Three Essays in Applied Microeconomics

Inaugural-Dissertation

zur Erlangung des Grades eines Doktors
der Wirtschafts- und Gesellschaftswissenschaften

durch

die Rechts- und Staatswissenschaftliche Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Klara Maria Röhrl

aus Nürnberg

Bonn

2022

Dekan:	Prof. Dr. Jürgen von Hagen
Erstreferentin:	Prof. Dr. Pia Pinger
Zweitreferent:	Prof. Dr. Hans-Martin von Gaudecker
Tag der mündlichen Prüfung:	28. Januar 2022

Acknowledgements

Firstly, I thank my advisors Pia Pinger and Hans-Martin von Gaudecker for their guidance. Discussing research with them has always sharpened my view for what is important. Also, I am very grateful for the liberty they have granted me in everything. Furthermore, I want to thank Thomas Dohmen for helpful discussions on early research ideas, his contagious passion for research and for being part of my dissertation committee.

I am extremely grateful to Janoś Gabler for all his support throughout the years: For being an excellent programming teacher, for showing me where I can still grow and always encouraging me to push my limits and strive for excellence, for helping me take the bird's eye view when I felt stuck and for always being an incredibly dependable friend.

Further thanks goes to the Bonn Graduate School of Economics and the following institutions for all the financial, administrative and institutional support they have provided me over the years: the Collaborative Research Center Transregio 224, briq and the Institute for Applied Microeconomics. Special thanks goes to IZA for their support to communicate research results to the media, policy makers and the general public. I also gratefully acknowledge funding by the Google Cloud CoViD-19 research credits program.

The ph.d. is a challenging time and it can be difficult to maintain perspective and balance during this phase of growth and obstacles. I am deeply grateful to the jazz choir of the university and all the friends I have found there for being an anchor all these years. They have been a family to me and a great source of joy and balance. Going to rehearsals and giving concerts with them always allowed me to reconnect and replenish my energy.

Lastly, I want to thank my parents. You raised me to have all the skills to tackle this big endeavor. I am forever grateful for all the opportunities you have given me and all the ones that are yet to come as the fruits of the love, support and lessons you have sown in me.

Contents

List of Figures

List of Tables

Introduction

Policies shape our lives from the cradle to the grave. It starts with the prenatal care we get and the parental leave our parents are granted. From there it continues with the education we receive, the taxes we pay and the health care we have access to – all the way to the types of burials we can choose from.

Just like the circumstances we are born into, many policies shape our lives in profound ways. However, in contrast to these circumstances, policies are directly under a society's control. This makes them powerful tools for reaching societal goals, such as equality of opportunity or overcoming challenges that require coordination.

At the same time, choosing a policy to reach a particular goal is a difficult task. Many policies can be used to move towards a single goal but their effects can be difficult to predict and can have widely varying costs. Thus, it is no wonder that the evaluation of public policies and predicting the effects of alternative policies is an important focus of economics (Heckman and Vytlacil, 2005).

This dissertation contributes to two fields where policies play an important role: Chapter 1 adds to the literature on equality of opportunity. It reveals that the current playing field is far from level when it comes to attaining happiness for children from different socio-economic backgrounds. Using an experiment, it also shows that policies that improve the social environment of disadvantaged children can correct this inequality. Chapters 2 and 3 contribute to the evaluation of epidemic containment policies. Chapter 2 proposes a model that is designed to evaluate a variety of typical policies to contain CoViD-19 and Chapter 3 applies the model to evaluate the role that vaccinations and rapid tests played in Germany's CoViD-19 pandemic in the spring of 2021.

The first chapter, **"Inequality and the Pursuit of Happiness: How Poor Children Grow Unhappy and What We Can Do About It"**, studies the formation of life satisfaction. Life satisfaction has become an accepted measure of individual well-being that has been shown to have a sizable, stable, trait-like component in adulthood (Diener, 2009, pp. 75-102). However, there are important gaps in our knowledge on how life satisfaction develops during childhood when the trait-like component is likely to emerge. We draw on a unique household panel to report several key features of life satisfaction during childhood and adolescence:

We find that life satisfaction becomes increasingly stable and reaches the stability usually reported for adults by age 15. Children from disadvantaged families experience a much steeper decline than their more advantaged peers leading to a sizable gap. This divide seems to persist over the entire life course when looking at current adult surveys. Using a randomized mentoring intervention we show that this gap is preventable. Treated children suffer a much smaller decline than their peers from the control group. Instead, their trajectories are more similar to those of the advantaged children. This evidence shows that the socio-economic gap in life satisfaction is likely preventable and that interventions in childhood can have persistent effects on life satisfaction.

The second chapter, **"People Meet People – An Epidemiological Model for Evaluating Non-Pharmaceutical Interventions"** presents an agent-based epidemiological model that is designed to evaluate and predict the effect of non-pharmaceutical interventions (NPIs). Most epidemiological models lack the detailed representation of human meeting patterns to credibly predict the effects of the diverse and nuanced policy measures that governments worldwide have been adopting to contain the spread of CoViD-19. We build on state-of-the-art agent-based simulation models, increasing detail and realism in key areas to make the model (1) suitable to predict the effect of a variety of policies, (2) able to use a range of data sources to calibrate most of the model's parameters and (3) capable to fit long running time series.

For this our model has three important features:
Firstly, our population of agents shares many features with human populations. Individuals are part of households and work teams, have differing degrees to which they can work from home, can react to events (such as a symptomatic household member) and their age and job influences when they are vaccinated. These features allow us to calibrate many population parameters from surveys and other empirical data.
Secondly, we model contacts in great detail. This means that many implemented policies can be translated directly into our model from decrees, laws and empirical data. It also enables us to add policies of interest directly into the model to predict their effects.
Thirdly, our model includes a detailed representation of the case detection process: agents can seek rapid and PCR tests in response to information and, as a result, the share of cases that is detected varies between age groups and over time. This is essential to fit empirical time series that include periods where testing capacities were rationed and phases where rapid testing became available.

Thanks to the many parameters that can be calibrated and a computationally tractable two stage matching algorithm, it is computationally feasible to estimate the parameters that cannot be calibrated. That our model has been estimated on nine months of data and is able to achieve an excellent fit on the whole time period with a plateau and two waves for an entire country is what sets it most apart from similar models that usually focus on only showcasing mechanisms such as Aleta et al. (2020) or are at most calibrated to much smaller populations and much shorter time frames (Abueg et al., 2021).

The third chapter, **"How To Beat SARS-CoV-2? The Role of Rapid Tests, Vaccinations, and NPIs to Contain CoViD-19"** uses the model from Chapter 2 to shed light on the role that vaccinations, rapid tests and seasonality played in the decline in CoViD-19 cases in Germany in the spring of 2021. Especially the role of rapid tests is an important subject of scientific inquiry as the role they could play to manage the pandemic was contended. Furthermore, emerging infectious diseases are becoming more common and are expected to increase (Jones et al., 2008; Dobson et al., 2020). Thus, an evaluation of the German rapid test strategy is important to inform the response to possible future pandemics – including possible immune escape variants of SARS-CoV-2 – whose pathogens are likely to share characteristics with SARS-CoV-2.

We calibrate and fit the model detailed in Chapter 2 to German data from August 2020 to June 2021. This very long time frame with two waves allows us to fit the infection probabilities of different contact types and the stringency of some policies. The remaining parameters, such as the home office share over time and the number of pre-pandemic contacts of each contact type, are calibrated from other sources.

We find that during a period in which vaccination rates rose from 5% to 40%, seasonality and rapid testing had the largest effect on reducing infection numbers. Vaccinations contribute only 16% to the difference caused by vaccinations, rapid tests and seasonality jointly. Seasonality contributes 43% of the cumulative difference. A similar-sized effect, 42%, is attributed to rapid testing. This is especially impressive as tests were never mandatory for employees; even in late April only 30% of individuals reported to test themselves at least once per week.

Lastly, we show that stringent testing can be a cheaper alternative to the two most contentious NPIs: work from home mandates and schooling restrictions. The effect of having 10% more workers stay home is small, changing the number of total infections by less than 5%, while mandatory testing at the workplace would have reduced infections by over 20%. Similarly, opening schools at full capacities with two tests per week would have only led to 6% more cases.

The chapters of this thesis illustrate the efficiency improvements that are possible when we rigorously evaluate policies and choose them based on their scientific merit.

This is especially important as inefficient policy responses can be very costly. The CoViD-19 pandemic is a prime example of this: Racine et al. (2021) conclude that anxiety and depression among children and adolescents increased by over 20% during the pandemic. School closures are likely a major contributor to this that could have been avoided had the effects of workplace testing and tests at schools been known.

References

Abueg, Matthew, Robert Hinch, Neo Wu, Luyang Liu, William Probert, Austin Wu, Paul Eastham, Yusef Shafi, Matt Rosencrantz, Michael Dikovsky, Zhao Cheng, Anel Nurtay, Lucie Abeler-Dörner, David Bonsall, Michael V. McConnell, Shawn O'Banion, and Christophe Fraser. 2021. "Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state." *npj Digital Medicine* 4 (1): [3]

Aleta, Alberto, David Martín-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E. Dean, M. Elizabeth Halloran, Ira M. Longini Jr, Stefano Merler, Alex Pentland, Alessandro Vespignani, Esteban Moro, and Yamir Moreno. 2020. "Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19." *Nature Human Behaviour* 4 (9): 964–71. [3]

Diener, Edward. 2009. *The science of well-being: The collected works of Ed Diener*. Vol. 37, Springer. [1]

Dobson, Andrew P., Stuart L. Pimm, Lee Hannah, Les Kaufman, Jorge A. Ahumada, Amy W. Ando, Aaron Bernstein, Jonah Busch, Peter Daszak, Jens Engelmann, Margaret F. Kinnaird, Binbin V. Li, Ted Loch-Temzelides, Thomas Lovejoy, Katarzyna Nowak, Patrick R. Roehrdanz, and Mariana M. Vale. 2020. "Ecology and economics for pandemic prevention." *Science* 369 (6502): 379–81. [3]

Heckman, James J., and Edward Vytlacil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73 (3): 669–738. [1]

Jones, Kate E., Nikkita G. Patel, Marc A. Levy, Adam Storeygard, Deborah Balk, John L. Gittleman, and Peter Daszak. 2008. "Global trends in emerging infectious diseases." *Nature* 451 (7181): 990–93. [3]

Racine, Nicole, Brae Anne McArthur, Jessica E. Cooke, Rachel Eirich, Jenney Zhu, and Sheri Madigan. 2021. "Global Prevalence of Depressive and Anxiety Symptoms in Children and Adolescents During COVID-19." *JAMA Pediatrics* 175 (11): 1142. [4]

Chapter 1

Inequality and the Pursuit of Happiness: How Poor Children Grow Unhappy and What We Can Do About It

Joint with Armin Falk, Fabian Kosse and Pia Pinger

1.1 Introduction

Life satisfaction is an important measure of individual well-being. Not only is it a nearly universal parenting goal, it also predicts several important later outcomes and behaviors, such as physical aggression and resilience in adolescents (Suldo and Huebner, 2004; Proctor, Linley, and Maltby, 2008; Proctor, Linley, and Maltby, 2018) and longevity in adults (Diener, Inglehart, and Tay, 2013).

