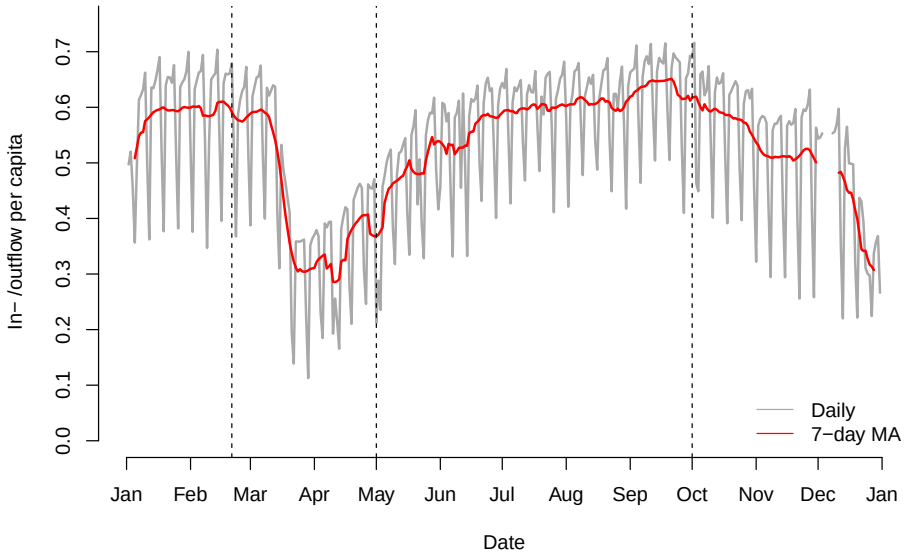
To measure mobility patterns in Germany, we use cell phone data provided on a daily basis. To ensure connectivity phones switch between cell towers, when cell phone users move. These cell tower switches are used to estimate the number of trips taking place between two geographic areas which is a proxy of human mobility over time and space (Oliver et al. 2020). We obtain data on the daily number of trips between and within counties in Germany for the period from January 1st, 2020 to December 31st, 2020 from Teralytics, a business partner of Telefónica.[20]

Panel A in Figure 3.3 describes the average daily number of inflows per capita in a county between January and December 2020. By late March the number of trips decreased substantially by about 50%. During the summer months mobility increased again and remained stable at a level comparable to the number of trips per capita in January and February. Starting in October, we observe a reduction in mobility again. Panel B in Figure 3.3 outlines the average daily number of trips within a county. The change in mobility over time follows a similar pattern as for mobility

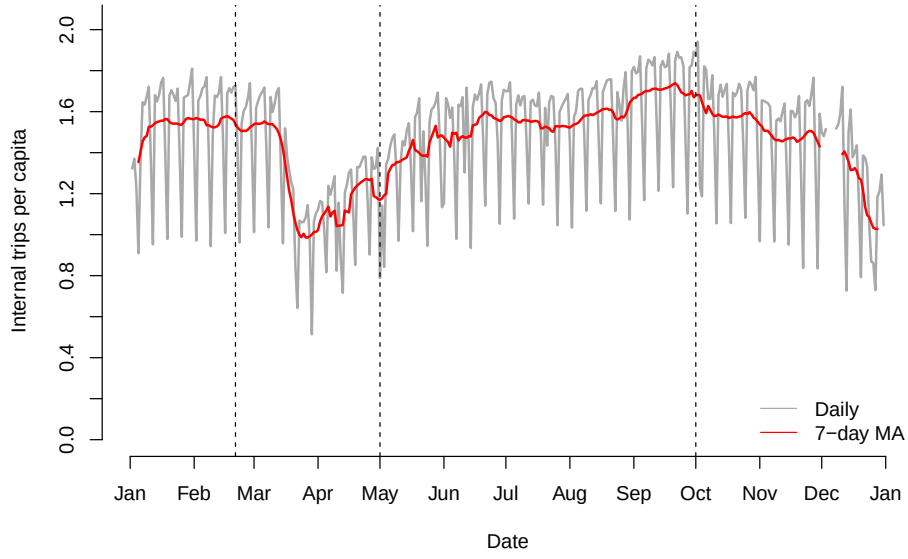
[20]Note that origin-destination pairs that have less than 5 trips taking place between counties are not included in the data set.

Figure 3.3: Mobility Patterns

(a) Inflow/Outflow per capita



(b) Internal Trips per capita



Notes: The figures display the daily number of trips per capita (grey line) and the 7-day moving average of the number of trips (red line) between January 1st, 2020 and December 31st, 2020. Panel A illustrates the average inflow (outflow) per capita into (out of) a county. Panel B summarizes the mean number of trips per capita within a county. The vertical dashed lines outline different phases of the pandemic.

between counties. Appendix Table C.1 presents the summary statistics for the number of trips per capita by phase of the pandemic and federal states.

3.3.3 Additional Data

We supplement the data on disease incidence and mobility with information on local NPIs. Data on local NPIs in Germany are provided by an interdisciplinary consortium that works on behalf of the German Ministry of Economics and Energy.[21] The consortium collects information on NPIs based on data provided on official websites of state governments. Appendix Table C.3 presents summary statistics of the NPIs considered in this study. The consortium also publishes information on county characteristics, such as the share of AfD votes, use of public transport, and the share of college educated, that we use to study heterogeneous behavior in response to local outbreaks. Appendix Table C.2 presents the summary statistics of additional variables.

3.4 Empirical Strategy

In this section, we describe our empirical strategy for distinguishing different sources of information, identifying unexpected outbreaks, and studying the aggregate behavioral response to the unexpected outbreaks. We start by describing the strategy used to disentangle changes in mobility caused by different sources of information (Section 3.4.1). In Section 3.4.2, we outline a model for predicting the incidence rate in each county on each day in Germany and classify outbreaks as an excess of cases above what is predicted by the model. In Section 3.4.3, we present both event study and difference-in-differences models for estimating the change in mobility due to the outbreak. Section 3.4.4 assesses the identification strategy.

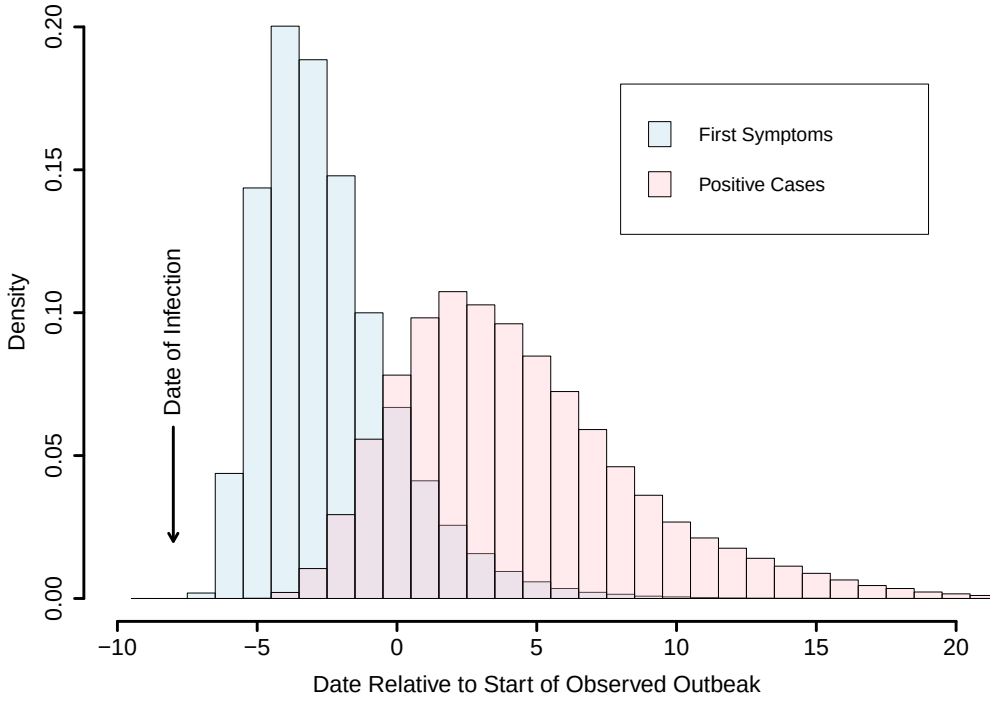
3.4.1 Identification of Information Sources

The timing of disease progression and delays in testing allow us to define distinct periods during which individuals receive information from two different sources: (i) Private information on people falling sick and (ii) Public information on reported Covid-19 cases.

If a local event sparks an unexpected outbreak, individuals may obtain private information on people falling sick about 4 to 7 days (inter-quartile range) after the event (Lauer et al. 2020). Data on disease incidence, further, document that individuals get tested about 5 days after first feeling symptoms. Hence, there is a delay of about a week between private information arriving based on individuals feeling sick and public information on excess cases being reported by the national

[21] The interdisciplinary consortium comprises "infas-Institut für angewandte Sozialwissenschaft", "infas 360 GmbH", and the Institute for Hygiene and Public Health of the University of Bonn. The data sets are gathered for research purposes. For more information see https://www.corona-datenplattform.de. Accessed on 3rd March 2021.

Figure 3.4: Timing of First Symptoms and Cases



Notes: This figure shows the distribution of the relative date of first symptoms and positive cases reported publicly for infections that occur at $\Delta t = -8$. The relative date ($\Delta t = 0$) is chosen with respect to the first decile of the positive case information. The log-normal distribution of first symptoms relative to the infection date is taken from Appendix Table 2 in Lauer et al. (2020). The delay between first symptoms and positive test results reported by the national public health agency is based on author's calculations from RKI dataset (see Section 3.3).

public health agency.[22] Combining high-frequency data with knowledge of incubation and testing delays allows us to disentangle changes in behavior caused by private information from changes induced by public information.

Figure 3.4 illustrates the timing of first symptoms and observed cases graphically. Consider an infection cluster that occurs at relative date $\Delta t = -8$. We show the distribution of the incubation time (i.e. time period between infection and first feeling symptoms) based on estimates using the log-normal distribution from Lauer et al. (2020).[23]

To calculate the delay between first symptoms and cases being reported by the national public health agency, we draw on information from the RKI dataset. We define $\Delta t = 0$ as the start of the *observed* outbreak which is specified as the day of the first decile of positive cases.[24] We define the period $\Delta t \in (-7, -1)$ as the private information period, where behavior may change

[22] Recall from Section 3.3.1, that it takes about 2 days until the information on positive test results is passed on from the local to the national health department.

[23] We use the parameter estimates in Appendix Table 2 in Lauer et al. (2020).

[24] The first decile is our approximation of the first day of excess positive test results (see Section 3.4.2).

due to private information about individuals getting sick. Figure 3.4 shows that more than 82% of individuals feels symptoms by $\Delta t = -1$, but only 10% of the cases are reported by the national health agency at that point. The public information period includes the days $\Delta t \in (0, 6)$, when behavior may start to change due to excess cases being reported. Finally, we define a post period as $\Delta t \in (7, 21)$, when we expect that there is no additional information about the outbreak.[25]

3.4.2 Identification of Outbreaks

We identify outbreaks by comparing the observed number of cases in a seven-day period to the number of cases predicted by a simple epidemiological model. The model used to predict the expected incidence $i_{c,s,t}$ (cases per 100,000 inhabitants) in county c, state s, day-of-the-week $dotw_t$, phase p_t, and day t is:

$$i_{c,s,t} = \sum_{j=1}^{7} \beta_j i_{c,s,t-j} + \gamma^i_{c,dotw_t,p_t} + \delta^i_{s,t} + \eta_{c,t}, \tag{3.1}$$

where β_j captures how the incidence depends on the seven-day incidence history in the same county. The model includes county times day-of-the-week times phase fixed effects ($\gamma^i_{c,dotw_t,p_t}$) and state times day fixed effects ($\delta^i_{s,t}$).[26] The fixed effects account for differences in testing regimes across counties, the incidence rate in the surrounding state, and the effect of any state-level policy changes. Importantly, individuals might expect the incidence to increase in a county if the incidence is already high in the surrounding state. Our procedure identifies outbreaks that are unexpected with respect to both the history of cases in the county and the number of cases in the state.

We define an outbreak as a seven-day period where the observed incidence exceeds the expected incidence based on our simple model ($\hat{i}_{c,s,t}$). We consider two different criteria due to sampling variation when the expected number of cases is low. For counties where the expected number of cases is at least five cases per day (henceforth: *Small-Count Threshold*), we take the ratio of the observed incidence to the predicted over seven days. If this ratio is greater than a threshold X, then we identify an outbreak in county c in the seven-day period (see Equation 3.2a, *Ratio Criterion*). As we are using count data, the ratio will be sensitive to sampling variation when the expected cases are low.[27] For this reason, we use a fixed number of cases as the threshold when the

[25] About 80% of observed cases have been reported by $\Delta t = 7$.

[26] County times day-of-the-week FE are allowed to vary by phase of the epidemic in Germany (1st wave, summer, 2nd wave) due to differences in testing and reporting regimes across counties and across phases. As Figure 3.2 shows, there is important day-of-the-week variation even in aggregate data.

[27] Consider the Poisson distribution. If the expected number of cases on a given day is five, then observing at least twice the number of cases has a probability of 1.4%. For this reason, we fix the threshold in terms of number of cases when the expected number of cases is five cases or lower. Five is the number when the Poisson probability of observing twice the number of cases is at least one percent.

expected number of cases is below five per day on average (see Equation 3.2b, *Fixed Criterion*).[28] Let $I_{c,s,t} \equiv i_{c,s,t} * pop_c$ be the number of cases observed in a county, where pop_c is the population divided by 100,000 inhabitants. The decision rule is then

$$\underbrace{\frac{\sum_{j=0}^{6} i_{c,t+j}}{\sum_{j=0}^{6} \hat{i}_{c,t+j}} \geq X \quad \text{if} \quad \underbrace{\sum_{j=0}^{6} \hat{I}_{c,t+j} \geq 35}_{\text{Small-Count Threshold}}}_{\text{Ratio Criterion}} \tag{3.2a}$$

$$\underbrace{\sum_{j=0}^{6} I_{c,t+j} \geq X * 35 \quad \text{if} \quad \underbrace{\sum_{j=0}^{6} \hat{I}_{c,t+j} < 35}_{\text{Small-Count Threshold}}}_{\text{Fixed Criterion}} . \tag{3.2b}$$

In determining the threshold X, we consider different percentiles of the distribution of the ratio $\frac{\sum_{j=0,6} i_{c,t+j}}{\sum_{j=0,6} \hat{i}_{c,t+j}}$ conditional on $\sum_{j=0,6} \hat{I}_{c,t+j} \geq 35$ in the data.

Equations 3.2a and 3.2b identify the seven-day period of an outbreak. We identify the first day of excess cases (i.e. first day of the observed outbreak) as the first day out of the seven when the number of observed cases is above the 90$^{\text{th}}$ percentile of the Poisson distribution (henceforth: *Poisson Threshold*) given the number of expected cases from the model. Finally, if two outbreaks occur within six weeks of each other in the same county, we ignore the second outbreak as it is likely part of a single large outbreak.

3.4.3 Event Study and Difference-in-Differences Designs

To investigate whether and to what extent information about local outbreaks affects mobility patterns, we use both event study and difference-in-differences designs. Recall from Section 3.4.1, that we can define three periods during which individuals may respond to different sources of information relative to first day of excess cases ($\Delta t = 0$): private information period ($\Delta t \in (-7, -1)$), public information period ($\Delta t \in (0, 6)$), and the post period ($\Delta t \in (7, 21)$).

The main estimation equation is

$$\log(\text{trips}_{c,s,t}) = \sum_k \alpha_k T^k_{c,t} + \lambda \cdot \boldsymbol{NPI}_{c,t} + \gamma_{c,dotw_t,p_t} + \delta_{s,t} + \epsilon_{c,s,t}, \tag{3.3}$$

where we define the treatment variable T^k_{ct} differently depending if we are estimating the event study model or the difference-in-differences model. For the event study, the treatment variable is

[28] We also consider small-count thresholds of four and seven in our robustness exercises and it does not affect our results. See Section 3.5.3.

defined as

$$T_{c,t}^{k}(\text{ES}) = \begin{cases} \sum_{j=-\infty}^{-22} D_{c,t-j} & \text{if } k = -22 \\ D_{c,t-k} & \text{if } -21 \leq k \leq 21 \\ \sum_{j=22}^{\infty} D_{c,t-j} & \text{if } k = 22 \end{cases},$$

where $D_{c,t}$ is an indicator that is 1 if it is the first day of an outbreak in county c at time t and 0 otherwise. We normalize $\alpha_{-7}^{\text{ES}} = 0$. In other words, all effects are relative to seven days before the first day of excess cases. For the difference-in-differences model, we aggregate the pre-period, private information period, public information period, and the post-period,

$$T_{c,t}^{k}(\text{DiD}) = \begin{cases} \sum_{j=-\infty}^{-22} D_{c,t-j} & \text{if } k = 0 \\ \sum_{j=-21}^{-8} D_{c,t-j} & \text{if } k = 1 \\ \sum_{j=-7}^{-1} D_{c,t-j} & \text{if } k = 2 \\ \sum_{j=0}^{6} D_{c,t-j} & \text{if } k = 3 \\ \sum_{j=7}^{21} D_{c,t-j} & \text{if } k = 4 \\ \sum_{j=22}^{\infty} D_{c,t-j} & \text{if } k = 5 \end{cases},$$

where we normalize $\alpha_{1}^{\text{DiD}} = 0$. In other words, all effects are relative to the period 8 to 21 days before the first day of excess cases (pre-period).