In addition, youths' life satisfaction deserves special attention. Firstly, because they are a particularly vulnerable group – their lives are still largely determined by others and adolescence is considered a particularly challenging phase of life (Shek and Liu, 2014; Proctor, Linley, and Maltby, 2018). Secondly, life satisfaction has a significant trait-like component, making it quite stable during adulthood (Fujita and Diener (2005), Lucas and Donnellan (2007) and Headey and Muffels (2018)). It seems likely that this trait-like component emerges during childhood and adolescence. Consequently, policies during childhood and adolescence might be a particular effective way to increase a population's life satisfaction in the long term if they can influence the development of that trait-like component.

Yet, longitudinal and causal evidence on the formation of life satisfaction during childhood and adolescence is scarce (Proctor, Linley, and Maltby, 2008; Shek and

Liu, 2014; Casas and González-Carrasco, 2018; Proctor, Linley, and Maltby, 2018).[1] Therefore no studies on the emergence of that trait-like component exist to our knowledge. And most long-run evaluations of childhood interventions focus on later earnings and educational outcomes rather than the participants' life satisfaction (Garces, Thomas, and Currie, 2002; Belfield et al., 2006).

Following nearly 500 children over seven years, this chapter shows how life satisfaction stabilizes during adolescence and that a mentoring intervention in mid childhood can prevent the socio-economic gap that emerges during adolescence. This longitudinal design shows that life satisfaction is still rather volatile before puberty but reaches the stability found in adulthood by age 15. It also confirms the marked decline with age commonly found in cross-sectional studies (Jewell and Kambhampati, 2014; Shek and Liu, 2014; Casas and González-Carrasco, 2018). The longitudinal nature of our data reveals that this decline is neither driven by the happiest children nor by children who already report low happiness at baseline. Furthermore, this chapter contributes to the discussion on the inconsistent correlations between socio-economic factors and life satisfaction (Frijters, Johnston, and Shields, 2014; Shek and Liu, 2014; Schnitzlein and Wunder, 2016; Proctor, Linley, and Maltby, 2018). Following children from families of different socio-economic status (SES), it appears that life satisfaction is unaffected by socio-economic background when children are eight. However, a sizable SES gap of over 0.4 points on a zero to ten scale emerges as children grow older. Lastly, exploiting a randomized control trial in the data shows that long-run life satisfaction can be affected through the social environment during childhood and the SES gap can be prevented to a large extent. Treated children in our data follow a different trajectory than their untreated disadvantaged peers and report higher life satisfaction years after the intervention ended.

This chapter proceeds as follows: Section 1.2 provides an overview over what is known about children's life satisfaction. Section 1.3 presents our data set and its relevant variables. The following section presents the results: It starts by documenting an increasing stability of children's life satisfaction and confirming the decline as children age (Section 1.4.1). Next, Section 1.4.2 shows that a large socio-economic gradient emerges during adolescence. Lastly, the chapter investigates whether the mentoring intervention that a subset of the low SES children were randomized to can improve children's long-run life satisfaction in Section 1.4.3. Section 1.5 concludes.

1. Jewell and Kambhampati (2014), Shek and Liu (2014), and Jung and Choi (2017) are notable exceptions but in comparison to us, observe less points in time.

1.2 Literature Review

Over the past decades life satisfaction has become a well-accepted measure of individual well-being (Diener, Inglehart, and Tay, 2013). This is not just the case for adults but also for adolescents (Jovanović, 2016) and children (Gilman and Huebner, 2003; Gadermann, Guhn, and Zumbo, 2010). In most cases a single item scale has shown similar desirable psychometric properties to multi item scales (Cheung and Lucas, 2014; Jovanović, 2016; Mazur et al., 2018).

Life satisfaction can be seen as an important outcome in itself. Moreover, it is an antecedent of many other desirable outcomes in both adults and adolescents. For example, higher life satisfaction predicts better physical health in adults, including increased longevity and a lower risk of disability, chronic disease and dementia (Koivumaa-Honkanen et al., 2004; Lyyra et al., 2006; Chida and Steptoe, 2008; Siahpush, Spittal, and Singh, 2008; Feller et al., 2013; Rauma et al., 2014; Peitsch et al., 2016).[2] In adolescents, life satisfaction is associated with a reduced risk of later violent and delinquent behavior (Jung and Choi, 2017; Hanniball et al., 2018), externalizing and internalizing behaviors and peer victimization (Suldo and Huebner, 2004; Haranin, Huebner, and Suldo, 2007; Martin, Huebner, and Valois, 2008).

Regarding its malleability, research of the last decades has shown that life satisfaction has a considerable constant component: 30-45% of adult life satisfaction appears to be fixed (Lucas and Donnellan, 2007; Frijters, Johnston, and Shields, 2014) and twin studies suggest a sizable genetic component (Neve et al., 2012). Nevertheless, important life events can have mid to long-term effects on life satisfaction (Lucas, 2007) and there is a bi-directional longitudinal association between personality and life satisfaction (e.g. Boyce, Wood, and Powdthavee (2012), Specht, Egloff, and Schmukle (2012), and Hounkpatin et al. (2018)).

This suggests that childhood could be a sensitive period for life satisfaction. Firstly, because it seems likely that the non-genetic part of the trait-like component observed in adulthood emerges during childhood. Secondly, because personality is more malleable during childhood before it becomes stabler during adolescence and young adulthood (Caspi, Roberts, and Shiner, 2005; Klimstra et al., 2009; Costa, McCrae, and Löckenhoff, 2019).

Additionally, life satisfaction during adolescence deserves special attention because it is well established that life satisfaction falls during this time (e.g. Jewell and Kambhampati (2014), Shek and Liu (2014), and Casas and González-Carrasco (2018)).

2. For a paper on potential causal channels, see Smyth et al. (2016).

Despite its widespread occurrence, many questions, such as whether the decline is driven by a catch-up of more "naïve" children, remain unanswered.

Awareness for the fragility of adolescents and the potential to foster their well-being through interventions during childhood and adolescence has been recognized by the literature (Seligman et al., 2009; Proctor and Linley, 2014; Shek and Liu, 2014; Proctor, Linley, and Maltby, 2018) and some evaluations of interventions such as Ruini et al. (2009) and Proctor, Tsukayama, et al. (2011) do exist. However, long-term evaluations of interventions on life satisfaction are extremely rare, both for adults (Weiss, Westerhof, and Bohlmeijer, 2016) and youths. The most long-term evaluation of an intervention for youths that we are aware of is Gillham, Hamilton, et al. (2006) with a two year follow-up. Thus, the time frames of interventions have been insufficient to detect long-term changes in life satisfaction.

Regarding potential ways to intervene against the decline, two correlates of youths' life satisfaction stand out: relationship quality and personality. Firstly, supportive relationships are strongly correlated with children's and adolescents' well-being – whether it be measured in the form of family time, an authoritative parenting style, supportive parent-child relationships, or other adult support (Paxton et al., 2006; Proctor and Linley, 2014; Shek and Liu, 2014). Furthermore, the perceived quality of family relationships decreases during adolescence (Goede, Branje, and Meeus, 2008). Secondly, several personality factors and character strengths are strongly associated with life satisfaction, in both youths (Gillham, Adams-Deutsch, et al., 2011; Proctor and Linley, 2014; Suldo, Minch, and Hearon, 2014; Weber and Huebner, 2015; Blanca et al., 2017) and adults (Lounsbury et al., 2004; Peterson et al., 2007; Proyer et al., 2011; Boyce, Wood, and Powdthavee, 2012).

Lastly, it is important to note that the role of socio-economic background on children's life satisfaction is a contentious question in the literature on children's life satisfaction (Cho, 2018; Proctor, Linley, and Maltby, 2018). For example, Jewell and Kambhampati (2014) found no significant effect of household income on life satisfaction and Layard et al. (2014) found that family income accounts only for a very small fraction of the variance in life satisfaction. On the other hand, Frijters, Johnston, and Shields (2014) showed that having a father in the lowest social class substantively lowered life satisfaction and Shek and Liu (2014) also found that economic disadvantage lowered life satisfaction.

1.3 Data and Measures

This chapter draws on the briq family panel, which has been following over 500 children – some of which were randomized to a mentoring program – for over seven years. The dataset is unique in several ways: Firstly, the panel has seven waves, i.e. participating families were interviewed at up to seven points in time, with approximately one year between interviews. Thus, we observe each child's life satisfaction up to seven times from mid childhood to mid adolescence. Secondly, the detailed yearly interviews with both children and caregivers yield a rich description of the children's developing personality and the environment they grow up in. Thirdly, the study design includes a subset of low SES children that were randomized to participate in a mentoring intervention which provides us with exogenous variation in the social environment in the low SES group. For a detailed description of the dataset and intervention, see Kosse et al. (2020, Section 2).

Throughout our analysis, children's self-reported life satisfaction is used as outcome variable which was surveyed using age-appropriate questionnaire items. In the first two waves (child ages 7-10), children answered an age-adjusted life satisfaction question using a smiley scale. Starting from wave three (child ages 11-15), children were asked the standard overall life satisfaction question that is also contained, in the World Value Survey and the German socio-economic panel (SOEP) (Bjørnskov, 2010; Pagán-Rodríguez, 2010).[3,4,5]

There are two issues that arise from the subjectivity and ordinality of life satisfaction. Firstly, there are likely individual differences in the interpretation of the scale as life satisfaction is subjective. The long-running nature of our panel allows us to control for the individual location of the scale of each individual (and constant factors in their environment) by controlling for baseline life satisfaction.

Secondly, life satisfaction is ordinal which implies that any encoding that preserves the order of the variable is a valid representation. The variable is encoded using the same numeric values zero to ten with which it was presented to the respondents in the later waves of the survey. Other encodings were only used to verify a basic robustness of our main results.[6] As Bond and Lang (2019) have shown the sign of differences between group means are only robust to any valid numeric encoding if the distribution of one group first order statistically dominates the other. Appendix 1.E shows the distributions of our main groups of interest and argues that only very particular encodings would lead to a reversal of our reported effects.

Families were categorized to have high or low socio-economic status (SES) based on their income, education and single parent status. Our definition is child centric

3. Single-item life satisfaction measures are very valid as shown by Cheung and Lucas (2014).

4. For a discussion on how these two measures are made comparable over time, see Appendix 1.A.

5. For evidence that children as young as eight understand and can provide thoughtful answers on their life satisfaction, see Gadermann, Guhn, and Zumbo (2010).

6. Results available upon request.

and aims to capture differences in resources available to the child from the family level. A family is considered to be low SES if it fulfills any of the following criteria at baseline: One, the family's income lies below the 30th percentile of the German household income distribution. Two, neither of the parents have obtained a high school degree that qualifies for university studies. Three, the child lives with a single parent. Households that did not fulfill any of these criteria at baseline were classified as high SES.

The briq family panel includes a subset of low SES children that were randomized to participate in the well-established non-profit mentoring program called "Balu und Du". In this program, children meet weekly with a mentor for the duration of one year. Mentors are mostly university students who usually report high life satisfaction. All mentors are volunteers, motivated to enrich the life of a disadvantaged child.[7] Conceptually, the idea of the program is twofold. Firstly, the program aims to encourage informal learning through joint activities. Secondly, children experience an additional adult, who is well-adjusted and treats the child in a responsive, attuned and encouraging way. As a result the mentor serves as friend and role model. Interviews and the mentoring intervention were scheduled such that the first wave is the pre-treatment interview, the second interview is the post-treatment interview, followed by a series of yearly interviews. This chapter includes six post-treatment interviews.

Thus, there are three main groups: a High SES Control group, a Low SES Control group and a Low SES Treatment group. All three comparisons are highly interesting. The two control groups show the role socio-economic background plays in children's and adolescents' life satisfaction. Comparing the treatment group to the Low SES Control group yields the treatment effect of our intervention and the comparison of the Treatment to the high SES Control group gives a sense to which degree the intervention is able to level the playing field.

1.4 Results

This section presents our main results. It first provides evidence for the emergence of a trait-like component by showing that life satisfaction becomes increasingly stable during adolescence. While life satisfaction becomes more stable over time, the reported level of life satisfaction declines as children age. This trend affects all children irrespective of the happiness they reported when young and thus does not appear to be the result of some form of catch-up of more "naïve" children with more "mature"

7. "Balu and Du" translates to Baloo and You. Baloo refers to Baloo the bear from *The Jungle Book*. Prior evidence shows that participation enhances children's prosociality (Kosse et al. (2020)).

peers. Next, we show that the age decline is much more pronounced for disadvantaged children. We provide suggestive evidence that the inequality that emerges is rarely overcome in adulthood and can be expected to be somewhat permanent by showing that a similar sized parental SES gap is also present in the adult German population over the entire life span. Having shown that the SES gap is problematic from an equality of opportunity perspective, the section continues with the evaluation of the randomized intervention in the briq family panel and finds that it affected the life satisfaction trajectory of the treated children. This effect is so pronounced that the treated low SES children and their high SES peers are statistically indistinguishable. Lastly, the section investigates possible treatment effect heterogeneities. The results on treatment heterogeneity are mixed and do not provide conclusive evidence.

Due to superior properties in finite samples, standard errors are bootstrapped unless stated otherwise. To account for different interpretations of the life satisfaction scale regressions usually control for baseline life satisfaction. Using the balanced sample or unbalanced sample[8] hardly changes the results. This is unsurprising as there is little evidence for selective attrition (see Appendix 1.B).

1.4.1 Life Satisfaction Becomes More Stable and Declines During Childhood and Adolescence

In line with the emergence of a trait-like component, life satisfaction becomes increasingly stable during adolescence. Figure 1.4.1 shows the correlations of life satisfaction across waves and compares them to the correlation of life satisfaction between adjacent waves of adults in the German Socio-Economic Panel (SOEP, Goebel et al. (2018) and Liebig et al. (2019)). Life satisfaction starts out quite volatile, with a correlation of less than 0.25 between the first and the second wave and increases to the level of 0.5 found in adults by the age of 14.[9]

8. The balanced sample only includes children that provided all required variables in all waves. The unbalanced sample also includes children that dropped out of the study at some point or did not always provide all variables.

9. The 2nd and 3rd wave were further apart than the other waves. Thus, it seems likely that life satisfaction was becoming more stable and the lack of an increase in the second column is due to the longer interval between interviews.

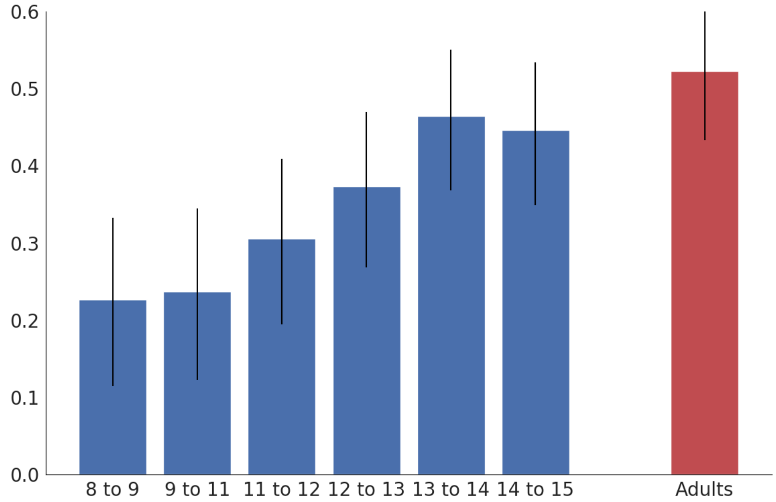


Figure 1.4.1. Development of the Correlation Of Life Satisfaction Between Waves Over Time

Note: The figure shows the correlation of children's life satisfaction between waves over time. The red bar on the right shows the same correlation found in German adults (data from the German Socio-Economic Panel, Goebel et al. (2018) and Liebig et al. (2019)). Whiskers represent 95% confidence intervals.

While life satisfaction becomes more stable, the reported levels decrease markedly in our sample (Figure 1.4.2). This is in line with earlier studies that also reported such declines (Jewell and Kambhampati, 2014; Shek and Liu, 2014; Casas and González-Carrasco, 2018). Our longitudinal dataset is able to show that this decline is not the result of a catching-up process but affects children irrespective of their baseline happiness.

All children start with high levels of life satisfaction. In the first wave, 73% in the Low SES Control Group and 65% in the High SES Control Group report the highest or second highest value. For comparison, less than one in four German adults reported life satisfaction this high in 2016 (Goebel et al. (2018) and Liebig et al. (2019)).

Life satisfaction falls from this extraordinarily high level by on average 1.2 points until age 14. This is more than one standard deviation in terms of the initial life satisfaction distribution. Relating it to event studies in adults, this drop is larger than the life satisfaction drop associated with disability and similar to the drop associated with the death of a spouse (Lucas, 2007).

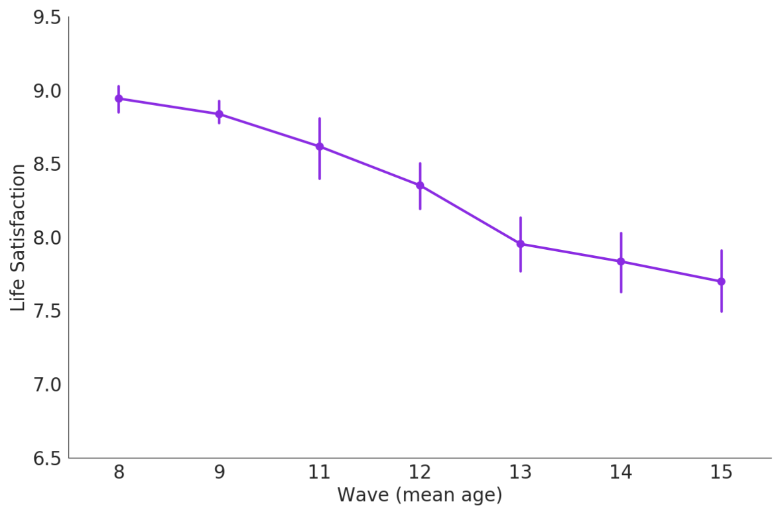


Figure 1.4.2. Development of Life Satisfaction Over Time

Note: Points show mean life satisfaction for each wave with the mean age during the wave's interview on the x axis. Whiskers show 95% confidence intervals.

The panel structure allows to investigate how this decline is distributed. Is it that "naïve" children catch up with more "mature" peers? Do struggling children turn into desolate adolescents when the challenges of puberty arise? Figure 1.4.3 and Table 1.4.1 show that neither is the case. Instead, children who reported the highest, second highest or a lower level of life satisfaction at baseline all show very similar trends over the years. The three groups' age trends are estimated to be -0.16, -0.21 and -0.19 points per year, respectively. These are roughly similar in absolute terms and also statistically indistinguishable (p=0.40, N=3015, F-Test). Thus, the decline appears to be a universal phenomenon and not driven by children who displayed low or high levels of life satisfaction when young.

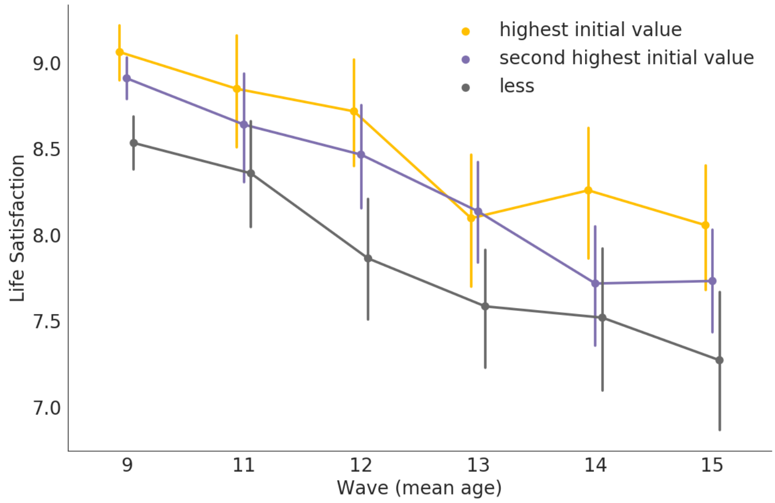


Figure 1.4.3. Development of Life Satisfaction by Baseline Life Satisfaction During Adolescence

Note: The figure shows mean life satisfaction from the second wave onwards for children who reported the highest, second highest or a lower value in the first wave. Points show mean life satisfaction of each group in each wave. Whiskers show 95% confidence intervals. There are between 336 and 426 observations per wave. Life satisfaction is measured on a 0 to 10 point scale.