The parameters of interest are a set of dummies α_j indicating a change in mobility relative to the pre-period.[29] We include county times day-of-the-week times phase fixed effects ($\gamma_{c,dotw_t,p_t}$), which accounts for seasonal and day of the week variation in mobility at the county level. State times day fixed effects ($\delta_{s,t}$) account for any state-level policies and seasonal variation at the state level. Finally, we include a vector of indicators on different non-pharmaceutical interventions ($\boldsymbol{NPI}_{ct}$) at the county-day level, which accounts for policy changes that happen within states, (e.g. NPI's implemented in response to the outbreak).

As outcome variables we consider within and between regional mobility, measured by the logarithm of the daily number of trips. To investigate between-county mobility, we sum the number of inter-county trips that begin or end in the county of interest.

3.4.4 Assessing the Identification Strategy

Identification is achieved within a county over time. The identifying assumption is that – conditional on the set of fixed effects – the timing of a local outbreak is exogenous.

Threats to identification include (1) county-specific time-varying unobservable characteristics (omitted variables) that correlate with outbreaks and mobility and (2) reverse causality, a change in mobility patterns affects local incidence rates.

[29] We bin periods at the endpoints of the event window. Hence, the parameters α_{-22}^{ES} and α_{0}^{DiD} account for all outbreaks occurring 22 or more days before the first day of excess cases. The parameters α_{22}^{ES} and α_{5}^{DiD} account for all outbreaks occurring 22 or more days after the first day of an outbreak (Schmidheiny and Siegloch 2019).

Local events (e.g. carnival, seasonal work in agriculture, private celebrations) may induce increased mobility in a particular county and thus, cause local Covid-19 outbreaks. In the case of publicly known local events, people might perceive an increased risk of infection prior to the event and thus, adjust their mobility in anticipation of an outbreak. Hence, the exogeneity assumption of local outbreaks may be violated.

To tackle the concern of potential outbreak endogeneity, we first exploit that the incubation time of Covid-19 takes on average four to five days (Gandhi et. al 2020b) and the time between first symptoms and positive test results is two to seven days. Individuals expecting a local event (e.g. wedding, religious gathering), that may increase the risk of infection, likely change their behavior already around the time of the event. Hence, there is plausibly a delay of at least seven days between a change in mobility caused by knowledge of a local event and one induced by information from an increase in incidence rates. Therefore, combining high frequency data on a daily basis with disease-specific characteristics (i.e. incubation time) allows us to disentangle changes in mobility due to a local event from ones due to information from the reported incidence rates. Second, to identify local outbreaks we use the predictions derived from the incidence model specified in Equation 3.1 (see Section 3.4.2 for a detailed description of the identification of local outbreaks). This approach allows us to identify *unexpected* changes in the number of locally reported cases.

Finally, it is noteworthy that we control for local NPIs imposed in the aftermath of the outbreaks. For example, child care facilities and schools were closed to curb the outbreak in a meat processing plant in Gütersloh.

3.5 Results

In this section, we present our empirical results on unanticipated outbreaks in Germany and the mobility response to the outbreaks. Section 3.5.1 describes the 259 outbreaks we identify in Germany. Section 3.5.2 presents our main results on the change in mobility due to information about the outbreaks. Finally, in Section 3.5.3, we discuss the sensitivity of our results to different criteria of outbreaks.

3.5.1 Unanticipated Covid-19 Outbreaks in Germany

We identify unanticipated outbreaks by looking for excess incidence rates relative to expected incidence rates based on a simple model that includes lagged incidence, state times day fixed effects, and county times day-of-the-week fixed effects (see Equation 3.1).[30] Appendix Table C.4 presents the estimates of the incidence rate model. The fixed effects alone explain about 74% of the variation in the incidence rates. This is not surprising because there is a strong day-of-the-week variation in the reporting of cases and the state-level incidence rate is a good predictor

[30] See Section 3.4.2 for more details.

Table 3.1: Summary of Outbreaks Identified

	No. of Outbreaks		
	Total	**Identified by Eq. 3.2a**	**Identified by Eq. 3.2b**
1st wave	105	77	28
Summer	80	36	44
2nd wave	74	48	26
Total	259	161	98

Notes: The table summarizes the number of outbreaks per phase of the pandemic. Column 1 presents the total number of outbreaks per phase. Columns 2 and 3 display the number of outbreaks identified by Equation 3.2a and Equation 3.2b, respectively.

of the incidence rate in the counties. Including state times day fixed effects is important as an increase in incidence rates that can be predicted by the incidence rate at the state level is unlikely to be unanticipated. The lagged incidence rate in a county is also an important predictor of future incidence rates as can be seen in the second column of Appendix Table C.4.[31] We use the incidence rate model to predict incidence rates in each county on each day.

We define an outbreak as an excess in the observed incidence rates compared to the expected incidence rates from our model using Equations 3.2a and 3.2b. Appendix Figure C.1 shows the distribution of the ratio of observed to expected incidence rates in the data. Our preferred threshold is the 98th percentile (i.e. $X = 1.54$), which identifies 259 outbreaks between February and November. In other words, we identify outbreaks in counties that have more than a 54% excess in a seven-day period, or more than 54 cases when the expected number of cases is less than 35 in a seven-day period. The Ratio Criterion (i.e. Equation 3.2a) identifies 161 outbreaks and 98 are identified by the Fixed Criterion (i.e. Equation 3.2b). Table 3.1 shows the number of outbreaks we identify in each phase and by each criterion.

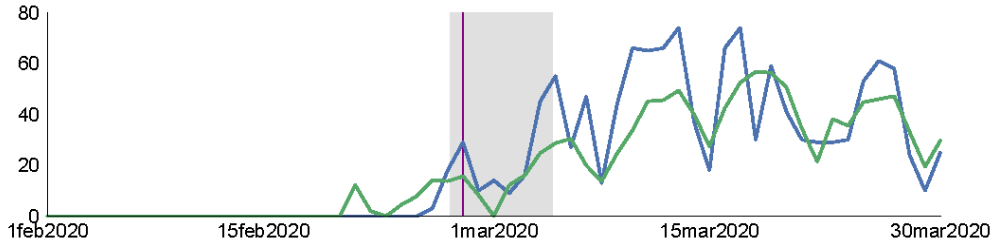
To better understand how unanticipated outbreaks are identified, we present the time-series figures for four well-known examples of outbreaks in Germany. Figure 3.5 displays the daily observed cases (blue line), the daily expected cases (green line), the earliest seven-day period that an excess is observed (grey shading), and the first day of excess cases determined by the Poisson Threshold (purple vertical line).[32]

[31] With the lagged incidence rates, the model explains about 78% of the variation in incidence rates.

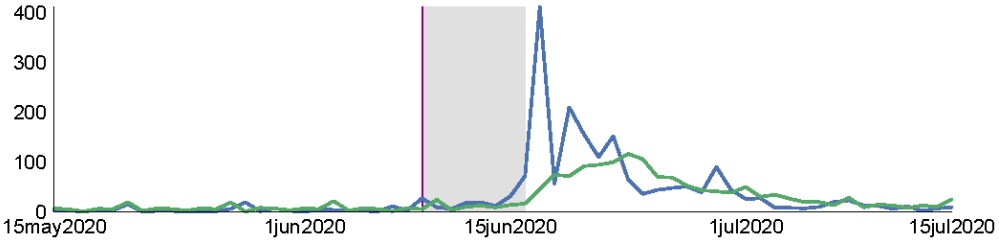
[32] Perhaps the dates chosen for Gütersloh are surprising, but it is important to note that this was a particularly large outbreak and the first day of excess cases are larger than they appear due to the differences in scale.

Figure 3.5: Examples of Well-Known Outbreaks in Germany

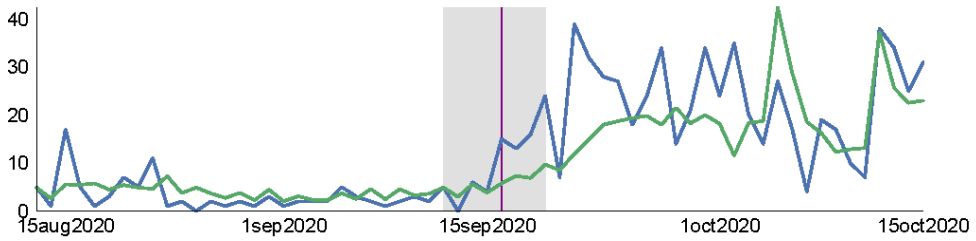
(a) March Outbreak in Heinsberg



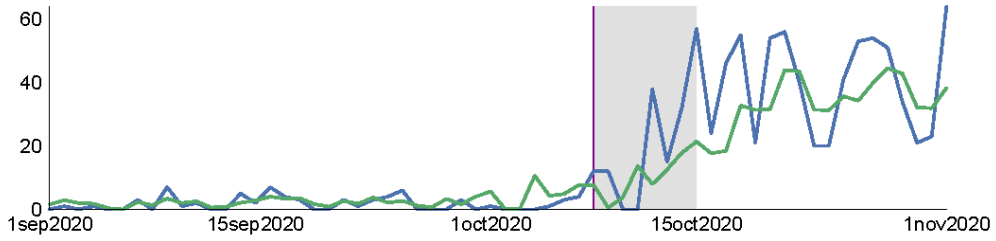
(b) June Outbreak in Gütersloh



(c) September Outbreak in Hamm



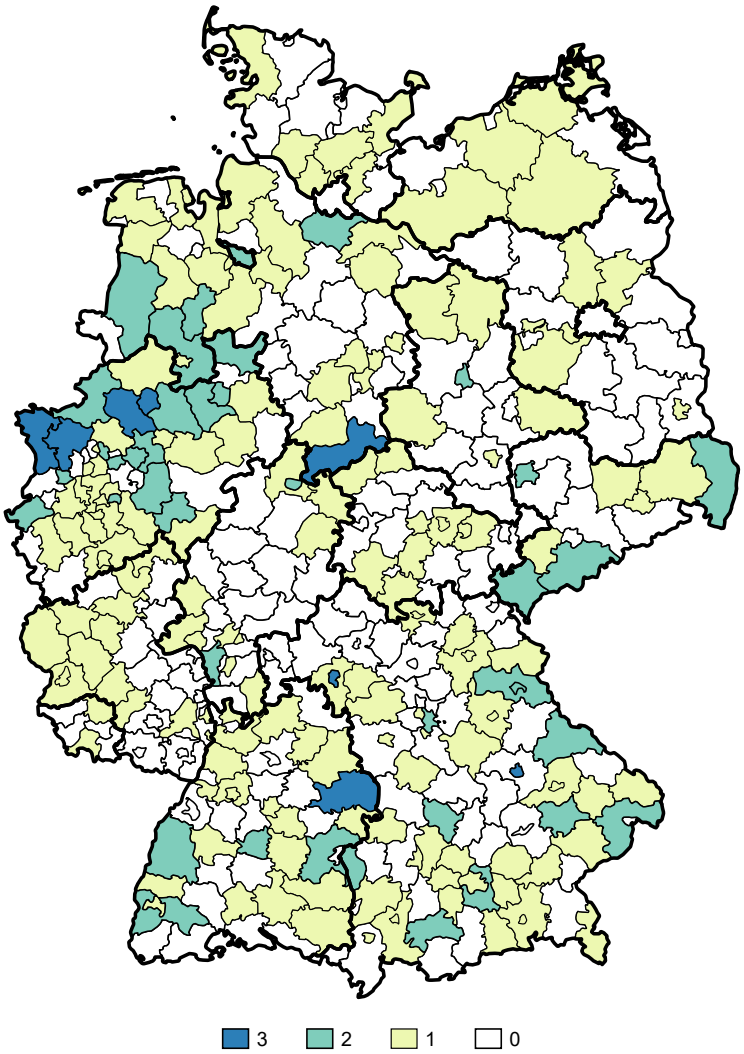
(d) October Outbreak in Berchtesgaden



Notes: These figures show four examples of well-known outbreaks that are also identified by our procedure. The blue time-series line shows the cases reported to the national public health institute on each day and the green time-series line shows the cases predicted by our model. The grey shaded area represents the seven-day period of the beginning of the outbreak and the purple vertical line shows the first date of the outbreak identified by our procedure. All four outbreaks are identified by the Ratio Criterion Equation 3.2a.

Figure 3.6 shows the spatial distribution of outbreaks across Germany, while Appendix Figure C.2 presents the temporal distribution of outbreaks. The outbreaks are distributed relatively evenly across the counties of Germany and across time.

Figure 3.6: Counties with Outbreaks



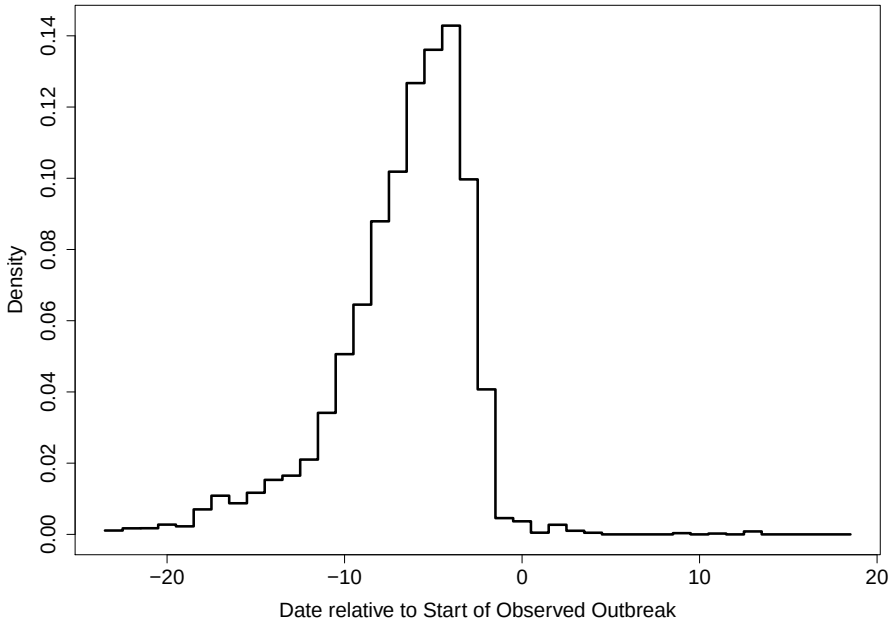
Notes: This figure shows how the outbreaks are distributed across Germany. Some counties have more than one outbreak in the February to November period.

3.5.2 Aggregate Mobility Responses to Outbreaks

We now study the mobility response to the outbreaks identified in Section 3.5.1. Recall that due to the progression of the disease and delays in getting tested, we can interpret changes in mobility on different days as being driven by different sources of information. Let the relative date be represented by Δt, where $\Delta t = 0$ is the first day of excess cases. Individuals may reduce mobility due to information about a risky event (e.g. a wedding or festival) in the period $\Delta t \in (-14, -7)$, due to private information about people falling ill in the period $\Delta t \in (-7, -1)$, and public information on excess cases in the period $\Delta t \in (0, 6)$.[33]

We construct the distribution of the relative date of first symptoms for cases that were reported on the first day of the observed outbreak ($\Delta t = 0$). Figure 3.7 shows the distribution of first reported symptoms for our set of outbreaks. Most individuals with positive cases start having symptoms in the seven days before the start of the observed outbreak. This distribution further motivates our definition of the private information period. In other words, the private information period is the period when infected individuals start to have symptoms, but when there is no public information about the outbreak yet.

Figure 3.7: Distribution of First Symptoms

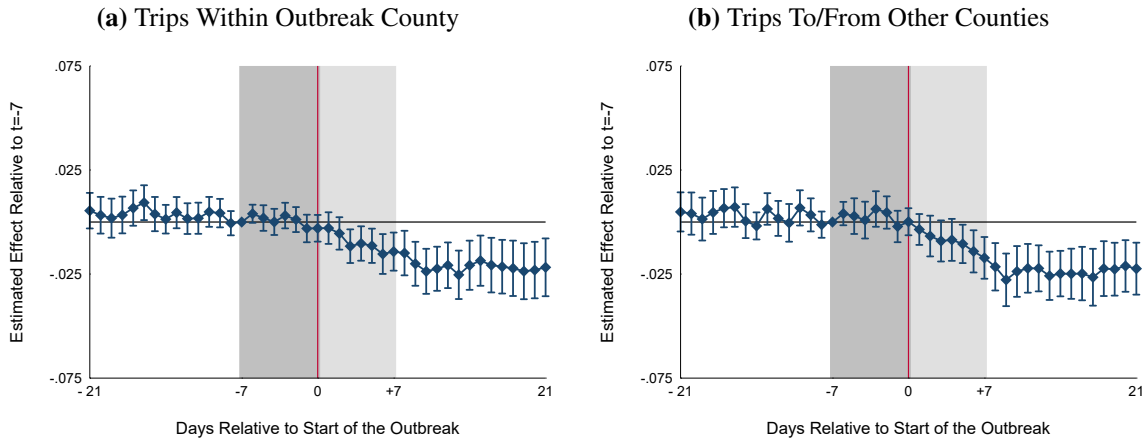


Notes: This figure shows the distribution of the relative date of first symptoms for the cases that were reported on the first day of the observed outbreaks ($\Delta t = 0$). Note that the relative date includes a two-day delay between individuals getting tested and the cases being publicly reported by the national public health agency (See Section 3.3).