Table 1.4.1. Age Trends in Life Satisfaction by Reported Baseline Life Satisfaction

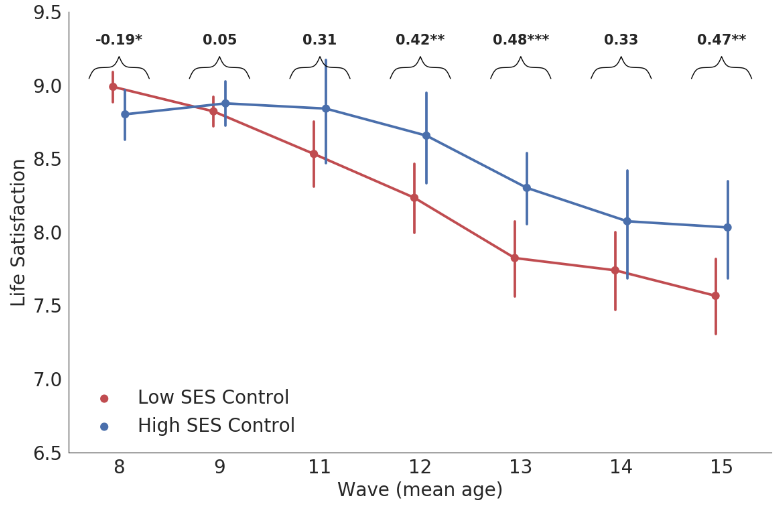
	(1)
	Life Satisfaction
Low Baseline Life Satisfaction	-0.28
2nd Highest Baseline Life Satisfaction	0.27
Age	-0.16***
Low Baseline Life Satisfaction x Age	-0.03
2nd Highest Baseline Life Satisfaction x Age	-0.05
Resident in Cologne	0.04
Constant	10.53***
Observations	3015

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

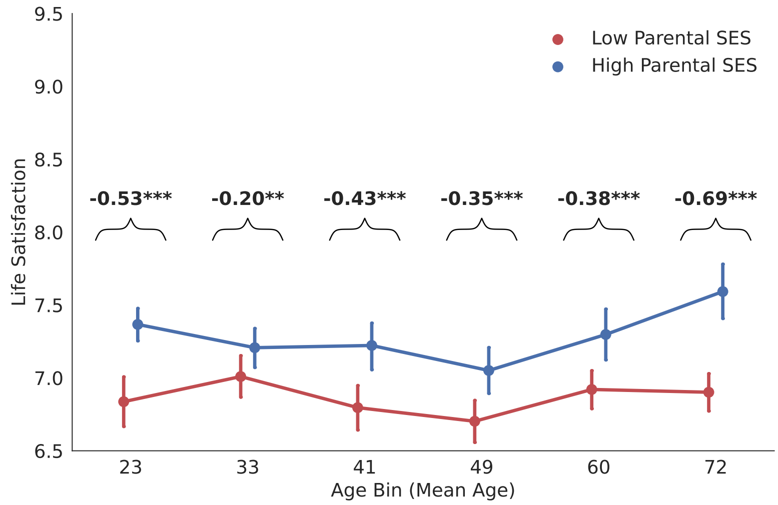
Note: Coefficients from an unbalanced random effects regression model controlling for place of residence (Cologne or Bonn). Intercepts and age trends are estimated for children that reported the highest, second highest or a lower life satisfaction at baseline, respectively. All three groups have very similar trends (p=0.40). Statistical significance is denoted by: *: p<0.1, **: p<0.05, ***: p<0.01.

1.4.2 The Emergence of a Socio-Economic Gradient in Life Satisfaction

We now show that a sizable SES gap emerges during our study period and that this gap is comparable to the gap present in the German adult population.



(a) Children's Life Satisfaction by SES Over Time



(b) Life Satisfaction During Adulthood by Parental SES

Figure 1.4.4. Evidence for the Emergence of a Life Long SES Gap in Life Satisfaction

Note: The upper figure shows mean reported life satisfaction in the Low and High SES Control group in each wave during childhood and adolescence in the briq family panel. The lower figure shows the same life satisfaction variable but among adults that grew up in the top and bottom quartiles of the parental SES distribution in a cross-section of the German socio-economic panel (data source: Wagner et al. (2011), N=6255). In both cases, life satisfaction is measured on a 0 to 10 point scale. Whiskers show 95% confidence intervals. Mean differences between high and low SES groups are displayed above each age bin. Stars represent statistical significance of two-sided t-tests: *: p<0.1, **: p<0.05, ***: p<0.01.

As can be seen in Figure 1.4.4a the two groups start out very similar but disadvantaged children suffer a much steeper decline than their more privileged peers. At age eight, the Low SES Control and High SES Control groups display similar levels of life satisfaction with the High SES Control children even reporting slightly lower levels of life satisfaction.[10] However, they follow very different trajectories. For children in the High SES Control group life satisfaction declines at a rate of 0.13 points per year. For children in the Low SES Control group that rate is 0.21 (p=0.01, N=3698). By age 15, the more privileged children report 0.44 points higher life satisfaction than their disadvantaged peers (p=0.03, two-sided t-test). To draw again a comparison to event studies in adults: High and Low SES adolescents in our sample differ as much in their life satisfaction as adults four years before their divorce compared to freshly divorced adults (Lucas, 2007).

The clear difference in trends and the gradual emergence of a SES gap could be one factor in explaining the contended role of socio-economic status in the literature on children's life satisfaction (Frijters, Johnston, and Shields, 2014; Jewell and Kambhampati, 2014; Layard et al., 2014; Shek and Liu, 2014; Cho, 2018; Proctor, Linley, and Maltby, 2018). Since our measure of socio-economic status is an aggregate of three different dimensions, Appendix 1.C.1 shows the degree to which the reported SES gap is robust to different SES classifications.

An important question is whether this gap can be expected to close in early adulthood. This could happen because of two reasons: Firstly, as adolescents become increasingly self-reliant and make their own choices the disadvantaged youths may be able to catch up with their more privileged peers. Secondly, there may be habituation, i.e. adolescents may become accustomed and grow content with their social status leading to a leveling of life satisfaction.

To investigate this we compare our results with a cross-section of the adults in the German socio-economic panel (SOEP), also splitting them by their parents' SES.[11] Figure 1.4.4b strongly indicates that the gap in life satisfaction, that opens up during

10. The means are 9.0 (Low SES Control), 8.8 (High SES Control). The difference is 0.19 (p=0.07, two-sided t-test).

11. The SOEP measure of adults' parental SES is different from the one for children in the briq family panel. Instead of education, income and single-parent status (which have all changed substantially over the last 60 years), the adults' parental SES in the SOEP uses the "International Socio-Economic Index of Occupational Status (ISEI)" developed by Ganzeboom, De Graaf, and Treiman (1992). The ISEI is based on a categorization of the International Standard Classification of Occupations (ISCO) codes, which are available for the parents of most SOEP participants. In order to obtain a binary SES classification comparable to the one in our study, we first aggregate maternal and paternal ISEI to a single parental ISEI. SOEP participants whose parents are in the bottom quartile are classified as "Low Parental SES". SOEP participants whose parents are in the top quartile are classified as "High Parental SES". Only the bottom and top quartile are displayed in the figure. The inner quartiles lie between the bottom and the top quartile.

adolescence, can be expected to persist well into adulthood and beyond. Over the entire life course, life satisfaction is consistently lower among adults who grew up in households in the bottom relative to the top quartile of the parental SES distribution (average difference: 0.42 points, p=0.00, two-sided t-test). In addition, this gap between top and bottom quartile is of similar size as that in the briq family panel.[12]

The substantial and suggestively long-lasting life satisfaction gap between individuals from different socio-economic backgrounds raises important questions about our ability to provide equal opportunities to all children to lead fulfilling lives. While the consequences of socio-economic background on education and health are well documented and whole literatures exist on how different policies affect these socio-economic gaps,[13] there is hardly any awareness for the huge divisions in the less visible dimension of life satisfaction. The size of these gaps show that there could be a lot of room for public policy to increase equality of opportunity in its arguably most fundamental dimension.

1.4.3 Interventions in Childhood Can Affect Long-Run Life Satisfaction and Prevent the SES Gap

This section shows that the mentoring intervention which some of the disadvantaged children in the briq family panel participated in, is able to prevent the socio-economic gap in life satisfaction. For this the Low SES Treatment group is compared to the Low and High SES Control groups. Exploiting the full panel structure allows to estimate separate intercepts and age trends for each of the three groups (see Table 1.4.2). This corroborates the SES gap from the previous section and provides clear evidence for a strong treatment effect: The Low SES Control group loses on average $^1/_5$ of a point per year, while the Low SES Treatment group loses on average only $\approx {}^1/_8$ of a point per year (p=0.00 , N=3698), a 40% reduction. Given this improvement, the High SES Control and Low SES Treatment group have very similar age trends and no gap emerges between them.

Because the SES gap in life satisfaction emerges as children grow older, the treatment effect only becomes apparent during adolescence. In Table 1.C.2 we show that both the SES gap and the treatment effect are also present in cross-sectional regressions. We show that the treatment effect is robust to using life satisfaction of one or both of the last two waves. In addition, we show that the treatment effect is still

12. Brüderl, Kratz, and Bauer (2019) do a similar analysis and also find a strong SES gap in the SOEP and conclude that the gap even increases from adulthood to old age.

13. For example Adler et al. (1993), Mackenbach et al. (2008), and Cesare et al. (2013) on health inequalities and Waltenberg and Vandenberghe (2007), Schütz, Ursprung, and Wößmann (2008), and Doorn, Pop, and Wolbers (2011) for education.

visible when measuring life satisfaction with the Students' Life Satisfaction Score (SLSS, Weber, Ruch, and Huebner (2013)) instead of our single-item question.

Table 1.4.2. Age Trends in Life Satisfaction by SES and Treatment Status

	(1)	(2)	(3)
Age Trend of the Low SES Control Group	-0.21***	-0.20***	-0.22***
Diff. in Trends High to Low SES	0.08***	0.07**	0.10***
Diff. in Trends Treatment to Low SES Control	0.08***	0.07***	0.08***
Sample	Unbalanced	Unbalanced	Balanced
Control for Baseline Life Sat.	No	Yes	Yes
Observations	3703	3698	2950

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows the age trends from a random effects regression model for each of the three treatment categories: Low SES Control, High SES Control and (Low SES) Treatment. The second and third column also control for baseline life satisfaction. The third column only uses the balanced panel. For details on the control variables and the size of the other coefficients, see Table 1.C.1.

Given these impressive long-term effects, we investigated possible channels by looking for treatment heterogeneity in different dimensions – with inconclusive results. The detailed results are shown in Appendix 1.D.

Recall that the mentoring program provides two types of resources that are presumably often lacking in low SES families: Firstly, the child has an additional adult as friend and role model. Mentors are usually highly satisfied with their life, well-adjusted and interact with the children in an attuned, responsive and encouraging manner. Secondly, the child gets to experience many enriching activities. This gives us three suggestive channels:

Firstly, the effect of the program could be due to having a role model of high life satisfaction. Children of caregivers who are less satisfied with their lives might learn behaviors and skills relevant for their later life satisfaction from their mentor that are not shown to them by their caregivers and thus benefit from the intervention especially. Appendix 1.D.1 provides some supportive evidence that children with less satisfied caregivers benefit more from the intervention but the difference between treatment effects is not consistently statistically significant. These are nevertheless encouraging results that life satisfaction is not an inherited trait but can be changed through investments during childhood.

Secondly, experiencing joint activities with a supportive adult may be the channel through which children learn skills or develop character strengths. These are important for their later life satisfaction (Proctor, Tsukayama, et al., 2011; Jewell and Kambhampati, 2014). In that case, children that have few opportunities for such joint activities with their caregiver would be expected to benefit from the intervention especially. However, the data reveals no such pattern. The differences in the treatment effect between children with many or few activities at baseline vary in size, change sign and are never statistically significant and in general much smaller than the treatment effect heterogeneity found by parental life satisfaction. This is robust to two different measures of joint activities (Appendix 1.D.2).

Thirdly, other aspects of having the mentor may foster children's socio-emotional development, allowing them to learn skills that prevent behavioral problems which are associated with lower life satisfaction (Frijters, Johnston, and Shields, 2014; Layard et al., 2014; Shek and Liu, 2014). Children from low SES backgrounds tend to perform worse on the *"Strengths and Difficulties Questionnaire"* (SDQ, Giannakopoulos et al. (2009)), an established measure of children's socio-emotional development (Goodman, 1997) which is included in the briq family panel. In addition, there is some evidence that a mentor may help children improve in these problem areas (Erdem et al., 2016). This, in turn, could lead to higher life satisfaction as children grow older and work towards more long-term goals and confront the emotional and peer challenges associated with puberty.

Layard et al. (2014) found emotional health and conduct to be the most powerful childhood predictors of later life satisfaction. This made the aggregate SDQ score as overall measure of a child's emotional health and the conduct problem score the most promising candidates for treatment heterogeneities. However, the differences in treatment effects across both dimensions are statistically insignificant and vary in both sign and size (see Appendix 1.D.3).

1.5 Conclusion

This chapter has presented some important characteristics of the development of life satisfaction during childhood and adolescence using a panel covering over seven years. Firstly, life satisfaction becomes increasingly stable reaching the stability of adults by age 15 which supports the formation of a trait-like component of life satisfaction. Secondly, an important gap emerges between children from different socio-economic backgrounds during adolescence. Looking at a large sample of German adults it seems likely that this gap will persist to at least some extent well into old age. Thirdly, enriching the social environment of children can have lasting effects on life satisfaction and compensate for other disadvantages: An intense one year mentoring intervention during mid childhood is able to effectively prevent the emergence of the socio-economic gradient in life satisfaction between the treatment group and their high SES peers.

Our evidence aligns with a large literature documenting important socio-economic gradients emerging during childhood in many important dimensions of well-being (Barnett, 1995; Reynolds et al., 2001; Campbell et al., 2014; Gertler et al., 2014). These gaps pose important challenges to providing equality of opportunity to all children. This chapter shows that even in an elusive dimension such as life satisfaction, early and high quality investments in disadvantaged children have large and long-lasting effects.

References

Adler, Nancy E., W. Thomas Boyce, Margaret A. Chesney, Susan Folkman, and S. Leonard Syme. 1993. "Socioeconomic Inequalities in Health: No Easy Solution." *JAMA* 269 (24): 3140–45. [17]

Barnett, W Steven. 1995. "Long-term effects of early childhood programs on cognitive and school outcomes." *future of children*, 25–50. [20]

Belfield, Clive R, Milagros Nores, Steve Barnett, and Lawrence Schweinhart. 2006. "The High/Scope Perry Preschool Program." XLI (1): 162–90. [6]

Bjørnskov, Christian. 2010. "How Comparable are the Gallup World Poll Life Satisfaction Data?" *Journal of Happiness Studies* 11 (1): 41–60. [9]

Blanca, María J., Marta Ferragut, Margarita Ortiz-Tallo, and Rebecca Bendayan. 2017. "Life Satisfaction and Character Strengths in Spanish Early Adolescents." *Journal of Happiness Studies* 19 (5): 1247–60. [8]

Bond, Timothy N., and Kevin Lang. 2019. "The Sad Truth about Happiness Scales." *Journal of Political Economy* 127 (4): 1629–40. [9]

Boyce, Christopher J., Alex M. Wood, and Nattavudh Powdthavee. 2012. "Is Personality Fixed? Personality Changes as Much as "Variable" Economic Factors and More Strongly Predicts Changes to Life Satisfaction." *Social Indicators Research* 111 (1): 287–305. [7, 8]

Brüderl, Josef, Fabian Kratz, and Gerrit Bauer. 2019. "Life course research with panel data: An analysis of the reproduction of social inequality." *Advances in Life Course Research* 41 (9): 100247. [17]

Campbell, Frances, Gabriella Conti, James J. Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan. 2014. "Early Childhood Investments Substantially Boost Adult Health." *Science* 343 (6178): 1478–85. eprint: https://science.sciencemag.org/content/343/6178/1478.full.pdf. [20]

Casas, Ferran, and Mònica González-Carrasco. 2018. "Subjective Well-Being Decreasing With Age: New Research on Children Over 8." *Child Development* 90 (2): 375–94. [6, 7, 12]

Caspi, Avshalom, Brent W. Roberts, and Rebecca L. Shiner. 2005. "Personality Development: Stability and Change." *Annual Review of Psychology* 56 (1): 453–84. [7]

Cesare, Mariachiara Di, Young-Ho Khang, Perviz Asaria, Tony Blakely, Melanie J Cowan, Farshad Farzadfar, Ramiro Guerrero, Nayu Ikeda, Catherine Kyobutungi, Kelias P Msyamboza, Sophal Oum, John W Lynch, Michael G Marmot, and Majid Ezzati. 2013. "Inequalities in non-communicable diseases and effective responses." *Lancet* 381 (9866): 585–97. [17]

Cheung, Felix, and Richard E. Lucas. 2014. "Assessing the validity of single-item life satisfaction measures: results from three large samples." *Quality of Life Research* 23 (10): 2809–18. [7, 9]

Chida, Yoichi, and Andrew Steptoe. 2008. "Positive Psychological Well-Being and Mortality: A Quantitative Review of Prospective Observational Studies." *Psychosomatic Medicine* 70 (7): 741–56. [7]

Cho, Esther Yin-Nei. 2018. "Links between Poverty and Children's Subjective Wellbeing: Examining the Mediating and Moderating Role of Relationships." *Child Indicators Research* 11 (2): 585–607. [8, 16]

Costa, Paul T., Robert R. McCrae, and Corinna E. Löckenhoff. 2019. "Personality Across the Life Span." *Annual Review of Psychology* 70 (1): 423–48. [7]

Diener, Ed, Ronald Inglehart, and Louis Tay. 2013. "Theory and Validity of Life Satisfaction Scales." *Social Indicators Research* 112 (3): 497–527. [5, 7]

Doorn, Majka van, Ioana Pop, and Maarten H. J. Wolbers. 2011. "Intergenerational Transmission of Education Across European Countries and Cohorts." *European Societies* 13 (1): 93–117. [17]

Erdem, Gizem, David L. DuBois, Simon Larose, David De Wit, and Ellen L. Lipman. 2016. "Mentoring Relationships, Positive Development, Youth Emotional and Behavioral Problems: Investigation of a Meditation Model." *Journal of Community Psychology* 44 (4): 464–83. [19]

Feller, Silke, Birgit Teucher, Rudolf Kaaks, Heiner Boeing, and Matthaeus Vigl. 2013. "Life Satisfaction and Risk of Chronic Diseases in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Germany Study." *PLoS ONE* 8 (8): edited by C. Mary Schooling, e73462. [7]

Frijters, Paul, David W. Johnston, and Michael A. Shields. 2014. "Does Childhood Predict Adult Life Satisfaction? Evidence from British Cohort Surveys." *Economic Journal* 124 (580): F688–F719. [6–8, 16, 19]

Fujita, Frank, and Ed Diener. 2005. "Life satisfaction set point: stability and change." *Journal of personality and social psychology* 88 (1): 158. [5]

Gadermann, Anne M., Martin Guhn, and Bruno D. Zumbo. 2010. "Investigating the Substantive Aspect of Construct Validity for the Satisfaction with Life Scale Adapted for Children: A Focus on Cognitive Processes." *Social Indicators Research* 100 (1): 37–60. [7, 9]

Ganzeboom, Harry BG, Paul M De Graaf, and Donald J Treiman. 1992. "A standard international socio-economic index of occupational status." *Social science research* 21 (1): 1–56. [16]

Garces, Eliana, Duncan Thomas, and Janet Currie. 2002. "Longer-Term Effects of Head Start." 92 (4): 999–1012. [6]

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeersch, Susan Walker, Susan M. Chang, and Sally Grantham-McGregor. 2014. "Labor market returns to an early childhood stimulation intervention in Jamaica." *Science* 344 (6187): 998–1001. eprint: https://science.sciencemag.org/content/344/6187/998.full.pdf. [20]

Giannakopoulos, George, Chara Tzavara, Christine Dimitrakaki, Gerasimos Kolaitis, Vasiliki Rotsika, and Yannis Tountas. 2009. "The factor structure of the Strengths and Difficulties Questionnaire (SDQ) in Greek adolescents." *Annals of General Psychiatry* 8 (1): 20. [19]