[33] If a risky event is large enough, we may observe an increase in mobility at the time of the event.

Our main results use an event study design to distinguish the effects of different sources of information on mobility. Figure 3.8 presents the event study on the effect of an outbreak on the number of within-county and between-county trips. We normalize effects relative to $\Delta t = -7$ as most of the infected individuals in the outbreak do not have any symptoms at this point. We shade the private information period and the public information periods. We don't find any pre-trend in the data that may be due to knowledge of a risky event, nor do we see a significant change in mobility during the private information period. It is not until public information of the outbreak becomes available that mobility begins to decline. The decrease in mobility plateaus after about seven days have passed since the first day of excess cases. The effects look similar for within-county travel and between-county travel, where the effect is a bit delayed and larger for between-county trips.

Figure 3.8: Event Study of County-Level Mobility



Notes: Event studies of the log of number of trips taken within a county with an outbreak (panel a) and between the outbreak county and other counties (panel b). The model controls for day times state fixed effects, county times day-of-the-week fixed effects and NPI indicators at the county-day level (see Section 3.4.3 for more information). Time $\Delta t = 0$ is the first day of excess cases in the outbreak (see Section 3.4.2). We normalize at $\Delta t = -7$ as we might expected mobility to change in the seven days before the observed outbreak due to private information. Hence, the left (dark grey) shaded area represents the period where we would expect changes in mobility due to private information. The right (light grey) shaded area represents the seven-day period when the excess of cases is reported and information about the outbreak is revealed by the public health authorities.

We present the results for the difference-in-differences specification in Table 3.2, where we pool the effects for the different periods. In order to understand if local NPIs are driving our results, we present results with and without controlling for local NPIs. We find that controlling for NPIs makes little difference in our estimates. As in the event study, we do not find any effect of the outbreak in the private information period. The average effect in the public information period is between a third (between-county) to half (within-county) of the full-effect in the post period. Within-county mobility decreases by about 2.2% and between-county mobility decreases a bit more, by about 2.4%.

Table 3.2: Difference-in-Difference Estimates

	Within-County		**Between-County**	
Private Information ($\Delta t \in (-7, -1)$)	-0.0007	-0.0007	0.0012	0.0012
	(0.003)	(0.003)	(0.003)	(0.003)
Public Information ($\Delta t \in (0, 6)$)	-0.0102*	-0.0100*	-0.0085*	-0.0084*
	(0.004)	(0.004)	(0.004)	(0.004)
Post Period ($\Delta t > 6$)	-0.0221***	-0.0217***	-0.0242***	-0.0237***
	(0.005)	(0.005)	(0.005)	(0.005)
Observations	140350	140350	140350	140350
Adj.R2.	0.994	0.994	0.992	0.992
Local NPI FE	No	Yes	No	Yes

Notes: This table presents the difference-in-difference estimates of the event studies shown in Figure 3.8. The dependent variable is the log of number of trips taken within a county with an outbreak and between the outbreak county and other counties. The model controls for day times state fixed effects, county times day-of-the-week fixed effects and NPI indicators at the county-day level (see Section 3.4.3 for more information). Time $\Delta t = 0$ is the first day of excess cases in the outbreak (see Section 3.4.2). We normalize $\Delta t \in (-21, -8)$ as we might expected mobility to change in the seven days before the observed outbreak due to private information of infected individuals falling ill. Standard errors are clustered at the county level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Heterogeneity Analysis There are a number of reasons why the mobility response to an outbreak may vary across counties and across time. Counties may differ in the average costs of reducing mobility (e.g. more factory workers than programmers), mobility may expose populations in certain counties to greater risk (e.g. higher usage of subways than cars), or the average beliefs about the risk of the virus may vary across populations. In the following, we provide suggestive evidence on each of the three channels.

Table 3.3 shows how effects vary by continuous proxies: share of college educated, hotel beds per capita, public transportation trips per capita, and share of voters who voted for the party AfD in the last state election. The continuous variables are standardized with mean zero and standard deviation one to ease interpretation and comparability of the results. Table 3.4 shows how effects vary by discrete proxies: stage of the pandemic and level of urbanization of the county.

The costs of adjusting mobility may vary if a large share of workers can work from home, if a large share of the mobility is discretionary, or if individuals are tired of restrictions on mobility. We investigate each of these in turn. If the likelihood of being able to work from home is larger for college-educated workers, then we may expect the effects to be larger in counties where the fraction of individuals with tertiary degrees is higher. Interestingly, we do not find that the effect on within-county mobility changes depends on education. The effect for between-county mobility is stronger in counties with a higher fraction of individuals with a tertiary degree. The effect is fifty percent larger in a county with a one standard deviation higher share of tertiary degrees.

Table 3.3: Heterogeneity Analysis: Part I

	Tertiary School Degree		**Public Transport**		**AFD Vote Share**		**Hotel Beds per Capita**	
	Within	**Between**	**Within**	**Between**	**Within**	**Between**	**Within**	**Between**
Private Information ($\Delta t \in (-7,-1)$)	-0.0008	0.0012	-0.0009	0.0012	-0.0008	0.0014	-0.0014	0.0008
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Public Information ($\Delta t \in (0,6)$)	-0.0099*	-0.0080	-0.0102*	-0.0082	-0.0098*	-0.0079	-0.0110**	-0.0090*
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Post Period ($\Delta t > 6$)	-0.0214***	-0.0230***	-0.0219***	-0.0233***	-0.0205***	-0.0225***	-0.0229***	-0.0245***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Private Information × Tertiary School Degree	0.0016	0.0002						
	(0.002)	(0.002)						
Public Information × Tertiary School Degree	-0.0015	-0.0052						
	(0.003)	(0.003)						
Post Period × Tertiary School Degree	-0.0037	-0.0113*						
	(0.004)	(0.005)						
Private Information × Public Transport			0.0024	0.0003				
			(0.002)	(0.002)				
Public Information × Public Transport			0.0018	-0.0016				
			(0.004)	(0.004)				
Post Period × Public Transport			0.0025	-0.0050				
			(0.005)	(0.006)				
Private Information × AfD Vote Share					-0.0003	0.0008		
					(0.003)	(0.003)		
Public Information × AfD Vote Share					0.0011	0.0019		
					(0.005)	(0.004)		
Post Period × AfD Vote Share					0.0048	0.0048		
					(0.007)	(0.006)		
Private Information × Hotel Beds per Capita							-0.0028	0.0011
							(0.005)	(0.004)
Public Information × Hotel Beds per Capita							-0.0155	-0.0094
							(0.008)	(0.006)
Post Period × Hotel Beds per Capita							-0.0368***	-0.0257***
							(0.009)	(0.007)
Observations	140350	140350	140350	140350	140350	140350	140350	140350
Adj.R2.	0.994	0.992	0.994	0.992	0.994	0.992	0.994	0.992
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table studies heterogeneity in response to local outbreaks along different dimensions. Column 1 analyzes how people's response in mobility vary by the share of individuals with a tertiary school degree in a county. Column 2 investigates differences in response behavior on the daily number of trips by public transportation per capita inhabitants, Column 3 on the vote share for AfD in the last state elections. Column 4 studies differences in response to the number of hotel beds per capita in a county. All four variables are standardized to have zero mean and standard deviation of 1.0. Standard errors are clustered at the county level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Trips taken due to tourism or business travel may be easier to reduce compared to other kinds of trips, like commuting to work. We study how the effects vary by the importance of the travel industry in a county, proxied by hotel beds per capita. Indeed, we find that the effect of an outbreak on mobility is much larger in counties where the travel industry is important. Counties with one standard deviation more hotel beds per capita have mobility effects that are two and a half times larger for within-county travel and twice as large for between-county travel.

"Pandemic fatigue" could increase the perceived cost of decreasing mobility as more time has passed since the beginning of the pandemic. To study this, we estimate the effects separately by phase. On one hand, we do find that the response to an outbreak is about half as strong in the summer period compared to the first wave. On the other hand, we find that the mobility response is largest during the second wave. We interpret these results as showing that "pandemic fatigue" was not important in the mobility response to unanticipated outbreaks. Alternatively, the response may have been weaker during the summer as people were less attentive while the national incidence rate was low.[34]

If the average risk of a trip is higher, we may expect the mobility response to be larger. The average risk may be larger in counties with higher usage of public transportation or in counties with a higher population density. Interestingly, as seen in Table 3.3, we do not find that the mobility response depends on the number of public transportation trips per capita. In Table 3.4, we separately estimate the effects for four different categories of population density. We do find that the effects are larger in "Large Cities" and do not find a significant effect in "Cities" and "Rural Areas with Small Cities".[35] Paradoxically, we find equally large effects in the "Rural" areas—least dense category—as in "Large Cities". Many of the rural counties in our sample are also places with high levels of tourism. If we control for the interaction with hotel beds per capita, we find that the mobility response in rural areas decreases by half, while it increases in cities. We interpret this as evidence that the increased effect in rural areas is mostly due to the lower costs of discretionary travel, while the effect in cities is consistent with the increased risk of a higher population density.

Finally, beliefs about the risk of Covid-19 in a county may lead to stronger or weaker mobility responses to an outbreak. A number of studies in the US (e.g. Allcott et al. 2020; Andersen 2020; Barrios and Hochberg 2020; Grossman et al. 2020; Painter and Qiu 2021) have shown how counties with higher Republican support were less likely to change their behavior in response to

[34]For example, a simple query in google trends for "rki corona" or "rki corona fallzahlen" shows that the number of searches in Germany about the coronavirus were higher during the first and second wave compared to the summer period, reaching four to five times higher at the peaks in the middle of our first and third phases.

[35]We refer to the classification into different types of urbanization proposed by BBSR (2019). "Large Cities" denote *Kreisfreie Großstädte* with more than 100,000 inhabitants. "Cities" present *Städtische Kreise* with a population density of at least 150 inhabitants/km^2 and at least 50% of the population living in a city. "Rural Areas with Small Cities" include (i) counties with a population density less than 150 inhabitants/km^2 and at least 50% of the population living in a city and (ii) counties with a population density of at least 100 inhabitants/km^2 and less 50% of the population living in a city. "Rural Areas" denote counties with a population density smaller than 100 inhabitants/km^2 and less than 50% of the population living in a city.

Table 3.4: Heterogeneity Analysis: Part II

	Phase of Pandemic		**Type of Urbanization**			
	Within	**Between**	**Within**	**Between**	**Within**	**Between**
Private Information × 1st Wave	-0.005	-0.000				
	(0.004)	(0.004)				
Private Information × Summer	-0.001	0.002				
	(0.003)	(0.004)				
Private Information × 2nd Wave	0.006	0.002				
	(0.008)	(0.005)				
Public Information × 1st Wave	-0.019**	-0.012				
	(0.007)	(0.007)				
Public Information × Summer	-0.002	-0.001				
	(0.004)	(0.005)				
Public Information × 2nd Wave	-0.007	-0.012				
	(0.011)	(0.008)				
Post Period × 1st Wave	-0.024**	-0.028**				
	(0.008)	(0.010)				
Post Period × Summer	-0.012**	-0.012*				
	(0.005)	(0.006)				
Post Period × 2nd Wave	-0.029*	-0.032**				
	(0.014)	(0.010)				
Private Information × Large Cities			-0.001	-0.001	-0.002	-0.001
			(0.005)	(0.005)	(0.005)	(0.005)
Private Information × Cities			-0.003	-0.003	-0.003	-0.003
			(0.003)	(0.003)	(0.003)	(0.003)
Private Information × Rural w/ Small Cities			0.010	0.003	0.009	0.002
			(0.008)	(0.008)	(0.007)	(0.007)
Private Information × Rural			-0.008	0.009	-0.006	0.010
			(0.009)	(0.007)	(0.009)	(0.006)
Public Information × Large Cities			-0.018*	-0.019*	-0.025**	-0.024*
			(0.008)	(0.009)	(0.009)	(0.010)
Public Information × Cities			-0.005	-0.005	-0.009	-0.008
			(0.005)	(0.005)	(0.006)	(0.005)
Public Information × Rural w/ Small Cities			0.002	-0.001	0.003	-0.000
			(0.007)	(0.010)	(0.008)	(0.009)
Public Information × Rural			-0.022	-0.007	-0.011	0.000
			(0.014)	(0.010)	(0.011)	(0.009)
Post Period × Large Cities			-0.037***	-0.044***	-0.050***	-0.053***
			(0.011)	(0.013)	(0.011)	(0.013)
Post Period × Cities			-0.003	-0.007	-0.013	-0.015*
			(0.006)	(0.006)	(0.007)	(0.006)
Post Period × Rural w/ Small Cities			-0.015	-0.016	-0.011	-0.013
			(0.008)	(0.011)	(0.008)	(0.011)
Post Period × Rural			-0.047**	-0.039**	-0.023*	-0.021*
			(0.017)	(0.013)	(0.011)	(0.010)
Observations	140350	140350	140350	140350	140350	140350
Adj.R2.	0.994	0.992	0.994	0.992	0.994	0.992
FE	Yes	Yes	Yes	Yes	Yes	Yes
Period × Hotel Beds Interactions	No	No	No	No	Yes	Yes

Notes: The table studies heterogeneity in response to local outbreaks during different phases of the pandemic (Columns 2 and 3) and types of urbanization in counties (Columns 4 and 5). In addition, the model with urbanization is estimated additionally controlling for interactions with the standardized "hotel beds per capita" variable from Table 3.3 (Columns 6 and 7). Standard errors are clustered at the county level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Covid-19. Similarly in Germany, the AfD party was critical of the government handling of Covid-19 and one may conjecture that counties with high AfD vote shares would, likewise, respond less to an outbreak. Unlike in the US, we do not find any variation in the effect in counties where the AfD vote share is larger.

3.5.3 Sensitivity Analysis

We estimate a number of alternative specifications to understand the robustness of our results. We vary the X threshold (97th, 98th, and 99th percentiles), the small-count threshold (four, five, and seven expected cases per day), and the Poisson threshold (p-values of 0.1, 0.01, and 0.001) for identifying outbreaks (see Section 3.4.2). Appendix Table C.5 shows the number of outbreaks identified using the different thresholds and for the different criteria, where the number of outbreaks identified varies from 140 to 353. We estimate the difference-in-differences model using each of these sets of outbreaks in Appendix Table C.6. The estimated effect is smaller when lowering the small-sample threshold, this is expected as we are introducing some outbreaks that are identified from the small-count noise. The estimated effects when raising the small-count threshold are nearly the same as our preferred specification. Lowering the X threshold leads to slightly lower estimates for the effect, but the difference with preferred specification is negligible. Increasing the X threshold on the other hand, leads to a substantial increase in the estimated effects. The higher X threshold only includes larger outbreaks, and so it may not be surprising that the responses are larger in this case. Finally, the estimates do not change much when varying the Poisson threshold, but this is expected as the value of Poisson threshold only determines which day the outbreak occurs. A tighter Poisson threshold leads to a one to two day delay in the identification of the first day of an outbreak, which has a small effect on the event study results. In addition, we find that behavioral responses to local outbreaks identified by the Ratio Criterion (Equation 3.2a) and by the Fixed Criterion (Equation 3.2b) are nearly the same (see Appendix Table C.8).

We also consider outbreaks that are mentioned in the daily reports of the national public health agency (RKI) during the summer (see Appendix Table C.7).[36] If we use the RKI-defined outbreaks, we find similar between-county effects and slightly smaller within-county effects compared to our main specification (1.2%). If we consider the outbreaks that are identified both in the RKI and in our sample, then we find larger effects that are similar to ones in the first and third phase of the pandemic (compare with results in Table 3.4). The main conclusion is that we do not find substantially different effects if we consider the outbreaks identified by the RKI as opposed to

[36] See Robert Koch Institute (2021a). The RKI listed outbreaks only during the summer period in their daily reports, presumably when they could use contract tracing to identify transmission clusters. We could not find a precise rule for how the RKI defined outbreaks during the summer period. They mention outbreaks in counties with incidence rates relatively high to the rest of Germany. However as the baseline incidence rate varies in Germany, the cutoff appears to also change.

the outbreaks identified using the procedure in Section 3.4.2. Our procedure though can be applied to the entire sample period.

Finally, we perform a placebo study, where, for each outbreak we identify, we randomly choose a county and then repeat the analysis. Appendix Figure C.4 shows the event study analysis using placebo outbreaks. We do not find any effect in the placebo outbreak sample.