Gillham, Jane, Zoe Adams-Deutsch, Jaclyn Werner, Karen Reivich, Virginia Coulter-Heindl, Mark Linkins, Breanna Winder, Christopher Peterson, Nansook Park, Rachel Abenavoli, Angelica Contero, and Martin E.P. Seligman. 2011. "Character strengths predict subjective well-being during adolescence." *Journal of Positive Psychology* 6 (1): 31–44. [8]

Gillham, Jane E., John Hamilton, Derek R. Freres, Ken Patton, and Robert Gallop. 2006. "Preventing Depression Among Early Adolescents in the Primary Care Setting: A Randomized Controlled Study of the Penn Resiliency Program." *Journal of Abnormal Child Psychology* 34 (2): 195–211. [8]

Gilman, Rich, and Scott Huebner. 2003. "A review of life satisfaction research with children and adolescents." *School Psychology Quarterly* 18 (2): 192–205. [7]

Goebel, Jan, Markus M. Grabka, Stefan Liebig, Martin Kroh, David Richter, Carsten Schröder, and Jürgen Schupp. 2018. "The German Socio-Economic Panel (SOEP)." *Journal of Economics and Statistics* 239 (2): 345–60. [11, 12]

Goede, Irene H. A. De, Susan J. T. Branje, and Wim H. J. Meeus. 2008. "Developmental Changes in Adolescents' Perceptions of Relationships with Their Parents." *Journal of Youth and Adolescence* 38 (1): 75–88. [8]

Goodman, Robert. 1997. "The Strengths and Difficulties Questionnaire: a research note." *Journal of child psychology and psychiatry* 38 (5): 581–86. [19]

Hanniball, Katherine B., Jodi L. Viljoen, Catherine S. Shaffer, Gira Bhatt, Roger Tweed, Lara B. Aknin, Nathalie Gagnon, Kevin S. Douglas, and Stephen Dooley. 2018. "The Role of Life Satisfaction in Predicting Youth Violence and Offending: A Prospective Examination." *Journal of Interpersonal Violence* 36 (11-12): 5501–29. [7]

Haranin, Emily C., E. Scott Huebner, and Shannon M. Suldo. 2007. "Predictive and Incremental Validity of Global and Domain-Based Adolescent Life Satisfaction Reports." *Journal of Psychoeducational Assessment* 25 (2): 127–38. [7]

Headey, Bruce, and Ruud Muffels. 2018. "A theory of life satisfaction dynamics: Stability, change and volatility in 25-year life trajectories in Germany." *Social Indicators Research* 140 (2): 837–66. [5]

Hounkpatin, Hilda Osafo, Christopher J. Boyce, Graham Dunn, and Alex M. Wood. 2018. "Modeling bivariate change in individual differences: Prospective associations between personality and life satisfaction." *Journal of Personality and Social Psychology* 115 (6): e12–e29. [7]

Jewell, Sarah, and Uma S. Kambhampati. 2014. "Are Happy Youth Also Satisfied Adults? An Analysis of the Impact of Childhood Factors on Adult Life Satisfaction." *Social Indicators Research* 121 (2): 543–67. [6–8, 12, 16, 19]

Jovanović, Veljko. 2016. "The validity of the Satisfaction with Life Scale in adolescents and a comparison with single-item life satisfaction measures: a preliminary study." *Quality of Life Research* 25 (12): 3173–80. [7]

Jung, Sunyoung, and Eunsil Choi. 2017. "Life satisfaction and delinquent behaviors among Korean adolescents." *Personality and Individual Differences* 104 (1): 104–10. [6, 7]

Klimstra, Theo A., William W. Hale, Quinten A. W. Raaijmakers, Susan J. T. Branje, and Wim H. J. Meeus. 2009. "Maturation of personality in adolescence." *Journal of Personality and Social Psychology* 96 (4): 898–912. [7]

Koivumaa-Honkanen, H., M. Koskenvuo, R. J. Honkanen, H. Viinamäki, K. Heikkilä, and J. Kaprio. 2004. "Life dissatisfaction and subsequent work disability in an 11-year follow-up." *Psychological Medicine* 34 (2): 221–28. [7]

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk. 2020. "The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy* 128 (2): 434–67. eprint: https://doi.org/10.1086/704386. [9, 10]

Layard, Richard, Andrew E. Clark, Francesca Cornaglia, Nattavudh Powdthavee, and James Vernoit. 2014. "What Predicts a Successful Life? A Life-Course Model of Well-Being." *Economic Journal* 124 (580): F720–F738. [8, 16, 19]

Liebig, Stefan, Jan Goebel, David Richter, Carsten Schröder, Jürgen Schupp, Charlotte Bartels, Alexandra Fedorets, Andreas Franken, Marco Giesselmann, Markus Grabka, Jannes Jacobsen, Selin Kara, Peter Krause, Martin Kroh, Hannels Kröger, Maria Metzing, Jana Nebelin, Diana Schacht, Paul Schmelzer, Christian Schmitt, Daniel Schnitzlein, Rainer Siegers, Knut Wenzig, and Stefan Zimmermann. 2019. "German Socio-Economic Panel Study (SOEP), version 34." [11, 12]

Lounsbury, John W., Soo-Hee Park, Eric Sundstrom, Jeanine M. Williamson, and Anne E. Pemberton. 2004. "Personality, Career Satisfaction, and Life Satisfaction: Test of a Directional Model." *Journal of Career Assessment* 12 (4): 395–406. [8]

Lucas, Richard E. 2007. "Adaptation and the Set-Point Model of Subjective Well-Being: Does Happiness Change After Major Life Events?" *Current Directions in Psychological Science* 16 (2): 75–79. [7, 12, 16]

Lucas, Richard E., and M. Brent Donnellan. 2007. "How stable is happiness? Using the STARTS model to estimate the stability of life satisfaction." *Journal of Research in Personality* 41 (5): 1091–98. [5, 7]

Lyyra, T.-M., T. M. Tormakangas, S. Read, T. Rantanen, and S. Berg. 2006. "Satisfaction With Present Life Predicts Survival in Octogenarians." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 61 (6): P319–P326. [7]

Mackenbach, Johan P., Irina Stirbu, Albert-Jan R. Roskam, Maartje M. Schaap, Gwenn Menvielle, Mall Leinsalu, and Anton E. Kunst. 2008. "Socioeconomic Inequalities in Health in 22 European Countries." *New England Journal of Medicine* 358 (23): 2468–81. [17]

Martin, Kellie, E. Scott Huebner, and Robert F. Valois. 2008. "Does life satisfaction predict victimization experiences in adolescence?" *Psychology in the Schools* 45 (8): 705–14. [7]

Mazur, Joanna, Monika Szkultecka-Dębek, Anna Dzielska, Mariola Drozd, and Agnieszka Małkowska-Szkutnik. 2018. "What does the Cantril Ladder measure in adolescence?" *Archives of medical science: AMS* 14 (1): 182–89. [7]

Neve, Jan-Emmanuel De, Nicholas A. Christakis, James H. Fowler, and Bruno S. Frey. 2012. "Genes, economics, and happiness." *Journal of Neuroscience, Psychology, and Economics* 5 (4): 193–211. [7]

Pagán-Rodríguez, Ricardo. 2010. "Onset of disability and life satisfaction: evidence from the German Socio-Economic Panel." *European Journal of Health Economics* 11 (5): 471–85. [9]

Paxton, Raheem J., Robert F. Valois, E. Scott Huebner, and J. Wazner Drane. 2006. "Opportunity for Adult Bonding/Meaningful Neighborhood Roles and Life-Satisfaction among USA Middle School Students." *Social Indicators Research* 79 (2): 291–312. [8]

Peitsch, Lorraine, Suzanne L. Tyas, Verena H. Menec, and Philip D. St. John. 2016. "General life satisfaction predicts dementia in community living older adults: a prospective cohort study." *International Psychogeriatrics* 28 (7): 1101–9. [7]

Peterson, Christopher, Willibald Ruch, Ursula Beermann, Nansook Park, and Martin E. P. Seligman. 2007. "Strengths of character, orientations to happiness, and life satisfaction." *Journal of Positive Psychology* 2 (3): 149–56. [8]

Proctor, Carmel, and P. Alex Linley. 2014. "Life Satisfaction in Youth." In *Cross-Cultural Advancements in Positive Psychology*. Springer Netherlands, 199–215. [8]

Proctor, Carmel, P. Alex Linley, and John Maltby. 2018. "Life Satisfaction." In. *Encyclopedia of Adolescence*. Edited by Roger J. R. Levesque. Springer International Publishing. [5, 6, 8, 16]

Proctor, Carmel, Eli Tsukayama, Alex M. Wood, John Maltby, Jennifer Fox Eades, and P. Alex Linley. 2011. "Strengths Gym: The impact of a character strengths-based intervention on the life satisfaction and well-being of adolescents." *Journal of Positive Psychology* 6 (5): 377–88. [8, 19]

Proctor, Carmel L., P. Alex Linley, and John Maltby. 2008. "Youth Life Satisfaction: A Review of the Literature." *Journal of Happiness Studies* 10 (5): 583–630. [5]

Proyer, René T., Fabian Gander, Tobias Wyss, and Willibald Ruch. 2011. "The Relation of Character Strengths to Past, Present, and Future Life Satisfaction among German-Speaking Women." *Applied Psychology: Health and Well-Being* 3 (3): 370–84. [8]

Rauma, Päivi H., Heli Koivumaa-Honkanen, Lana J. Williams, Marjo T. Tuppurainen, Heikki P. Kröger, and Risto J. Honkanen. 2014. "Life Satisfaction and Bone Mineral Density Among Postmenopausal Women." *Psychosomatic Medicine* 76 (9): 709–15. [7]

Reynolds, Arthur J., Judy A. Temple, Dylan L. Robertson, and Emily A. Mann. 2001. "Long-term Effects of an Early Childhood Intervention on Educational Achievement and Juvenile ArrestA 15-Year Follow-up of Low-Income Children in Public Schools." *JAMA* 285 (18): 2339–46. [20]

Ruini, Chiara, Fedra Ottolini, Elena Tomba, Carlotta Belaise, Elisa Albieri, Dalila Visani, Emanuela Offidani, Ernesto Caffo, and Giovanni A. Fava. 2009. “School intervention for promoting psychological well-being in adolescence.” *Journal of Behavior Therapy and Experimental Psychiatry* 40 (4): 522–32. [8]

Schnitzlein, Daniel D., and Christoph Wunder. 2016. “Are We Architects of Our Own Happiness? The Importance of Family Background for Well-Being.” *BE Journal of Economic Analysis & Policy* 16 (1): 125–49. [6]

Schütz, Gabriela, Heinrich W. Ursprung, and Ludger Wößmann. 2008. “Education Policy and Equality of Opportunity.” *Kyklos* 61 (2): 279–308. [17]

Seligman, Martin E. P., Randal M. Ernst, Jane Gillham, Karen Reivich, and Mark Linkins. 2009. “Positive education: positive psychology and classroom interventions.” *Oxford Review of Education* 35 (3): 293–311. [8]

Shek, Daniel T. L., and Ting Ting Liu. 2014. “Life satisfaction in junior secondary school students in Hong Kong.” *Social Indicators Research* 117 (3): 777–94. [5–8, 12, 16, 19]

Siahpush, Mohammad, Matt Spittal, and Gopal K. Singh. 2008. “Happiness and Life Satisfaction Prospectively Predict Self-Rated Health, Physical Health, and the Presence of Limiting, Long-Term Health Conditions.” *American Journal of Health Promotion* 23 (1): 18–26. [7]

Smyth, Joshua M., Matthew J. Zawadzki, Vanessa Juth, and Christopher N. Sciamanna. 2016. “Global life satisfaction predicts ambulatory affect, stress, and cortisol in daily life in working adults.” *Journal of Behavioral Medicine* 40 (2): 320–31. [7]

Specht, Jule, Boris Egloff, and Stefan C. Schmukle. 2012. “Examining Mechanisms of Personality Maturation.” *Social Psychological and Personality Science* 4 (2): 181–89. [7]

Suldo, Shannon M., and E. Scott Huebner. 2004. “Does life satisfaction moderate the effects of stressful life events on psychopathological behavior during adolescence?” *School Psychology Quarterly* 19 (2): 93–105. [5, 7]

Suldo, Shannon M., Devon R. Minch, and Brittany V. Hearon. 2014. “Adolescent Life Satisfaction and Personality Characteristics: Investigating Relationships Using a Five Factor Model.” *Journal of Happiness Studies* 16 (4): 965–83. [8]

Wagner, Gert. G., Joachim R. Frick, Juergen Schupp, Silke Anger, Jan Goebel, Markus M. Grabka, Elke Holst, Peter Krause, Martin Kroh, Elisabeth Liebau, Henning Lohmann, Christian Schmitt, and C. Katharina Spiess. 2011. “German Socio-Economic Panel Study (SOEP), version 26.” [15]

Waltenberg, Fábio D., and Vincent Vandenberghe. 2007. “What does it take to achieve equality of opportunity in education?: An empirical investigation based on Brazilian data.” *Economics of Education Review* 26 (6): 709–23. Economics of Education: Major Contributions and Future Directions - The Dijon Papers. [17]

Weber, Marco, and E. Scott Huebner. 2015. “Early adolescents’ personality and life satisfaction: A closer look at global vs. domain-specific satisfaction.” *Personality and Individual Differences* 83 (9): 31–36. [8]

Weber, Marco, Willibald Ruch, and E. Scott Huebner. 2013. “Adaptation and Initial Validation of the German Version of the Students’ Life Satisfaction Scale (German SLSS).” *European Journal of Psychological Assessment* 29 (2): 105–12. [18]

Weiss, L. A., G. J. Westerhof, and E. T. Bohlmeijer. 2016. “Can we increase psychological well-being? The effects of interventions on psychological well-being: a meta-analysis of randomized control trials.” *PLoS ONE* 11 (6): e0158092. [8]

Appendix 1.A Change in Answer Format

In the first two waves (when children were 7-9 and 8.5-10.5 years old), children were not asked the abstract standard question used in the SOEP and many other surveys. Instead, a more age appropriate question and scale was used. The scale can be seen in Figure 1.A.1. The question was “How satisfied are you most of the time?” (Wie zufrieden bist du die meiste Zeit?).

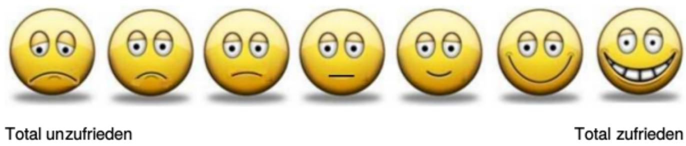


Figure 1.A.1. Life Satisfaction Scale of the First Two Waves

Note: The scale goes from "totally unsatisfied" on the left to "totally satisfied" on the right. Starting in the third wave children answered on an abstract 0 to 10 scale without stylized faces.

Starting in the third wave, all children were asked the standard question on a zero to ten scale. Thus, the question became “All things considered, how satisfied are you with your life as a whole these days?” (in German: Wie zufrieden bist Du gegenwärtig, alles in allem, mit Deinem Leben?).

This raises the question whether the change in wording or the change in the scale leads children to answer differently and how their answers can be made comparable.

The question asked on the smiley scale was coded numerically as going from zero to six. For the later waves we use the zero to ten coding as it was shown to respondents. Simply adding four points to the early variable (such that it goes from four to ten) leads to a very smooth age trend in life satisfaction as can be seen in Figure 1.A.2.

In addition, Figure 1.4.1 shows that despite the larger time lapse between waves 2 and 3 (19 months versus 12 in all later waves) and the change in the question and answer format, the correlation between waves does not decline. For comparison the correlation between answers falls by more than 0.1 for parents. As the parents’ question and scale remained the same over all waves, this fall can be exclusively attributed to the increase in time between interviews.

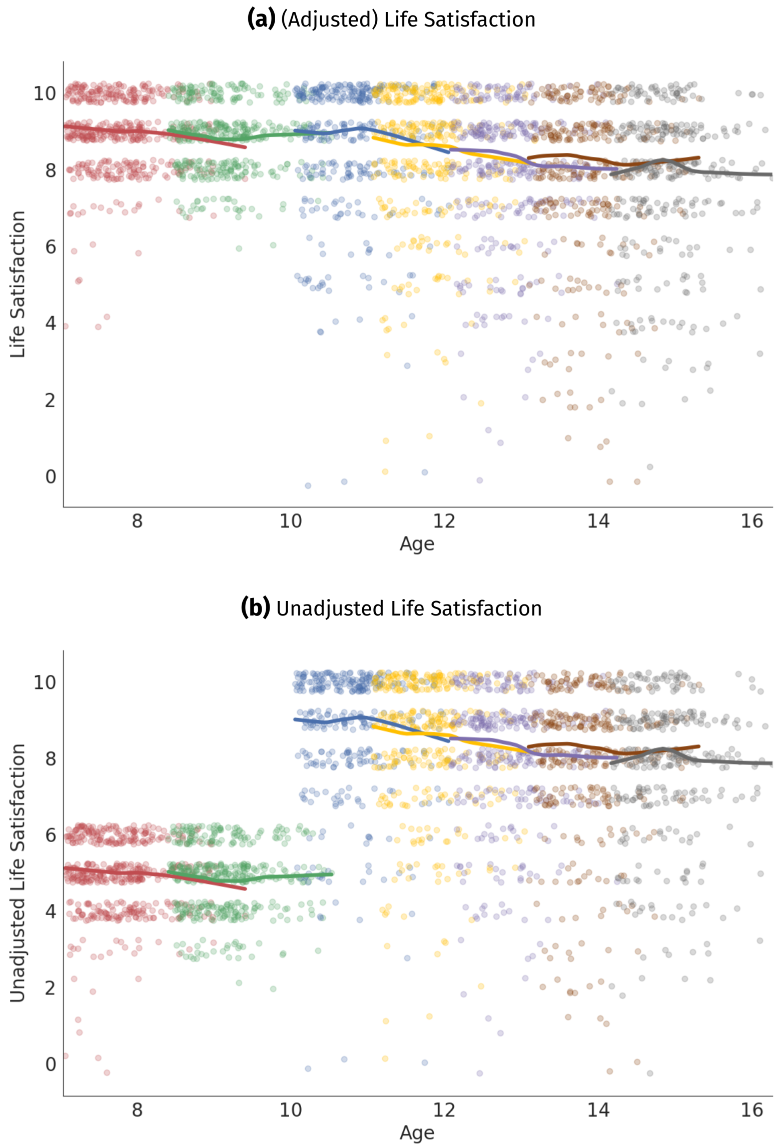


Figure 1.A.2. Life Satisfaction Responses Across the Question Change

Note: The raw zero to six and zero to ten score (Bottom) and transformed (Top) as reported by children across waves by age. Each wave is represented in a different color. The simple addition of four points leads to no visible difference across the change in response possibilities. The number of observations is 702 in the first and 485 in the last wave.

Appendix 1.B Evidence Against Selective Attrition

The briq family panel has very low levels of attrition and there is very little evidence for selective attrition. Over two thirds of children still participate in the 7th wave, the last one used in this study (see Table 1.B.1). In addition, there is little evidence for selective attrition since treatment status, baseline happiness and their interaction explain only 1% of the variance in who decides to drop out. This is very much in line with the fact that our main results are not affected by whether we use the balanced or unbalanced panel (see Table 1.4.2, columns (2) and (3)).

Table 1.B.1. Do Treatment Status and Life Satisfaction at Baseline Predict Dropping out of the Study?

	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7
Intercept	-0.18	0.08	0.05	0.03	0.18	0.16
Low SES Treatment	0.44*	0.62**	0.62**	0.82***	0.38	0.46
Baseline Life Satisfaction	0.04**	0.03	0.03	0.03	0.02	0.02
Treatment x Baseline Life Satisfaction	-0.05*	-0.07*	-0.07*	-0.09**	-0.04	-0.05
R-squared	0.008	0.006	0.006	0.011	0.002	0.004
Number of Children Lost	110	229	217	241	242	247

Note: Each column shows the OLS regression coefficients predicting which individuals have dropped out in the respective wave for treatment status, baseline life satisfaction and their interaction.

Appendix 1.C Robustness Checks

In this section we provide evidence on the robustness of all of our main results.

1.C.1 Robustness of the SES Gap with Respect to the SES Classification

We categorized families as low SES if they meet at least one of the following criteria:

- **Low income**: Equivalence income of the household is lower than 1065 Euro, i.e. the 30% percentile of the German income distribution.
- **Low education**: Both parents are not qualified for university studies.

- **Single parent status**: The child lives in a single-parent household.

Here we check if the SES gap is robust to using each dimension on its own to classify children as high or low SES.