3.6 Conclusion

This paper underscores the importance of public information as a policy tool for mitigating public health risks. We study this in the context of Covid-19, where behavioral responses have been important to contain the spread of the virus and, hence, the costs of the pandemic. To identify behavioral responses induced by public information versus other sources, we combine high-frequency data with facts about the incubation period of Covid-19 and reporting time in Germany. We first develop a simple epidemiological model that allows us to identify unexpected local outbreaks by comparing the observed number of cases to the expected number. Based on the model, we find 259 outbreaks at the county level that are distributed relatively evenly across the counties in Germany and across time. Using an event study design on local unexpected outbreaks, we find that mobility significantly decreases by about 2 to 3% in response to public information about the outbreak, while private knowledge about people falling sick does not appear to cause a change in behavior. There are important heterogeneities in the behavioral responses, where responses are stronger in counties with high population density, with more hotels per capita, and with a higher share of college educated. These findings are consistent with behavioral changes depending on the relative risk and costs of changing mobility.

The study provides evidence that people respond to information on local outbreaks to protect themselves towards an increased risk of infection. Having instruments to control the spread of the pandemic beyond enforced social distancing measures will be particularly important for policy makers in times that people exhibit "pandemic fatigue" while at the same time herd immunity by the vaccination strategy has not been achieved. More generally, this paper highlights the importance of providing detailed real-time information to improve population health.

Appendix C

C1 Descriptive Statistics

Table C.1: Summary Statistics Part I: Infections and Trips

				New Infections, Mean (S.D.)		Trips per capita, Mean (S.D.)		
State	# of Counties	Phase	Obs	Cases	Incidence	Outgoing	Incoming	Internal
All states	401	before	20,050			0.68 (0.29)	0.68 (0.29)	1.46 (0.41)
		1st wave	28,070	5.79 (13.25)	2.83 (5.18)	0.50 (0.26)	0.50 (0.26)	1.24 (0.39)
		Summer	61,353	2.15 (6.76)	0.89 (1.94)	0.68 (0.28)	0.68 (0.28)	1.54 (0.48)
		2nd wave	36,892	39.57 (73.07)	18.21 (17.63)	0.61 (0.28)	0.61 (0.28)	1.46 (0.45)
Baden-Württemberg	44	before	2,200			0.62 (0.26)	0.62 (0.26)	1.55 (0.40)
		1st wave	3,080	10.51 (15.05)	4.47 (6.27)	0.44 (0.23)	0.44 (0.23)	1.29 (0.37)
		Summer	6,732	2.64 (4.43)	1.00 (1.47)	0.59 (0.23)	0.59 (0.23)	1.52 (0.38)
		2nd wave	4,048	47.47 (45.58)	18.66 (14.32)	0.54 (0.25)	0.54 (0.24)	1.52 (0.43)
Bayern	96	before	4,800			0.91 (0.33)	0.90 (0.33)	1.42 (0.45)
		1st wave	6,720	6.40 (14.60)	4.71 (7.87)	0.64 (0.32)	0.63 (0.32)	1.16 (0.41)
		Summer	14,688	1.76 (6.15)	1.12 (2.59)	0.89 (0.31)	0.89 (0.30)	1.50 (0.53)
		2nd wave	8,832	29.59 (50.69)	20.95 (17.92)	0.81 (0.34)	0.81 (0.33)	1.43 (0.45)
Berlin	1	before	50			0.22 (0.04)	0.22 (0.03)	1.78 (0.29)
		1st wave	70	84.97 (76.02)	2.33 (2.09)	0.16 (0.05)	0.16 (0.05)	1.35 (0.38)
		Summer	153	58.86 (52.57)	1.61 (1.44)	0.22 (0.04)	0.21 (0.03)	1.56 (0.28)
		2nd wave	92	903.95 (501.08)	24.80 (13.75)	0.19 (0.04)	0.18 (0.04)	1.46 (0.30)
Brandenburg	18	before	900			0.74 (0.26)	0.72 (0.24)	1.40 (0.48)
		1st wave	1,260	2.33 (5.19)	1.43 (2.89)	0.59 (0.24)	0.59 (0.24)	1.30 (0.47)
		Summer	2,754	0.52 (1.32)	0.34 (0.86)	0.86 (0.26)	0.86 (0.26)	1.75 (0.64)
		2nd wave	1,656	23.51 (26.81)	17.48 (20.80)	0.73 (0.25)	0.73 (0.25)	1.72 (0.65)
Bremen	2	before	100			0.57 (0.14)	0.56 (0.13)	1.52 (0.29)
		1st wave	140	6.46 (12.65)	1.40 (2.28)	0.43 (0.17)	0.42 (0.16)	1.21 (0.32)
		Summer	306	5.11 (7.63)	1.41 (2.06)	0.52 (0.14)	0.52 (0.14)	1.40 (0.26)
		2nd wave	184	60.88 (67.49)	14.79 (10.71)	0.49 (0.17)	0.49 (0.16)	1.35 (0.30)
Hamburg	1	before	50			0.32 (0.06)	0.32 (0.05)	1.73 (0.31)
		1st wave	70	69.36 (60.72)	3.77 (3.30)	0.22 (0.09)	0.22 (0.08)	1.28 (0.33)
		Summer	153	20.89 (23.52)	1.13 (1.28)	0.29 (0.06)	0.28 (0.05)	1.51 (0.29)
		2nd wave	92	319.12 (158.07)	17.33 (8.59)	0.25 (0.06)	0.25 (0.06)	1.42 (0.32)
Hessen	26	before	1,300			0.60 (0.22)	0.61 (0.22)	1.40 (0.37)
		1st wave	1,820	4.61 (6.89)	1.97 (2.91)	0.44 (0.20)	0.45 (0.20)	1.18 (0.33)
		Summer	3,978	2.72 (5.04)	1.03 (1.54)	0.58 (0.19)	0.58 (0.19)	1.44 (0.36)

		2nd wave	2,392	49.97 (52.73)	19.83 (15.60)	0.52 (0.19)	0.52 (0.19)	1.37 (0.38)
Mecklenburg-Vorpommern	8	before	400			0.50 (0.22)	0.49 (0.22)	1.72 (0.39)
		1st wave	560	1.24 (2.13)	0.67 (1.28)	0.41 (0.21)	0.40 (0.21)	1.53 (0.41)
		Summer	1,224	0.43 (1.21)	0.22 (0.68)	0.69 (0.24)	0.68 (0.25)	2.50 (0.91)
		2nd wave	736	15.04 (17.74)	7.49 (8.72)	0.53 (0.24)	0.52 (0.25)	2.02 (0.68)
Niedersachsen	45	before	2,250			0.60 (0.22)	0.60 (0.22)	1.47 (0.39)
		1st wave	3,150	3.28 (7.55)	1.58 (2.80)	0.46 (0.21)	0.46 (0.21)	1.25 (0.37)
		Summer	6,885	1.49 (4.08)	0.72 (1.57)	0.60 (0.21)	0.60 (0.21)	1.52 (0.41)
		2nd wave	4,140	21.52 (34.56)	11.40 (11.57)	0.54 (0.22)	0.54 (0.22)	1.46 (0.42)
Nordrhein-Westfalen	53	before	2,650			0.53 (0.16)	0.53 (0.15)	1.61 (0.34)
		1st wave	3,710	8.92 (13.33)	2.65 (3.58)	0.38 (0.16)	0.39 (0.16)	1.32 (0.33)
		Summer	8,109	4.67 (9.01)	1.35 (2.47)	0.50 (0.14)	0.50 (0.14)	1.56 (0.32)
		2nd wave	4,876	66.93 (59.47)	19.86 (13.85)	0.44 (0.14)	0.44 (0.14)	1.43 (0.35)
Rheinland-Pfalz	36	before	1,800			0.79 (0.27)	0.79 (0.26)	1.26 (0.33)
		1st wave	2,520	2.41 (3.83)	2.07 (3.12)	0.59 (0.25)	0.59 (0.25)	1.07 (0.28)
		Summer	5,508	0.87 (1.96)	0.71 (1.46)	0.80 (0.24)	0.80 (0.23)	1.32 (0.34)
		2nd wave	3,312	18.87 (19.88)	16.38 (14.65)	0.70 (0.24)	0.70 (0.24)	1.22 (0.35)
Saarland	6	before	300			0.61 (0.17)	0.61 (0.17)	1.47 (0.31)
		1st wave	420	6.05 (11.36)	3.24 (4.52)	0.43 (0.18)	0.43 (0.18)	1.16 (0.30)
		Summer	918	0.90 (1.62)	0.54 (0.95)	0.59 (0.17)	0.59 (0.17)	1.44 (0.30)
		2nd wave	552	30.23 (32.74)	17.53 (12.92)	0.54 (0.18)	0.54 (0.18)	1.39 (0.31)
Sachsen	13	before	650			0.49 (0.18)	0.49 (0.18)	1.54 (0.37)
		1st wave	910	5.20 (7.78)	1.66 (2.43)	0.39 (0.19)	0.39 (0.19)	1.36 (0.38)
		Summer	1,989	1.39 (2.94)	0.45 (1.02)	0.53 (0.19)	0.53 (0.20)	1.66 (0.39)
		2nd wave	1,196	108.95 (103.10)	36.29 (34.21)	0.46 (0.20)	0.46 (0.20)	1.59 (0.44)
Sachsen-Anhalt	14	before	700			0.53 (0.21)	0.54 (0.21)	1.65 (0.46)
		1st wave	980	1.60 (3.15)	0.97 (1.77)	0.45 (0.20)	0.45 (0.20)	1.51 (0.43)
		Summer	2,142	0.53 (1.45)	0.31 (0.78)	0.58 (0.20)	0.59 (0.21)	1.79 (0.46)
		2nd wave	1,288	22.01 (25.58)	13.86 (15.46)	0.53 (0.22)	0.54 (0.22)	1.73 (0.49)
Schleswig-Holstein	15	before	750			0.59 (0.19)	0.58 (0.18)	1.38 (0.30)
		1st wave	1,050	2.58 (4.77)	1.21 (1.99)	0.43 (0.19)	0.43 (0.19)	1.17 (0.31)
		Summer	2,295	0.93 (2.16)	0.49 (1.33)	0.61 (0.19)	0.60 (0.18)	1.55 (0.50)
		2nd wave	1,380	14.71 (17.52)	7.04 (7.00)	0.53 (0.19)	0.52 (0.19)	1.38 (0.38)
Thüringen	23	before	1,150			0.60 (0.21)	0.60 (0.21)	1.35 (0.43)
		1st wave	1,610	1.45 (2.87)	1.52 (2.92)	0.48 (0.21)	0.49 (0.21)	1.22 (0.39)
		Summer	3,519	0.51 (1.38)	0.58 (1.69)	0.63 (0.21)	0.63 (0.22)	1.43 (0.42)
		2nd wave	2,116	18.66 (22.98)	20.67 (25.26)	0.57 (0.23)	0.58 (0.23)	1.41 (0.43)

Notes: The table reports the mean and standard deviation of the set of variables included in the analysis. The data is pooled over the period of observation and counties (N=401). Incidence is defined as the daily number of cases per 100,000 inhabitants.

Table C.2: Summary Statistics Part II: Additional Variables

	Mean	**S.D.**
No. of Beds in Hotels per Capita	4.26	5.24
Public Transport per Capita	0.29	0.09
Share of Tertiary Degree	11.85	3.23
Vote Share AfD	12.47	6.15
Large Cities	0.17	0.37
Cities	0.33	0.47
Rural Areas w/ Small Cities	0.25	0.43
Rural Areas	0.25	0.44

Notes: This table reports mean and standard deviation of variables used in the heterogeneity analysis. The data are cross-sectional and provided at the county level.

Table C.3: Summary Statistics Part III: Non-Pharmaceutical Interventions

	1st wave		**summer**		**2nd wave**	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
Child Care, School and Work						
Day Care	0.35	0.48	0.66	0.47	0.86	0.35
Primary Schools	0.42	0.49	0.89	0.31	0.91	0.29
Secondary Schools	0.35	0.48	0.91	0.28	0.92	0.27
Workplace	0.07	0.26	0.24	0.42	0.13	0.34
Travel						
Travel Restrictions Domestic	0.12	0.33	0.13	0.34	0.00	0.00
Travel Restrictions Foreign	0.04	0.20	0.00	0.00	0.00	0.00
Public Transport	0.00	0.00	0.00	0.00	0.00	0.00
General NPIs						
Mask Mandate	0.09	0.29	0.99	0.10	1.00	0.00
Social Distancing	0.45	0.50	0.87	0.34	0.95	0.22
Exit Restrictions	0.15	0.35	0.05	0.21	0.09	0.29
Contacts - Private Space	0.19	0.39	0.57	0.49	0.76	0.42
Contacts - Public Space	0.41	0.49	0.99	0.07	0.98	0.13
Testing	0.02	0.14	0.12	0.33	0.08	0.27
Service, Hotels, Restaurants						
Services	0.52	0.50	0.99	0.07	0.99	0.07
Hotels	0.52	0.50	0.96	0.20	0.93	0.25
Restaurants	0.57	0.50	1.00	0.00	0.97	0.18
Retails	0.49	0.50	0.91	0.28	0.97	0.16
Events, Sports and Culture						
Events Indoor	0.56	0.50	0.99	0.08	0.98	0.13
Events Outdoor	0.57	0.50	0.97	0.17	0.99	0.11
Night Life	0.55	0.50	1.00	0.00	1.00	0.02
Sports Indoor	0.55	0.50	0.96	0.20	0.96	0.19
Sports Outdoor	0.55	0.50	0.94	0.23	0.93	0.25
Culture and Education	0.54	0.50	0.96	0.20	0.97	0.16
Observations	28070		61353		36892	

Notes: This table reports mean and standard deviation of the NPIs considered in this study, by phase of the pandemic. The variables present the fraction of days a particular NPI was in place.

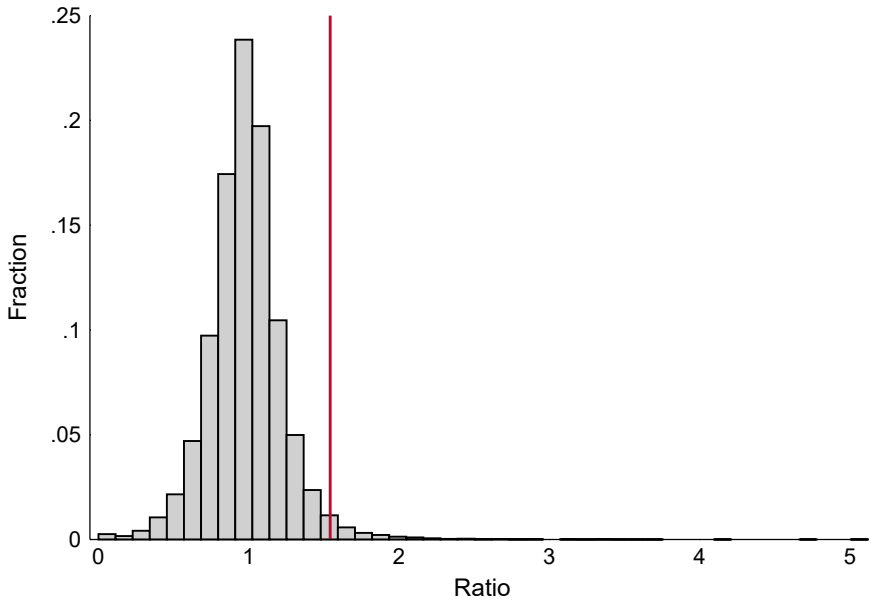
C2 Empirical Strategy

Table C.4: Incidence Model

		Incidence Rates$_t$
Incidence Rate$_{t-1}$		0.1465***
		(0.011)
Incidence Rate$_{t-2}$		0.0921***
		(0.011)
Incidence Rate$_{t-3}$		0.0956***
		(0.009)
Incidence Rate$_{t-4}$		0.0945***
		(0.011)
Incidence Rate$_{t-5}$		0.0709***
		(0.011)
Incidence Rate$_{t-6}$		0.0971***
		(0.009)
Incidence Rate$_{t-7}$		0.0468***
		(0.012)
Observations	131127	131127
Adj.R2.	0.742	0.777
FE	Yes	Yes

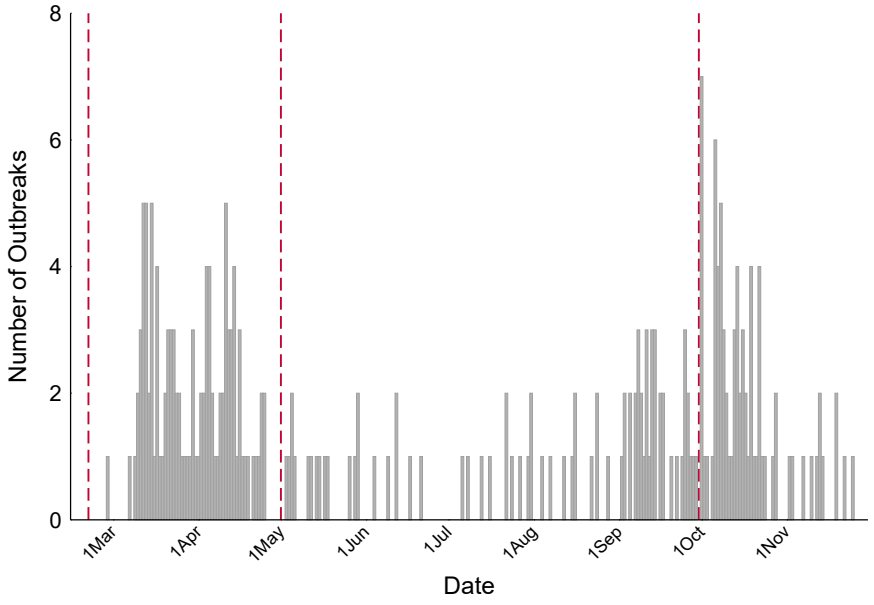
Notes: The table displays the estimates of Equation 3.1 in the text. Column 1 includes county times day-of-the-week times phase fixed effects ($\gamma^i_{c,dotw_t,p_t}$) and state times day fixed effects (δ^i_{st}). In Column 2, we additionally control for lagged incidence rates.

Figure C.1: Distribution of the Seven-Day Incidence Rate Relative to the Expected Seven-Day Incidence Rate



Notes: The figure illustrates the empirical distribution of the 7-day incidence rates relative to the expected 7-day incidence rates if the number of expected cases within that 7-day window exceeds 35. The red line displays the 98th percentile threshold.

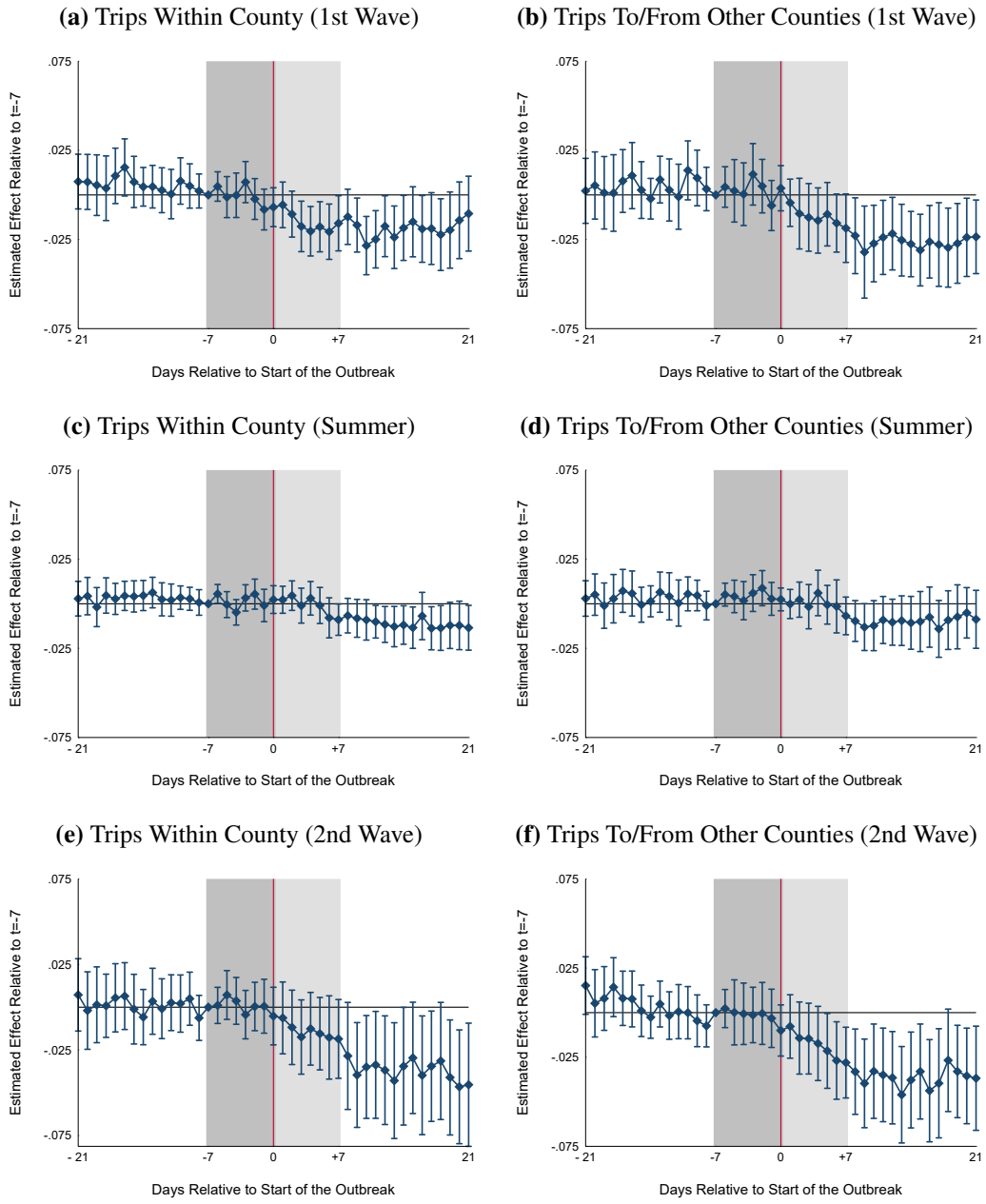
Figure C.2: Distribution of Outbreaks, per Day



Notes: The figure illustrates the number of outbreaks per day. The red dash vertical lines represent the classification into the phases of the pandemic.

C3 Results

Figure C.3: Event Study of County-Level Mobility for Different Phases



Notes: Event studies of the log of number of trips taken within a county with an outbreak and between the outbreak county and other counties. For three different phases of the Covid-19 epidemic in Germany. Model controls for day times state and county times day-of-the-week fixed effects, accounting for state-level policy changes. Time $\Delta t = 0$ is the first day we observe excess cases in an outbreak (see Section 3.4.2). We normalize at $\Delta t = -7$ as we might expected mobility to change in the seven days before the outbreak is observed in testing data due to private information. Hence, the left (dark grey) shaded area represents the period where we would expect changes in mobility due to private information. The right (light grey) shaded area represents the seven-day period when the excess of cases is reported and information about the outbreak is revealed by the case data.

Table C.5: Sensitivity Analysis: Description of Outbreaks

	Main Specification	**Change in Sum-Count Threshold**		**Change in X-Threshold**		**Change in Poisson-Threshold**	
		28	**49**	**97th**	**99th**	**0.01**	**0.001**
1st wave	105 (77, 28)	103 (74, 29)	98 (69, 29)	137 (104, 33)	61 (45, 16)	105 (77, 28)	105 (77, 28)
Summer	80 (36, 44)	98 (45, 53)	52 (23, 29)	101 (47, 54)	41 (15, 26)	79 (35, 44)	77 (33, 44)
2nd wave	74 (48, 26)	72 (59, 13)	81 (56, 25)	115 (85, 30)	38 (23, 15)	75 (49, 26)	77 (51, 26)
Total	259 (161, 98)	273 (178, 95)	231 (148, 83)	353 (236, 117)	140 (83, 57)	259 (161, 98)	259 (161, 98)

Notes: The table summarizes the number of outbreaks per phase of the pandemic. The first number in the bracket refers to the number identified by Equation 3.2a and the second is identified by Equation 3.2b, respectively.

Table C.6: Sensitivity Analysis: Changes in Outbreak Definition

	Main Specification		**Change in Sum-Count Threshold**				**Change in X-Threshold**				**Change in Poisson-Threshold**			
			28		**49**		**97th**		**99th**		**0.01**		**0.001**	
	Within	**Between**	**Within**	**Between**	**Within**	**Between**	**Within**	**Between**	**Within**	**Between**	**Within**	**Between**	**Within**	**Between**
Private Information ($\Delta t \in (-7,-1)$)	-0.0007	0.0012	0.0003	0.0016	-0.0022	0.0006	-0.0010	0.0005	-0.0017	0.0006	0.0001	0.0010	-0.0016	-0.0007
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Public Information ($\Delta t \in (0,6)$)	-0.0100*	-0.0084*	-0.0076*	-0.0076*	-0.0099*	-0.0084*	-0.0099**	-0.0071*	-0.0131*	-0.0129*	-0.0119**	-0.0124**	-0.0128**	-0.0142***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)	(0.006)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)
Post Period ($\Delta t > 6$)	-0.0217***	-0.0237***	-0.0170***	-0.0205***	-0.0193***	-0.0228***	-0.0198***	-0.0216***	-0.0275***	-0.0335***	-0.0222***	-0.0245***	-0.0221***	-0.0238***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.006)	(0.005)	(0.004)	(0.004)	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	140350	140350	140350	140350	140350	140350	140350	140350	140350	140350	140350	140350	140350	140350
Adj.R2.	0.994	0.992	0.994	0.992	0.994	0.991	0.994	0.992	0.994	0.991	0.994	0.992	0.994	0.992
FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents results from a DiD regression using different definitions for outbreaks. Columns 1 and 2 refer to the main specification. Columns 3 to 6 summarize results when changing the decision rule such that the expected number of cases must be above 28 (49) within a week (i.e. Small-Count Threshold). Columns 7 to 10 present results defining the X-threshold by the 97th or 99th percentile (i.e. X-Threshold). In Column 11 to 14 we change the threshold used to define the first day out of the seven day window when the number of cases is above the 99th percentile or about the 99.9th percentile of the Poisson distribution given the number of expected cases from the model (i.e. Poisson-Threshold). Standard errors are clustered at the county level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.7: Sensitivity Analysis: Summer Outbreaks defined by RKI

	Outbreaks in RKI Reports		**Outbreaks only in Our Study**		**Outbreaks only in RKI Reports**		**Outbreaks in RKI Reports & Our Study**	
	Within	**Between**	**Within**	**Between**	**Within**	**Between**	**Within**	**Between**
Private Information ($\Delta t \in (-7, -1)$)	-0.0012	-0.0017	0.0023	0.0012	0.0018	-0.0009	-0.0048	0.0023
	(0.003)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.006)
Public Information ($\Delta t \in (0, 6)$)	-0.0056	-0.0081	0.0003	-0.0010	-0.0030	-0.0079	-0.0034	-0.0003
	(0.005)	(0.004)	(0.006)	(0.006)	(0.007)	(0.006)	(0.006)	(0.008)
Post Period ($\Delta t > 6$)	-0.0085	-0.0135**	-0.0023	0.0002	0.0001	-0.0103	-0.0203***	-0.0201*
	(0.005)	(0.005)	(0.006)	(0.007)	(0.007)	(0.007)	(0.006)	(0.008)
Observations	140350	140350	140350	140350	140350	140350	140350	140350
Adj.R2.	0.994	0.991	0.994	0.991	0.994	0.991	0.994	0.991
Basic FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

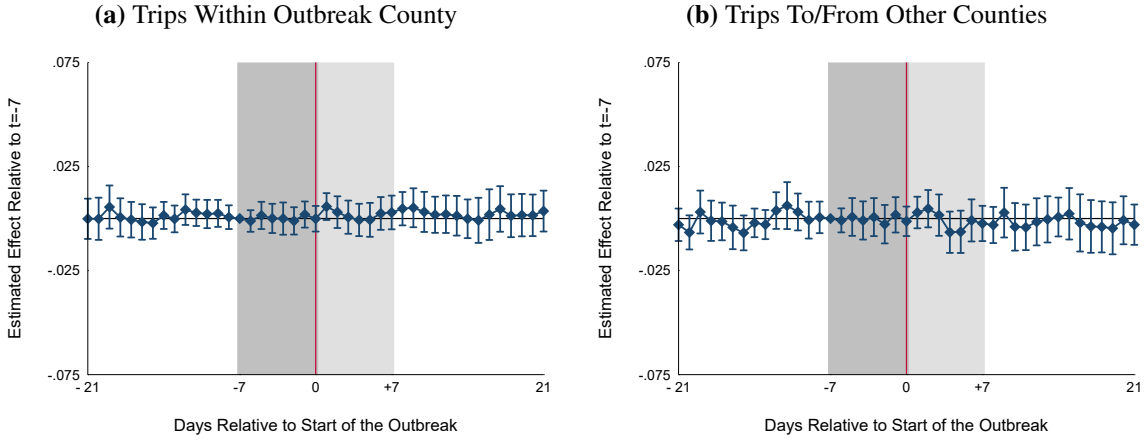
Notes: The table compares outbreaks listed in the daily reports of the German national health institute (Robert Koch Institute) and outbreaks identified by Equation 3.2a and Equation 3.2b between May, 2020 and September, 2020. Columns 1 and 2 consider the 103 outbreaks mentioned in the RKI reports. The date of the onset of an outbreak is defined by first time the outbreak is mentioned in the reports. Columns 3 and 4 include outbreaks identified by Equation 3.2a and Equation 3.2b, but not listed in the reports. Columns 5 to 6 consider outbreaks only identified by RKI reports, but not in our specification. Columns 7 to 8 include outbreaks identified by RKI reports and in our specification. Standard errors are clustered at the county level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table C.8: Sensitivity Analysis: Outbreaks identified by Ratio/Fixed Criterion

	Within	**Between**
Private Information × Ratio Criterion (Eq. 3.2a)	-0.002	-0.001
	(0.004)	(0.003)
Private Information × Fixed Criterion (Eq. 3.2b)	0.002	0.005
	(0.003)	(0.004)
Public Information × Ratio Criterion (Eq. 3.2a)	-0.011	-0.009
	(0.006)	(0.006)
Public Information × Fixed Criterion (Eq. 3.2b)	-0.009*	-0.007
	(0.004)	(0.005)
Post Period × Ratio Criterion (Eq. 3.2a)	-0.021**	-0.025***
	(0.007)	(0.007)
Post Period × Fixed Criterion (Eq. 3.2b)	-0.022***	-0.021**
	(0.005)	(0.007)
Observations	140350	140350
Adj.R2.	0.994	0.992
FE	Yes	Yes

Notes: The table studies compares behavioral responses to local outbreaks identified by Equation 3.2a and by Equation 3.2b. Standard errors are clustered at the county level. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure C.4: Sensitivity Analysis: Event Study using Placebo Outbreaks



Notes: Event studies of the log of number of trips taken within a county with a "placebo" outbreak. Placebo outbreaks are randomly assigned to another county as a real outbreak using the same date. Model controls for day times state and county times day-of-the-week fixed effects, accounting for state-level policy changes. Relative time $\Delta t = 0$ is the first day of the outbreak (see Section 3.4.2). We normalize at $\Delta t = -7$ as we might expect mobility to change in the seven days before the observed outbreak due to private information. Hence, the left (dark grey) shaded area represents the period where we would expect changes in mobility due to private information. The right (light grey) shaded area represents the seven-day period when the excess of cases is reported and information about the outbreak is revealed by the case data.

List of Tables

List of Figures

Bibliography

Adda, Jérôme. 2016. “Economic Activity and the Spread of Viral Diseases: Evidence from High Frequency Data.” *The Quarterly Journal of Economics* 131 (2): 891–941.

Ajzenman, Nicolás, Tiago Cavalcanti, and Daniel Da Mata. 2020. “More Than Words: Leaders’ Speech and Risky Behavior during a Pandemic.” *Available at SSRN:* https://ssrn.com/abstract=3582908.

Akerlof, George A. 1970. “The Market for "Lemons": Quality Uncertainty and the Market Mechanism.” *The Quarterly Journal of Economics* 84 (3): 488.

Akerman, Anders, Ingvil Gaarder, and Magne Mogstad. 2015. “The Skill Complementarity of Broadband Internet.” *The Quarterly Journal of Economics* 130 (4): 1781–1824.

Allcott, Hunt, Levi Boxell, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Yang. 2020. “Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic.” *Journal of Public Economics* 191:104254.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. “Childhood Circumstances and Adult Outcomes: Act II.” *Journal of Economic Literature* 56 (4): 1360–1446.

Almond, Douglas, and Bhashkar Mazumder. 2005. “The 1918 Influenza Pandemic and Subsequent Health Outcomes: An Analysis of SIPP Data.” *The American Economic Review* 95 (2): 258–262.

Alsan, Marcella, David Bloom, and David Canning. 2006. “The Effect of Population Health on Foreign Direct Investment Inflows to Low- and Middle-Income Countries.” *World Development* 34 (4): 613–630.

Amboss. 2021. *Influenza - Wissen für Mediziner.* Accessed March 5, 2021. https://www.amboss.com/de/wissen/Influenza.

Andersen, Martin. 2020. “Early Evidence on Social Distancing in Response to COVID-19 in the United States.” *Available at SSRN:* https://ssrn.com/abstract=3569368.

Arrow, Kenneth J. 1963. “Uncertainty and the Welfare Economics of Medical Care.” *American Economic Review* 53 (5): 941–973.

Ashenfelter, Orley. 2006. "Measuring the Value of a Statistical Life: Problems and Prospects." *The Economic Journal* 116 (510): C10–C23.

Baggett, Travis P., Harrison Keyes, Nora Sporn, and Jessie M. Gaeta. 2020. "Prevalence of SARS-CoV-2 Infection in Residents of a Large Homeless Shelter in Boston." *JAMA* 323 (21): 2191–2192.

Baird, Sarah, Joan Hamory Hicks, Michael Kremer, and Edward Miguel. 2016. "Worms at Work: Long-run Impacts of a Child Health Investment." *The Quarterly Journal of Economics* 131 (4): 1637–1680.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well–Being." *Journal of Political Economy* 116 (4): 709–745.

———. 2015. "Non-Cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program." *National Bureau of Economic Research Working Paper,* no. 21571.

Ball, Thomas M., Jose A. Castro-Rodriguez, Kent A. Griffith, Catherine J. Holberg, Fernando D. Martinez, and Anne L. Wright. 2000. "Siblings, Day-Care Attendance, and the Risk of Asthma and Wheezing during Childhood." *The New England Journal of Medicine* 343 (8): 538–543.

Ball, Thomas M., Catharine J. Holberg, Michael B. Aldous, Fernando D. Martinez, and Anne L. Wright. 2002. "Influence of Attendance at Day Care on the Common Cold From Birth Through 13 Years of Age." *Archives of Pediatrics & Adolescent Medicine* 156 (2): 121–126.

Banerjee, Abhijit, Marcella Alsan, Emily Breza, Arun G. Chandrasekhar, Abhijit Chowdhury, Esther Duflo, Paul Goldsmith-Pinkham, and Benjamin A. Olken. 2020. "Messages on COVID-19 Prevention in India Increased Symptoms Reporting and Adherence to Preventive Behaviors Among 25 Million Recipients with Similar Effects on Non-recipient Members of Their Communities." *National Bureau of Economic Research Working Paper,* no. 27496.

Bärnighausen, Till, David E. Bloom, Elizabeth T. Cafiero-Fonseca, and Jennifer Carroll O'Brien. 2014. "Valuing Vaccination." *Proceedings of the National Academy of Sciences of the United States of America* 111 (34): 12313–12319.

Barrios, John M., and Yael V. Hochberg. 2020. "Risk Perception Through the Lens of Politics in the Time of the COVID-19 Pandemic." *National Bureau of Economic Research Working Paper,* no. 27008.

Bauer, Cici, and Jon Wakefield. 2018. "Stratified Space–Time Infectious Disease Modelling, with an Application to Hand, Foot and Mouth Disease in China." *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 67 (5): 1379–1398.

Bauernschuster, Stefan, Timo Hener, and Helmut Rainer. 2016. "Children of a (Policy) Revolution: The Introduction of Universal Child Care and its Effect on Fertility." *Journal of the European Economic Association* 14 (4): 975–1005.

Bauernschuster, Stefan, and Martin Schlotter. 2015. "Public Child Care and Mothers' Labor Supply—Evidence from Two Quasi-Experiments." *Journal of Public Economics* 123:1–16.

Bayer, Ronald, and Sandro Galea. 2015. "Public Health in the Precision-Medicine Era." *The New England Journal of Medicine* 373 (6): 499–501.

BBSR. 2019. "Indikatoren und Karten zur Raum und Stadtentwickling - Erläuterungen Raumbezüge Inkar." Bundesinstitut für Bau-, Stadt- und Raumforschung. Technical Report.

Bertelsmann Stiftung. 2018. *ElternZOOM 2018 Schwerpunkt: Elternbeteiligung an der KiTa-Finanzierung*. Technical Report.

———. 2021a. *Betreuungsquote und Betreuungswunsch.* Accessed March 1, 2022. https://www.laendermonitor.de/de/vergleich-bundeslaender-daten/kinder-und-eltern/fokus-u3/betreuungsquote-und-betreuungswunsch-2?tx_itaohyperion_pluginview%5C%5Baction%5C%5D=chart&tx_itaohyperion_pluginview%5C%5Bcontroller%5C%5D=PluginView&cHash=05878fccc1a75e7278dd05a7a99fecfd.

———. 2021b. *Finanzierungsgemeinschaft für FBBE.* Accessed March 1, 2022. https://www.laendermonitor.de/de/vergleich-bundeslaender-daten/finanzen/finanzierung/finanzierungsgemeinschaft-fuer-fbbe-1?tx_itaohyperion_pluginview%5C%5Baction%5C%5D=chart&tx_itaohyperion_pluginview%5C%5Bcontroller%5C%5D=PluginView&cHash=a93eda82cb8721d9baf81759ddc7dc29.

———. 2021c. *KiTas nach Träger.* Accessed March 1, 2022. https://www.laendermonitor.de/de/vergleich-bundeslaender-daten/personal-und-einrichtungen/traeger/kitas-nach-traeger-1?tx_itaohyperion_pluginview%5C%5Baction%5C%5D=chart&tx_itaohyperion_pluginview%5C%5Bcontroller%5C%5D=PluginView&cHash=ebcae510dcaba92198a99eb01bce2e030.

Bhalotra, Sonia R., and Atheendar Venkataramani. 2015. "Shadows of the Captain of the Men of Death: Early Life Health Interventions, Human Capital Investments, and Institutions." *Available at SSRN:* https://ssrn.com/abstract=1940725.

Bien, Walter, Thomas Rauschenbach, and Birgit Riedel, eds. 2006. *Wer Betreut Deutschlands Kinder? DJI-Kinderbetreuungsstudie.* 1. Auflage. Weinheim: Beltz.

Biggerstaff, Matthew, Simon Cauchemez, Carrie Reed, Manoj Gambhir, and Lyn Finelli. 2014. "Estimates of the Reproduction Number for Seasonal, Pandemic, and Zoonotic Influenza: A Systematic Review of the Literature." *British Medical Journal Infectious Diseases* 14 (1): 1–20.

Bleakley, Hoyt. 2007. "Disease and Development: Evidence from Hookworm Eradication in the American South." *The Quarterly Journal of Economics* 122 (1): 73–117.

Bloom, David E., Michael Kuhn, and Klaus Prettner. 2020. "Modern Infectious Diseases: Macroeconomic Impacts and Policy Responses." *National Bureau of Economic Research Working Paper,* no. 27757.

Bollinger, Bryan, Phillip Leslie, and Alan Sorensen. 2011. "Calorie Posting in Chain Restaurants." *American Economic Journal: Economic Policy* 3 (1): 91–128.

Brockmann, D., L. Hufnagel, and T. Geisel. 2006. "The Scaling Laws of Human Travel." *Nature* 439 (7075).

Brosig-Koch, Jeannette, Heike Henning-Schmidt, Nadja Kairies-Schwarz, Johanna Kokot, and Daniel Wiesen. 2019. "Physician Performance Pay: Experimental Evidence." *Available at SSRN:* https://ssrn.com/abstract=3467583.

Brzezinski, Adam, Guido Deiana, Valentin Kecht, and David van Dijcke. 2020. "The COVID-19 Pandemic: Government vs. Community Action Across the United States." *Covid Economics: Vetted and Real-Time Papers* 7:115–156.

Bundesministerium für Familie, Senioren, Frauen und Jugend. 2012. *Dritter Zwischenbericht zur Evaluation des Kinderförderungsgesetzes.* Technical Report.

———. 2020. "Kindertagesbetreuung Kompakt. Ausbaustand und Bedarf 2019." Technical Report.

Bundesministerium für Gesundheit. 2021. *Coronavirus SARS-CoV-2: Chronik der bisherigen Maßnahmen.* Accessed March 5, 2021. https://www.bundesgesundheitsministerium.de/coronavirus/chronik-coronavirus.html.

Bütikofer, Aline, and Kjell G. Salvanes. 2018. "Disease Control and Inequality Reduction: Evidence from a Tuberculosis Testing and Vaccination Campaign." *Review of Economic Studies* 87 (5): 2087–2125.

Caldwell, John, Pat Caldwell, John Anarfi, Kofi Awusabo-Asare, James Ntozi, I. O. Orubuloye, Jeff Marck, Wendy Cosford, Rachel Colombo, and Elaine Hollings. 1999. *Resistances to Behavioural Change to Reduce HIV/AIDS Infection in Predominantly Heterosexual Epidemics in Third World Countries.* Health Transition Centre, National Centre for Epidemiology and Population Health, Australian National University.

Camehl, Georg, and H. Pete Frauke. 2017. "Je höher die Kita-Qualität, desto prosozialer das Verhalten von Kindern." *Deutsches Institut für Wirtschaftsforschung Wochenbericht* 84 (51/52): 1197–1204.

Camerer, Colin F., and Eric J. Johnson. 1991. "The Process-Performance Paradox in Expert Judgment: How can experts know so much and predict so badly?" In *Towards a General Theory of Expertise: Prospects and Limits,* edited by K. Anders Ericsson and Jacqui Smith, 195–217. New York: Cambridge University Press.

Case, Anne, Darren Lubotsky, and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review* 92 (5): 1308–1334.

Cauchemez, Simon, Alain-Jacques Valleron, Pierre-Yves Boëlle, Antoine Flahault, and Neil M. Ferguson. 2008. "Estimating the Impact of School Closure on Influenza Transmission from Sentinel data." *Nature* 452 (7188): 750–754.

Cawley, John, Alex M. Susskind, and Barton Willage. 2020. "Does Information Disclosure Improve Consumer Knowledge? Evidence from a Randomized Experiment of Restaurant Menu Calorie Labels." *National Bureau of Economic Research Working Paper,* no. 27126.

Center fors Disease Control and Prevention. 2020. *Vaccine Effectiveness: How Well Do the Flu Vaccines Work?* Accessed March 5, 2021. https://www.cdc.gov/flu/vaccines-work/vaccineeffect.htm.

Centers for Disease Control and Prevention. 2021. *People with Certain Medical Conditions.* Accessed March 5, 2021. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/people-with-medical-conditions.html.

Chan, Tat Y., Barton H. Hamilton, and Nicholas W. Papageorge. 2016. "Health, Risky Behaviour and the Value of Medical Innovation for Infectious Disease." *The Review of Economic Studies* 83 (4): 1465–1510.

Chen, Jonathan H., and Steven M. Asch. 2017. "Machine Learning and Prediction in Medicine — Beyond the Peak of Inflated Expectations." *The New England Journal of Medicine* 376 (26): 2507–2509.

Chen, Nanshan, Min Zhou, Xuan Dong, Jieming Qu, Fengyun Gong, Yang Han, Yang Qiu, Jingli Wang, Ying Liu, Yuan Wei, et al. 2020. "Epidemiological and Clinical Characteristics of 99 Cases of 2019 Novel Coronavirus Pneumonia in Wuhan, China: A Descriptive Study." *The Lancet* 395 (10223): 507–513.

Chu, Derek K., Elie A. Akl, Stephanie Duda, Karla Solo, Sally Yaacoub, Holger J. Schünemann, Amena El-harakeh, Antonio Bognanni, Tamara Lotfi, Mark Loeb, Anisa Hajizadeh, Anna Bak, Ariel Izcovich, Carlos A. Cuello-Garcia, Chen Chen, David J. Harris, Ewa Borowiack, Fatimah Chamseddine, Finn Schünemann, Gian Paolo Morgano, Schünemann, Giovanna E U Muti, Guang Chen, Hong Zhao, Ignacio Neumann, Jeffrey Chan, Joanne Khabsa, Layal Hneiny, Leila Harrison, Maureen Smith, Nesrine Rizk, Paolo Giorgi Rossi, Pierre AbiHanna, Rayane El-khoury, Rosa Stalteri, Tejan Baldeh, Thomas Piggott, Yuan Zhang, Zahra Saad, Assem Khamis, and Marge Reinap. 2020. "Physical Distancing, Face Masks, and Eye Protection to Prevent Person-to-Person Transmission of SARS-CoV-2 and COVID-19: A Systematic Review and Meta-Analysis." *The Lancet* 395 (10242): 1973–1987.

Clark, Alexander M., Marie DesMeules, Wei Luo, Amanda S. Duncan, and Andy Wielgosz. 2009. "Socioeconomic Status and Cardiovascular Disease: Risks and Implications for Care." *Nature Reviews Cardiology* 6 (11): 712–722.

Conroy, R. 2003. "Estimation of Ten-Year Risk of Fatal Cardiovascular Disease in Europe: The SCORE Project." *European Heart Journal* 24 (11): 987–1003.

Côté, Sylvana M., Amélie Petitclerc, Marie-France Raynault, Qian Xu, Bruno Falissard, Michel Boivin, and Richard E. Tremblay. 2010. "Short- and Long-term Risk of Infections as a Function of Group Child Care Attendance: An 8-Year Population-Based Study." *Archives of Pediatrics & Adolescent Medicine* 164 (12): 1132–1137.

COVID-19 Datenhub. 2021. *RKI COVID19.* Accessed March 5, 2021. https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/dd4580c810204019a7b8eb3e0b329dd6_0/data.

Cummings, Linda C., and Gregory S. Cooper. 2011. "Colorectal Cancer Screening: Update for 2011." *Seminars in Oncology* 38 (4): 483–489.

Currie, Janet. 2009. "Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development." *Journal of Economic Literature* 47 (1): 87–122.

Currie, Janet, and Hannes Schwandt. 2013. "Within-Mother Analysis of Seasonal Patterns in Health at Birth." *Proceedings of the National Academy of Sciences of the United States of America* 110 (32): 13228.

Cutler, David, Jonathan S. Skinner, Ariel Dora Stern, and David Wennberg. 2019. "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending." *American Economic Journal: Economic Policy* 11 (1): 192–221.

Dawes, Robyn M., David Faust, and Paul E. Meehl. 1989. "Clinical versus Actuarial Judgment." *Science* 243 (4899): 1668–1674.

Deyle, Ethan R., M. Cyrus Maher, Ryan D. Hernandez, Sanjay Basu, and George Sugihara. 2016. "Global Environmental Drivers of Influenza." *Proceedings of the National Academy of Sciences of the United States of America* 113 (46): 13081–13086.

Diederichs, Claudia, Hannelore Neuhauser, Viktoria Rücker, Markus A. Busch, Ulrich Keil, Anthony P. Fitzgerald, Peter U. Heuschmann, and Stefan Kiechl. 2018. "Predicted 10-year Risk of Cardiovascular Mortality in the 40 to 69 Year Old General Population Without Cardiovascular Diseases in Germany." *PLOS ONE* 13 (1).

Diekmann, Laura, and Michael Thöne. 2011. "Föderale Finanzierung des Kinderbetreuungsausbaus: Ermittlung der Lastenverteilung." Technical Report.

Dupas, Pascaline. 2011. "Do Teenagers Respond to HIV Risk Information? Evidence from a Field Experiment in Kenya." *American Economic Journal: Applied Economics* 3 (1): 1–34.

Endo, Akira, Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Sam Abbott, Adam J. Kucharski, and Sebastian Funk. 2020. "Estimating the Overdispersion in COVID-19 Transmission using Outbreak Sizes outside China." *Wellcome Open Research* 5:67.

Enke, Benjamin, and Thomas Graeber. 2021. "Cognitive Uncertainty." *National Bureau of Economic Research Working Paper,* no. 26518.

European Centre for Disease Prevention and Control. 2017. *Seasonal Influenza Vaccination in Europe: Vaccination Recommendations and Coverage Rates in the EU Member States for eight Influenza Seasons, 2007–2008 to 2014–2015.* Accessed March 1, 2022. https://www.ecdc.europa.eu/en/publications-data/seasonal-influenza-vaccination-europe-vaccination-recommendations-and-coverage-2007-2015.

———. 2020. "COVID-19 in Children and the Role of School Settings in Transmission - First Update." Technical Report.

European Centre for Disease Prevention and Control. 2021a. *Latest evidence on COVID-19.* Accessed March 5, 2021. https://www.ecdc.europa.eu/en/covid-19/latest-evidence.

———. 2021b. *Timeline of ECDC's Reponse to COVID-19.* Accessed March 5, 2021. https://www.ecdc.europa.eu/en/covid-19/timeline-ecdc-response.

Fang, Hanming, Long Wang, and Yang Yang. 2020. "Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China." *Journal of Public Economics* 191:104272.

Federal Institute for Drugs and Medical Devices. 2021. *Iris Institute.* Accessed March 5, 2021. https://www.dimdi.de/dynamic/en/classifications/iris-institute/.

Felfe, Christina, and Rafael Lalive. 2018. "Does Early Child Care Affect Children's Development?" *Journal of Public Economics* 159:33–53.

Felfe, Christina, and Larissa Zierow. 2018. "From Dawn Till Dusk: Implications of Full-Day Care for Children's Development." *Labour Economics* 55:259–281.

Ferguson, N., D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, ZULMA Cucunuba Perez, G. Cuomo-Dannenburg, A. Dighe, I. Dorigatti, H. Fu, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, L. Okell, S. van Elsland, H. Thompson, R. Verity, E. Volz, H. Wang, Y. Wang, P. Walker, P. Winskill, C. Whittaker, C. Donnelly, S. Riley, and A. Ghani. 2020. "Report 9: Impact of Non-Pharmaceutical Interventions (NPIs) to Reduce COVID19 Mortality and Healthcare Demand." *Imperial College London* 10 (77482): 491–497.

Finkenstädt, Bärbel, and Bryan Grenfell. 2000. "Time Series Modelling of Childhood Dieseases: A Dynamical Systems Approach." *Journal of the Royal Statistical Society. Series C: Applied Statistics,* 187–205.

Gail, Mitchell H., Louise A. Brinton, David P. Byar, Donald K. Corle, Sylvan B. Green, Catherine Schairer, and John J. Mulvihill. 1989. "Projecting Individualized Probabilities of Developing Breast Cancer for White Females Who Are Being Examined Annually." *JNCI: Journal of the National Cancer Institute* 81 (24): 1879–1886.

Gandhi, M., D. Yokoe, and D. Havlir. 2020a. "Asymptomatic Transmission, the Achilles' Heel of Current Strategies to Control Covid-19." *The New England Journal of Medicine* 382 (22): 2158–2160.

Gandhi, Rajesh, John Lynch, and Carlos Del Rio. 2020b. "Mild or Moderate Covid-19." *The New England Journal of Medicine* 383 (18): 1757–1766.

Gaudecker, Hans-Martin von, Radost Holler, Lena Janys, Bettina M. Siflinger, and Christian Zimpelmann. 2020. "Labour Supply in the Early Stages of the COVID-19 Pandemic: Empirical Evidence on Hours, Home Office, and Expectations." *IZA DP,* no. 13158.

Geis, Wido. 2018. "Familien müssen für die gleiche Betreuung in der Kita unterschiedlich viel zahlen - ein Vergleich der Gebührenordnung der größten Städte in Deutschland." Institut der deutschen Wirtschaft (IW) Köln, *IW Report,* no. 50.

Gesundheitsberichterstattung des Bundes. 2021. *Sterbefälle, Sterbeziffern (je 100.000 Einwohner, altersstandardisiert) (ab 1998). Gliederungsmerkmale: Jahre, Region, Alter, Geschlecht, Nationalität, ICD-10, Art der Standardisierung.* Accessed March 5, 2021. https://www.gbe-bund.de/gbe/pkg_isgbe5.prc_menu_olap?p_uid=gast&p_aid=13057327&p_sprache=D&p_help=0&p_indnr=6&p_indsp=6071&p_ityp=H&p_fid=.

Gilliam, Walter S., Amyn A. Malik, Mehr Shafiq, Madeline Klotz, Chin Reyes, John Eric Humphries, Thomas Murray, Jad A. Elharake, David Wilkinson, and Saad B. Omer. 2021. "COVID-19 Transmission in US Child Care Programs." *Pediatrics* 147 (1).

Girard, Marc P., Thomas Cherian, Yuri Pervikov, and Marie Paule Kieny. 2005. "A Review of Vaccine Research and Development: Human Acute Respiratory Infections." *Vaccine* 23 (50): 5708–5724.

Giustinelli, Pamela, Charles Manski, and Francesca Molinari. 2019. "Precise or Imprecise Probabilities? Evidence from Survey Response on Late-onset Dementia." *National Bureau of Economic Research Working Paper,* no. 26125.

Glaeser, Edward L., Caitlin Gorback, and Stephen J. Redding. 2020. "How Much does COVID-19 Increase with Mobility? Evidence from New York and Four Other U.S. Cities." *National Bureau of Economic Research Working Paper,* no. 27519.

Glogowsky, Ulrich, Emanuel Hansen, and Simeon Schächtele. 2020. "How Effective are Social Distancing Policies? Evidence on the Fight Against COVID-19 from Germany." *Available at SSRN:* https://ssrn.com/abstract=3619845.

Goff, D. C., D. M. Lloyd-Jones, G. Bennett, S. Coady, R. B. D'Agostino, R. Gibbons, P. Greenland, D. T. Lackland, D. Levy, C. J. O'Donnell, J. G. Robinson, J. S. Schwartz, S. T. Shero, S. C. Smith, P. Sorlie, N. J. Stone, P. W. Wilson, H. S. Jordan, L. Nevo, J. Wnek, J. L. Anderson, J. L. Halperin, N. M. Albert, B. Bozkurt, R. G. Brindis, L. H. Curtis, D. DeMets, J. S. Hochman, R. J. Kovacs, E. M. Ohman, S. J. Pressler, F. W. Sellke, W. K. Shen, and G. F. Tomaselli. 2014. "2013 ACC/AHA Guideline on the Assessment of Cardiovascular Risk: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines." *Circulation* 129 (25 Suppl 2): 49–73.

Gonzalez, M., M. C. Pirez, E. Ward, H. Dibarboure, A. García, and H. Picolet. 2000. "Safety and Immunogenicity of a Paediatric Presentation of an Influenza Vaccine." *Archives of Disease in Childhood* 83 (6): 488–491.

Grenfell, Bryan T., Ottar Bjùrnstad, and Matt Kappey. 2001. "Travelling Waves and Spatial Hierarchies in Measles Epidemics." *Nature* 414 (6865): 716–723.

Grossman, Guy, Soojong Kim, Jonah M. Rexer, and Harsha Thirumurthy. 2020. "Political Partisanship Influences Behavioral Responses to Governors' Recommendations for COVID-19 Prevention in the United States." *Proceedings of the National Academy of Sciences* 117 (39): 24144–24153.

Guan, Wei-Jie, Zheng-Yi Ni, Yu Hu, Wen-Hua Liang, Chun-Quan Ou, Jian-Xing He, Lei Liu, Hong Shan, Chun-Liang Lei, David S. C. Hui, Bin Du, Lan-Juan Li, Guang Zeng, Kwok-Yung Yuen, Ru-Chong Chen, Chun-Li Tang, Tao Wang, Ping-Yan Chen, Jie Xiang, Shi-Yue Li, Jin-Lin Wang, Zi-Jing Liang, Yi-Xiang Peng, Li Wei, Yong Liu, Ya-Hua Hu, Peng Peng, Jian-Ming Wang, Ji-Yang Liu, Zhong Chen, Gang Li, Zhi-Jian Zheng, Shao-Qin Qiu, Jie Luo, Chang-Jiang Ye, Shao-Yong Zhu, and Nan-Shan Zhong. 2020. "Clinical Characteristics of Coronavirus Disease 2019 in China." *The New England Journal of Medicine* 382 (18): 1708–1720.

Gupta, Sumedha, Thuy Nguyen, Felipe Lozano Rojas, Shyam Raman, Byungkyu Lee, Ana Bento, Kosali Simon, and Coady Wing. 2020. "Tracking Public and Private Responses to the COVID-19 Epidemic: Evidence from State and Local Government Actions." *National Bureau of Economic Research Working Paper,* no. 27027.

Gutierrez, Emilio, Adrian Rubli, and Tiago Tavares. 2020. "Information and Behavioral Responses during a Pandemic: Evidence from Delays in COVID-19 Death Reports." *Available at SSRN:* https://ssrn.com/abstract=3645317.

Hamner, Lea, Polly Dubbel, Ian Capron, Andy Ross, Amber Jordan, Jaxon Lee, Joanne Lynn, Amelia Ball, Simranjit Narwal, Sam Russell, Dale Patrick, and Howard Leibrand. 2020. "High SARS-CoV-2 Attack Rate Following Exposure at a Choir Practice - Skagit County, Washington, March 2020." *Morbidity and Mortality Weekly Report* 69 (19): 606–610.

Hay, A. J., V. Gregory, A. R. Douglas, and Y. P. Lin. 2001. "The Evolution of Human Influenza Viruses." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 356 (1416).

Heikkinen, Terho, Heli Silvennoinen, Ville Peltola, Thedi Ziegler, Raija Vainionpaa, Tytti Vuorinen, Leena Kainulainen, Tuomo Puhakka, Tuomas Jartti, Pia Toikka, Pasi Lehtinen, Taina Routi, and Taina Juven. 2004. "Burden of Influenza in Children in the Community." *The Journal of Infectious Diseases* 190 (8): 1369–1373.

Hoffman, Mitchell, Lisa B. Kahn, and Danielle Li. 2018. "Discretion in Hiring." *The Quarterly Journal of Economics* 133 (2): 765–800.

Hoffmann, Susanne, Francis G. Caro, Alison S. Gottlieb, Iris Kesternich, and Joachim K. Winter. 2014. "Contributions of Second Opinions, Outcome Forecasts, and Testimonials to Patient Decisions about Knee Replacement Surgery." *Medical Decision Making: An International Journal of the Society for Medical Decision Making* 34 (5): 603–614.

Huang, Chaolin, Lixue Huang, Yeming Wang, Xia Li, Lili Ren, Xiaoying Gu, Liang Kang, Li Guo, Min Liu, Xing Zhou, Jianfeng Luo, Zhenghui Huang, Shengjin Tu, Yue Zhao, Li Chen, Decui Xu, Yanping Li, Caihong Li, Lu Peng, Yong Li, Wuxiang Xie, Dan Cui, Lianhan Shang, Guohui Fan, Jiuyang Xu, Geng Wang, Ying Wang, Jingchuan Zhong, Chen Wang, Jianwei Wang, Dingyu Zhang, and Bin Cao. 2021. "6-Month Consequences of COVID-19 in Patients Discharged from Hospital: A Cohort Study." *The Lancet* 397 (10270): 220–232.

Huang, Chaolin, Yeming Wang, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan, Jiuyang Xu, Xiaoying Gu, Zhenshun Cheng, Ting Yu, Jiaan Xia, Yuan Wei, Wenjuan Wu, Xuelei Xie, Wen Yin, Hui Li, Min Liu, Yan Xiao, Hong Gao, Li Guo, Jungang Xie, Guangfa Wang, Rongmeng Jiang, Zhancheng Gao, Qi Jin, Jianwei Wang, and Bin Cao. 2020. "Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China." *The Lancet* 395 (10223): 497–506.

Hufnagel, Lars, Dirk Brockmann, and Theo Geisel. 2004. "Forecast and Control of Epidemics in a Globalized World." *Proceedings of the National Academy of Sciences of the United States of America* 101 (42): 15124–15129.

Hurd, Michael D. 2009. "Subjective Probabilities in Household Surveys." *Annual Review of Economics* 1 (1): 543–562.

Hut, Stefan, and Emily Oster. 2018. "Changes in Household Diet: Determinants and Predictability." *National Bureau of Economic Research Working Paper,* no. 24892.

James, Allison, Lesli Eagle, Cassandra Phillips, D. Stephen Hedges, Cathie Bodenhamer, Robin Brown, J. Gary Wheeler, and Hannah Kirking. 2020. "High COVID-19 Attack Rate Among Attendees at Events at a Church - Arkansas, March 2020." *Morbidity and Mortality Weekly Report* 69 (20): 632–635.

Kelly, Elaine. 2011. "The Scourge of Asian Flu." *Journal of Human Resources* 46 (4): 669–694.

Kermack and McKendrick. 1927. "A Contribution to the Mathematical Theory of Epidemics." *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115 (772): 700–721.

Khwaja, Ahmed, Dan Silverman, Frank Sloan, and Yang Wang. 2009. "Are Mature Smokers Misinformed?" *Journal of Health Economics* 28 (2): 385–397.

Khwaja, Ahmed, Frank Sloan, and Sukyung Chung. 2007. "The Relationship between Individual Expectations and Behaviors: Mortality Expectations and Smoking Decisions." *Journal of Risk and Uncertainty* 35 (2): 179–201.

Kim, Hyuncheol Bryant, Suejin A. Lee, and Wilfredo Lim. 2019. "Knowing is Not Half the Battle: Impacts of information from the National Health Screening Program in Korea." *Journal of Health Economics* 65:1–14.

King, James C., Ginny E. Cummings, Jeffrey Stoddard, Bernard X. Readmond, and Elizabeth Ruff. 2005. "A Pilot Study of the Effectiveness of a School-Based Influenza Vaccination Program." *Pediatrics* 116 (6): e868–e873.

Kleinberg, Jon, Himabindu Lakkaraju, Jure Leskovec, Jens Ludwig, and Sendhil Mullainathan. 2018. "Human Decisions and Machine Predictions." *The Quarterly Journal of Economics* 133 (1): 237–293.

Kniesner, Thomas J., and W. Kip Viscusi. 2019. "The Value of a Statistical Life." *Oxford Research Encyclopedia of Economics and Finance.*

Kremer, Michael. 1996. "Integrating Behavioural Choice into Epidemiological Models of Aids." *The Quarterly Journal of Economics* 111 (2): 550–573.

Lambert, Linda C., and Anthony S. Fauci. 2010. "Influenza Vaccines for the Future." *The New England Journal of Medicine* 363 (21): 2036–2044.

Lauer, Stephen A., Kyra H. Grantz, Qifang Bi, Forrest K. Jones, Qulu Zheng, Hannah R. Meredith, Andrew S. Azman, Nicholas G. Reich, and Justin Lessler. 2020. "The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application." *Annals of Internal Medicine* 172 (9): 577–582.

Lemieux, Jacob E., Katherine J. Siddle, Bennett M. Shaw, Christine Loreth, Stephen F. Schaffner, Adrianne Gladden-Young, Gordon Adams, Timelia Fink, Christopher H. Tomkins-Tinch, Lydia A. Krasilnikova, Katherine C. DeRuff, Katherine C. Melissa Rudy, Matthew R. Bauer, Kim A. Lagerborg, Erica Normandin, Sinéad B. Chapman, Steven K. Reilly, Melis N. Anahtar, Aaron E. Lin, Amber Carter, Cameron Myhrvold, Molly E. Kemball, Sushma Chaluvadi, Caroline Cusick, Katelyn Flowers, Anna Neumann, Felecia Cerrato, Maha Farhat, Damien Slater, Jason B. Harris, John A. Branda, David Hooper, Jessie M. Gaeta, Travis P. Baggett, James O'Connell, Andreas Gnirke, Tami D. Lieberman, Anthony Philippakis, Meagan Burns, Catherine M. Brown, Jeremy Luban, Edward T. Ryan, Sarah E. Turbett, Regina C. LaRocque, William P. Hanage, Glen R. Gallagher, Lawrence C. Madoff, Sandra Smole, Virginia M. Pierce, Eric Rosenberg, Pardis C. Sabeti, Daniel J. Park, and Bronwyn L. MacInnis. 2020. "Phylogenetic Analysis of SARS-CoV-2 in Boston highlights the Impact of Superspreading Events." *Science* 371 (6529): eabe3261.

Li, Qun, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy S. M. Leung, Eric H. Y. Lau, Jessica Y. Wong, Xuesen Xing, Nijuan Xiang, Yang Wu, Chao Li, Qi Chen, Dan Li, Tian Liu, Jing Zhao, Man Liu, Wenxiao Tu, Chuding Chen, Lianmei Jin, Rui Yang, Qi Wang, Suhua Zhou, Rui Wang, Hui Liu, Yinbo Luo, Yuan Liu, Ge Shao, Huan Li, Zhongfa Tao, Yang Yang, Zhiqiang Deng, Boxi Liu, Zhitao Ma, Yanping Zhang, Guoqing Shi, Tommy T. Y. Lam, Joseph T. Wu, George F. Gao, Benjamin J. Cowling, Bo Yang, Gabriel M. Leung, and Zijian Feng. 2020. "Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia." *The New England Journal of Medicine* 382 (13): 1199–1207.

Lindman, Anja S., Marit B. Veierød, Jan I. Pedersen, Aage Tverdal, Inger Njølstad, and Randi Selmer. 2007. "The Ability of the SCORE High-Risk Model to Predict 10-year Cardiovascular Disease Mortality in Norway." *European Journal of Cardiovascular Prevention & Rehabilitation* 14 (4): 501–507.

Lopez, Adriana S., Mary Hill, Jessica Antezano, Dede Vilven, Tyler Rutner, Linda Bogdanow, Carlene Claflin, Ian T. Kracalik, Victoria L. Fields, Angela Dunn, Jacqueline E. Tate, Hannah L. Kirking, Tair Kiphibane, Ilene Risk, and Cuc H. Tran. 2020. "Transmission Dynamics of COVID-19 Outbreaks Associated with Child Care Facilities — Salt Lake City, Utah, April–July 2020." *Morbidity and Mortality Weekly Report* 69 (37): 1319–1323.

Lowen, Anice C., and John Steel. 2014. "Roles of Humidity and Temperature in Shaping Influenza Seasonality." *Journal of Virology* 88 (14): 7692–7695.

Lundborg, Petter, and Björn Lindgren. 2004. "Do They Know What They are Doing? Risk Perceptions and Smoking Behaviour Among Swedish Teenagers." *Journal of Risk and Uncertainty* 28 (3): 261–286.

Majra, Dasha, Jayme Benson, Jennifer Pitts, and Justin Stebbing. 2021. "SARS-CoV-2 (COVID-19) superspreader events." *The Journal of Infection* 82 (1): 36–40.

Manski, Charles F. 2004. "Measuring Expectations." *Econometrica* 72 (5): 1329–1376.

———. 2018. "Credible Ecological Inference for Medical Decisions with Personalized Risk Assessment." *Quantitative Economics* 9 (2): 541–569.

———. 2019. *Patient Care under Uncertainty.* Princeton University Press.

Manski, Charles F., and Francesca Molinari. 2010. "Rounding Probabilistic Expectations in Surveys." *Journal of Business & Economic Statistics* 28 (2): 219–231.

Meiner, Christiane, Svenja Rieser, and Eva Strunz. 2015. "Bedarfsgerecht oder angebotsorientiert? Eine Analyse der monatsspezifischen Aufnahmen von unter 3-Jährigen in der Kindertagesbetreuung." *KomDat Jugendhilfe* 18 (1): 12–15.

Meyer, Sebastian, and Leonhard Held. 2017. "Incorporating Social Contact Data in Spatio-Temporal Models for Infectious Disease Spread." *Biostatistics* 18 (2): 338–351.

Molinari, Noelle-Angelique M, Ismael R Ortega-Sanchez, Mark L Messonnier, William W Thompson, Pascale M Wortley, Eric Weintraub, and Carolyn B Bridges. 2007. "The Annual Impact of Seasonal Influenza in the US: Measuring Disease Burden and Costs." *Vaccine* 25 (27): 5086–5096.

Mossong, Joël, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, Janneke Heijne, Malgorzata Sadkowska-Todys, Magdalena Rosinska, and W. John Edmunds. 2008. "Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases." *PLoS Medicine* 5 (3): e74.

Müller, Kai-Uwe, and Katharina Wrohlich. 2019. "Does Subsidized Care for Toddlers Increase Maternal Labor Supply?" *Labour Economics* 62:101776.

Obermeyer, Ziad, and Ezekiel J. Emanuel. 2016. "Predicting the Future — Big Data, Machine Learning, and Clinical Medicine." *New England Journal of Medicine* 375 (13): 1216–1219.

OECD. 2020. "Education at a Glance 2020." Organization for Economic Co-operation and Organization. Technical Report.

Oliver, Nuria, Bruno Lepri, Harald Sterly, Renaud Lambiotte, Sébastien Deletaille, Marco de Nadai, Emmanuel Letouzé, Albert Ali Salah, Richard Benjamins, Ciro Cattuto, Vittoria Colizza, Nicolas de Cordes, Samuel P. Fraiberger, Till Koebe, Sune Lehmann, Juan Murillo, Alex Pentland, Phuong N. Pham, Frédéric Pivetta, Jari Saramäki, Samuel V. Scarpino, Michele Tizzoni, Stefaan Verhulst, and Patrick Vinck. 2020. "Mobile Phone Data for Informing Public Health Actions across the COVID-19 Pandemic Life Cycle." *Science Advances* 6 (23).

Oster, Emily. 2005. "Sexually Transmitted Infections, Sexual Behavior, and the HIV/AIDS Epidemic." *The Quarterly Journal of Economics* 120 (2): 467–515.

———. 2012. "Routes of Infection: Exports and HIV Incidence in Sub-Saharan Africa." *Journal of the European Economic Association* 10 (5): 1025–1058.

———. 2017. "Does Disease Cause Vaccination? Disease Outbreaks and Vaccination Response." *Journal of Health Economics* 57:90–101.

———. 2018. "Diabetes and Diet: Purchasing Behavior Change in Response to Health Information." *American Economic Journal: Applied Economics* 10 (4): 308–348.

Painter, Marcus, and Tian Qiu. 2021. "Political Beliefs affect Compliance with Government Mandates." *Journal of Economic Behavior & Organization* 185:688–701.

Parri, Niccolò, Matteo Lenge, and Danilo Buonsenso. 2020. "Children with Covid-19 in Pediatric Emergency Departments in Italy." *The New England Journal of Medicine* 383 (2): 187–190.

Pauly, Mark V. 1968. "The Economics of Moral Hazard: Comment." *The American Economic Review* 58 (3): 531–537.

Peabody, John W., Jeff Luck, Peter Glassman, Timothy R. Dresselhaus, and Martin Lee. 2000. "Comparison of Vignettes, Standardized Patients, and Chart Abstraction: A Prospective Validation Study of 3 Methods for Measuring Quality." *The Journal of the American Medical Association* 283 (13): 1715.

Peabody, John W., Jeff Luck, Peter Glassman, Sharad Jain, Joyce Hansen, Maureen Spell, and Martin Lee. 2004. "Measuring the Quality of Physician Practice by Using Clinical Vignettes: A Prospective Validation Study." *Annals of Internal Medicine* 141 (10): 771.

Philipson, Tomas. 2000. "Economic Epidemiology and Infectious Diseases." Chap. 33 in *Handbook of Health Economics,* 1:1761–1799.

Piepoli, Massimo F., Arno W. Hoes, Stefan Agewall, Christian Albus, Carlos Brotons, Alberico L. Catapano, Marie-Therese Cooney, Ugo Corrà, Bernard Cosyns, Christi Deaton, Ian Graham, Michael Stephen Hall, F. D. Richard Hobbs, Maja-Lisa Løchen, Herbert Löllgen, Pedro Marques-Vidal, Joep Perk, Eva Prescott, Josep Redon, Dimitrios J. Richter, Naveed Sattar, Yvo Smulders, Monica Tiberi, van der Worp, H. Bart, Ineke van Dis, and W. M. Monique Verschuren. 2016. "European Guidelines on Cardiovascular Disease Prevention in Clinical Practice (Version 2016)." *European Heart Journal* 37 (29): 2315–2381.

Prina, Silvia, and Heather Royer. 2014. "The Importance of Parental Knowledge: Evidence from Weight Report Cards in Mexico." *Journal of Health Economics* 37:232–247.

Prinz, Daniel, Michael Chernew, David Cutler, and Austin Frakt. 2018. "Health and Economic Activity Over the Lifecycle: Literature Review." *National Bureau of Economic Research Working Paper,* no. 24865.

Pritsch, Michael, Katja Radon, Abhishek Bakuli, Ronan Le Gleut, Laura Olbrich, Jessica Michelle Guggenbuehl Noller, Elmar Saathoff, Noemi Castelletti, Mercè Garí, Peter Puetz, Yannik Schaelte, Turid Frahnow, Roman Wölfel, Camilla Rothe, Michel Pletschette, Dafni Metaxa, Felix Forster, Verena Thiel, Friedrich Riess, Maximilian Nikolaus Diefenbach, Guenter Froeschl, Jan Bruger, Simon Winter, Jonathan Frese, Kerstin Puchinger, Isabel Brand, Inge Kroidl, Jan Hasenauer, Christiane Fuchs, Andreas Wieser, and Michael Hoelscher. 2020. "Prevalence and Risk Factors of Infection in the Representative COVID-19 Cohort Munich." *Available at SSRN:* https://ssrn.com/abstract=3745128.

Read, Jonathan. M., W. J. Edmunds, S. Riley, Justin Lessler, and Derek A. Cummings. 2012. "Close Encounters of the Infectious Kind: Methods to Measure Social Mixing Behaviour." *Epidemiology and Infection* 140 (12): 1–14.

Reuss, Annicka M., Dietmar Walter, Marcel Feig, Lutz Kappelmayer, Udo Buchholz, Tim Eckmanns, and Gabriele Poggensee. 2010. "Influenza Vaccination Coverage in the 2004/05, 2005/06, and 2006/07 Seasons: A Secondary Data Analysis based on Billing Data of the German Associations of Statutory Health Insurance Physicians." *Deutsches Arzteblatt international* 107 (48): 845–850.

Rieck, Thorsten, Marcel Feig, Ole Wichmann, and Anette Siedler. 2017. "Aktuelles aus der KV-Impfsurveillance Impfquoten der Rotavirus-, Masern-, HPV- und Influenza- Impfung in Deutschland." Robert-Koch Institut, *Epidemiologisches Bulletin* 1:1–12.

———. 2018. "Aktuelles aus der KV-Impfsurveillance Impfquoten ausgewählter Schutzimpfungen in Deutschland." Robert Koch-Institut, *Epidemiologisches Bulletin* 1:1–14.

Riou, Julien, and Christian L. Althaus. 2020. "Pattern of Early Human-to-Human Transmission of Wuhan 2019 novel Coronavirus (2019-nCoV), December 2019 to January 2020." *Eurosurveillance* 25 (4).

Robert Koch Institute. 2012. *Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2011.* Berlin: Robert Koch-Institut.

———. 2013. *Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2012.* Berlin: Robert Koch-Institut.

———. 2016a. "Aktuelles aus der KV-Impfsurveillance Impfquoten der Rotavirus-, Masern-, HPV- und Influenza- Impfung in Deutschland." *Epidemiologisches Bulletin* 1:1–10.

———. 2016b. *Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2015.* Berlin: Robert Koch-Institut.

———. 2020a. *RKI-Navigation - Für wen wird die Impfung gegen die saisonale Influenza empfohlen?* Accessed March 5, 2021. https://www.rki.de/SharedDocs/FAQ/Impfen/Influenza/FAQ01.html?nn=2444038.

———. 2020b. *RKI-Ratgeber - Influenza (Teil 1): Erkrankungen durch saisonale Influenzaviren.* Accessed January 5, 2021. https://www.rki.de/DE/Content/Infekt/EpidBull/Merkblaetter/Ratgeber_Influenza_saisonal.html.

———. 2021a. *Aktueller Lage-/Situationsbericht des RKI zu COVID-19.* Accessed March 7, 2021. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Gesamt.html.

———. 2021b. *Antworten auf häufig gestellte Fragen zum Coronavirus SARS-CoV-2 / Krankheit COVID-19.* Accessed March 5, 2021. https://www.rki.de/SharedDocs/FAQ/NCOV2019/gesamt.html;jsessionid=B2523231178C46965BA8C8E07F394EBA.internet051.

———. 2021c. *Epidemiologischer Steckbrief zu SARS-CoV-2 und COVID-19.* Accessed March 5, 2021. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Steckbrief.html.

Rose, Sherri. 2018. "Machine Learning for Prediction in Electronic Health Data." *The Journal of the American Medical Association Network Open* 1 (4).

Sachs, Jeffrey. 2001. *Macroeconomics and Health: Investing in Health for Economic Development; Report of the Commission on Macroeconomics and Health.* Technical Report. World Health Organization.

Sampat, Bhaven N. 2019. "The Economics of Health Innovation: Looking Back and Forward." Chap. 2 in *Global Innovation Index 2019 Creating Healthy Lives — The Future of Medical Innovation,* 81–85. World Intellectual Property Organization.

Sattelmair, Jacob, Jeremy Pertman, Eric L. Ding, Harold W. Kohl, William Haskell, and I-Min Lee. 2011. "Dose-Response Between Physical Activity and Risk of Coronary Heart Disease: A Meta-Analysis." *Circulation* 124 (7): 789–795.

Schmidheiny, Kurt, and Sebastian Siegloch. 2019. "On Event Study Designs and Distributed-Lag Models: Equivalence, Generalization and Practical Implications." *CESifo Working Papers,* no. 7481.

Scholz, Stefan, Oliver Damm, Udo Schneider, Bernhard Ultsch, Ole Wichmann, and Wolfgang Greiner. 2019. "Epidemiology and Cost of Seasonal Influenza in Germany - A Claims Data Analysis." *British Medical Journal Public Health* 19 (1): 1090.

Schwandt, Hannes. 2017. "The Lasting Legacy of Seasonal Influenza in Utero Exposuer and Labor Market Outcomes: Labour Economics." *IZA Discussion Papers,* no. 10589.

Smith, James P. 1999. "Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status." *The Journal of Economic Perspectives* 13 (2): 144–166.

Spence, Micheal. 1973. "Job Market Signaling." *The Quarterly Journal of Economics* 87 (3): 355–374.

Staerk, Christian, Tobias Wistuba, and Andreas Mayr. 2020. "Estimating Effective Infection Fatality Rates during the Course of the COVID-19 Pandemic in Germany." *arXiv preprint.*

Stevenson, Megan, and Jennifer L. Doleac. 2019. "Algorithmic Risk Assessment in the Hands of Humans." *Available at SSRN:* https://ssrn.com/abstract=3489440.

Stigler, George J. 1961. "The Economics of Information." *Journal of Political Economy* 69 (3): 213–225.

Strunz, Eva. 2013. *Kita vor Ort. Betreuungsatlas auf Ebene der Jugendamtsbezirke 2011.* Technical Report. Deutsches Jugendinstitut e.V. und Dortmunder Arbeitsstelle Kinder- und Jugendhilfestatistik.

Thefeld, Wolfgang, Heribert Stolzenberg, and Bärbel-Maria Bellach. 1999. "Bundes-Gesundheitssurvey: Response, Zusammensetzung der Teilnehmer und Non-Responder-Analyse." *Gesundheitswesen* 61 (Sonderheft 2): 57–61.

van den Berg, Gerald, and Bettina M. Siflinger. 2020. "The Effects of Day Care on Health During Childhood: Evidence by Age." *CEPR Discussion Paper,* no. DP15036.

van Gaal, Luc F., Ilse L. Mertens, and Christophe E. de Block. 2006. "Mechanisms linking Obesity with Cardiovascular Disease." *Nature* 444 (7121): 875–880.

Viner, Russell M., Oliver T. Mytton, Chris Bonell, G. J. Melendez-Torres, Joseph Ward, Lee Hudson, Claire Waddington, James Thomas, Simon Russell, Fiona van der Klis, Archana Koirala, Shamez Ladhani, Jasmina Panovska-Griffiths, Nicholas G. Davies, Robert Booy, and Rosalind M. Eggo. 2020. "Susceptibility to SARS-CoV-2 Infection Among Children and Adolescents Compared With Adults: A Systematic Review and Meta-Analysis." *JAMA Pediatrics* 175 (2): 143–156.

Viscusi, W. Kip, and Joseph E. Aldy. 2003. "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World." *Journal of Risk and Uncertainty* 27 (1): 5–76.

Weil, David N. 2007. "Accounting for the Effect of Health on Economic Growth." *The Quarterly Journal of Economics* 122 (3): 1265–1306.

Winter, Joachim, and Amelie Wuppermann. 2014. "Do They Know What is at Risk? Health Risk Perception among the Obese." *Health Economics* 23 (5): 564–585.

Wisdom, Jessica, Julie S. Downs, and George Loewenstein. 2010. "Promoting Healthy Choices: Information versus Convenience." *American Economic Journal: Applied Economics* 2 (2): 164–78.

Wolf, Ingrid-Katharina, Markus Busch, Michael Lange, Panagiotis Kamtsiuris, Rüdiger Dölle, Almut Richter, Ronny Kuhnert, Thomas Ziese, Hildtraud Knopf, and Christa Scheidt-Nave. 2014. "Mortalitäts-Follow-up der Studie zur Gesundheit Erwachsener in Deutschland (DEGS)." *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz* 57:1331–1337.

Wolf, Ingrid-Katharina, Hildtraud Knopf, Christa Scheidt-Nave, and Bärbel-Maria Kurth. 2012. "Möglichkeiten und Grenzen retrospektiver Todesursachenrecherchen im Rahmen bundesweiter epidemiologischer Studien." *Bundesgesundheitsblatt-Gesundheitsforschung-Gesundheitsschutz* 55:431–435.

World Health Organization. 2020a. *Coronavirus Disease 2019 (COVID-19): Situation Report–46.* Accessed March 5, 2021. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200306-sitrep-46-covid-19.pdf?sfvrsn=96b04adf_4.

———. 2020b. *Influenza Vaccination.* https://www.euro.who.int/en/health-topics/communicable-diseases/influenza/vaccination.

———. 2021. *Coronavirus Disease (COVID-19): Similarities and Differences with Influenza.* Accessed March 5, 2021. https://www.who.int/news-room/q-a-detail/coronavirus-disease-covid-19-similarities-and-differences-with-influenza.

Wu, Zunyou, and Jennifer M. McGoogan. 2020. "Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention." *Journal of the American Medical Association* 323 (13): 1239.

Zentralinstitut für die Kassenärztliche Versorgung in Deutschland. 2021a. *Influenza-Impfraten bei Patienten über 60 Jahre - Basisbericht 2007/2008.* Accessed March 7, 2021. https://www.versorgungsatlas.de/themen/alle-analysen-nach-datum-sortiert/?tab=2&uid=2.

———. 2021b. *Trendanalyse zur Entwicklung der saisonalen Influenzastandardimpfraten zwischen 2009/2010 bis 2013/2014.* Accessed March 7, 2021. https://www.versorgungsatlas.de/themen/alle-analysen-nach-datum-sortiert/?tab=2&uid=68.

Zhao, Yu-miao, Yao-min Shang, Wen-bin Song, Qing-quan Li, Hua Xie, Qin-fu Xu, Jun-li Jia, Li-ming Li, Hong-li Mao, Xiu-man Zhou, Hong Luo, Yan-feng Gao, and Ai-guo Xu. 2020. "Follow-up Study of the Pulmonary Function and Related Physiological Characteristics of COVID-19 Survivors Three Months after Recovery." *EClinicalMedicine* 25:100463.

Zhou, Fei, Ting Yu, Ronghui Du, Guohui Fan, Ying Liu, Zhibo Liu, Jie Xiang, Yeming Wang, Bin Song, Xiaoying Gu, Lulu Guan, Yuan Wei, Hui Li, Xudong Wu, Jiuyang Xu, Shengjin Tu, Yi Zhang, Hua Chen, and Bin Cao. 2020. "Clinical Course and Risk Factors for Mortality of Adult Inpatients with COVID-19 in Wuhan, China: A Retrospective Cohort Study." *The Lancet* 395 (10229): 1054–1062.

Eidesstattliche Versicherung

Ich versichere hiermit eidesstattlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie mir gegebene Anregungen sind als solche kenntlich gemacht.

Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht. Sofern ein Teil der Arbeit aus bereits veröffentlichten Papers besteht, habe ich dies ausdrücklich angegeben.

München, den 11. März 2021

Corinna Hartung