The results are similar to our joint measure of socio-economic status, especially for parental income and single parenthood but in general less statistically significant. The SES gap is smaller and statistically insignificant if only parental education is used to classify families as low or high SES. Part of this might be driven by Germany's strong apprenticeship system that allows Germans to command higher wages and attain higher social status without the need to receive a university entrance diploma.

Children of more educated and less educated parents start out very similar to children of less educated parents reporting a slightly higher life satisfaction. This does not appear to change over time. By the time they are 15 no gap has emerged.

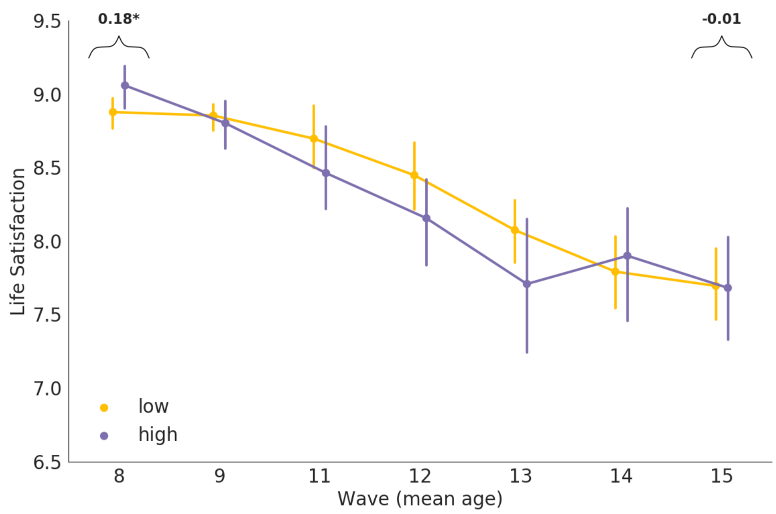


Figure 1.C.1. Life Satisfaction Over Time By Parental Education

Note: Points show mean life satisfaction of all children that answered the life satisfaction question in that wave, where the violet ("high") line shows the means of children of highly educated parents and the yellow ("low") line refers to children of parents where neither parent holds a university entrance diploma. Whiskers indicate 95% confidence intervals. The mean difference is shown above the first and last wave with stars indicating the statistical significance between the two groups. (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$)

As with the other dimensions of socio-economic status, children from higher earning and lower earning households start out very similarly with all children reporting similarly high life satisfaction. However, by the time they are 15 a weakly significant gap of 0.43 has opened between the two groups with richer children reporting higher life satisfaction during adolescence.

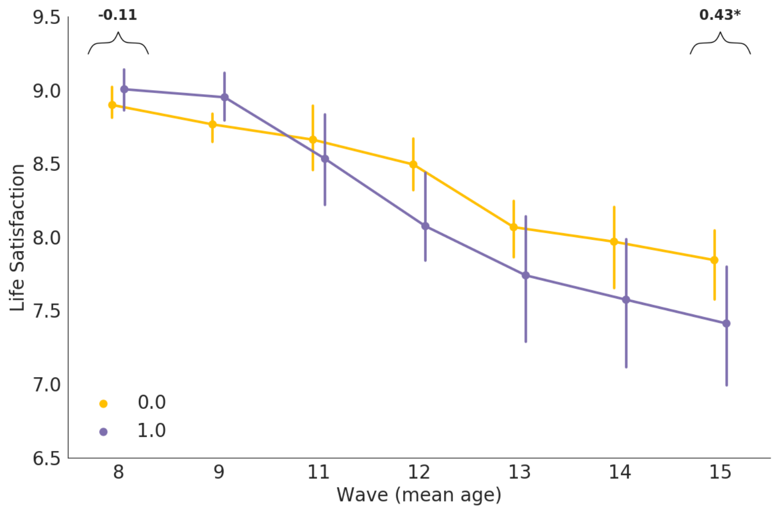


Figure 1.C.2. Life Satisfaction Over Time By Parental Income

Note: Points show mean life satisfaction of all children that answered the life satisfaction question in that wave. The violet (1) line shows the means of the children of poorer parents and the yellow (0) line shows the means of children of richer parents. Whiskers indicate 95% confidence intervals. The mean difference is shown above the first and last wave with stars indicating the statistical significance between the two groups. (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$)

Despite the negative effects that would be expected from children growing up in single-parent households, our data show only very small differences in the life satisfaction between children who live with one or more parents at baseline. However, as children grow older, children that lived in single parent homes report a $^1/_3$ point, albeit statistically insignificant, lower life satisfaction than their peers that live with more than one parent. This effect is smaller than that found in other studies where growing up in intact families was correlated with higher life satisfaction, e.g. Shek and Liu (2014). However, single parent status is likely to have changed for a significant fraction of the children who lived in single parent households at baseline.

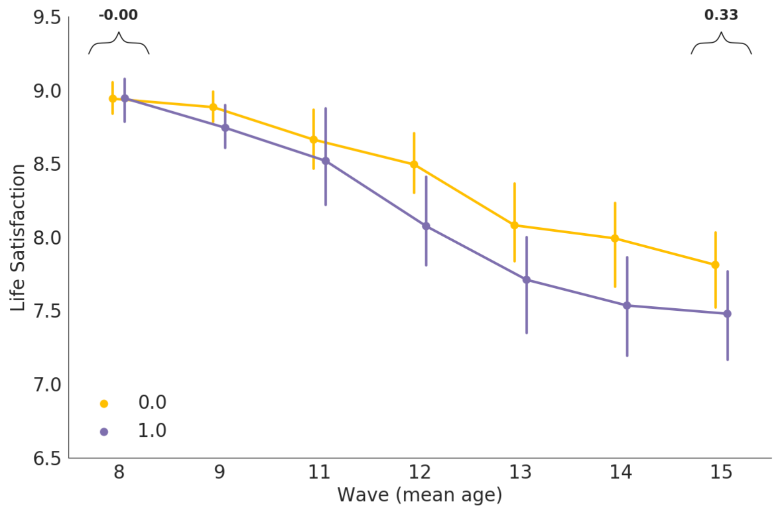


Figure 1.C.3. Life Satisfaction Over Time By Single Parent Status

Note: Points show mean life satisfaction of all children that answered the life satisfaction question in the respective wave. The violet (1) line shows the means of children who lived in single-parent homes at baseline. The yellow line (0) shows children growing up with more than one parent. Whiskers indicate 95% confidence intervals. The mean difference is shown above each wave with stars indicating the statistical significance between the two groups. (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$)

1.C.2 Robustness of the SES Gap and Treatment Effect

Next, we complement our analysis on the SES gap (Figure 1.4.4a) and the treatment effect (Table 1.4.2) with a more detailed regression analysis which is shown in Table 1.C.1. The age trends of the High SES Control group and the Low SES Treatment group are both substantively different from the Low SES Control group and highly statistically significant. This holds irrespective of whether we control for baseline life satisfaction (column 1 versus column 2) and whether we only include children that participated in all waves (column 3).

Table 1.C.1. SES Gap in Age Trends in Life Satisfaction

	(1)	(2)	(3)
	Life Satisfaction	Life Satisfaction	Life Satisfaction
Age	-0.21***	-0.20***	-0.22***
High SES Control x Age	0.08***	0.07**	0.10***
Low SES Treatment x Age	0.08***	0.07***	0.08***
Resident in Cologne	0.13**	0.11*	0.07
Baseline Life Satisfaction		0.40***	0.33***
Constant	10.56***	6.94***	7.71***
High SES Control	-0.67**	-0.52*	-0.79**
Low SES Treatment	-0.75***	-0.64***	-0.70**
Observations	3703	3698	2950

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from an unbalanced random effects regression model controlling for place of residence (Cologne or Bonn). Intercepts and age trends are separately estimated for children of the Low SES Control group, High SES Control and Treatment group. The second and third column also control for baseline life satisfaction. The third column only uses the balanced panel.

The treatment effect is not only visible when looking at trends but also when we look at the later waves separately. However, looking at the raw means (Figure 1.C.4) it is clear that given our sample size, the confidence intervals of the groups in each wave do overlap to some degree.

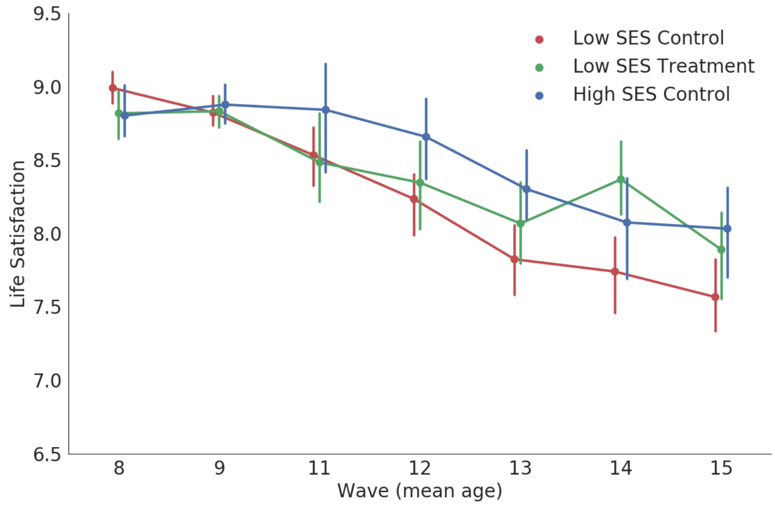


Figure 1.C.4. Life Satisfaction by Treatment Arm in Every Wave

Note: Points show group mean life satisfaction of all individuals that answered the life satisfaction question in the respective wave. The red line shows the Low SES Control group, the blue line the High SES Control group and the Treatment group's mean response is shown in green. Whiskers indicate 95% confidence intervals. The mean age in each wave is plotted on the x axis.

We find consistently large gaps between high and low SES children in the last two waves after controlling for baseline life satisfaction and age at the interview (see Table 1.C.2). In addition, for all specifications the treatment effect stays at least close to statistically significant. To show the robustness of our results to the exact framing of the question table 1.C.2 includes the Students' Life Satisfaction Score (SLSS) (Weber, Ruch, and Huebner (2013)) – which was only collected in the 7th wave – in column two. That column shows that both the treatment effect and the SES gap are still sizable if we use a multi item scale of life satisfaction. However, with the SLSS scale there does remain a gap between the High SES Control group and our treatment group.

Table 1.C.2. Differences in Life Satisfaction by SES and Treatment Status

	(1) Mean Life Satisfaction in Waves 6 and 7	(2) SLSS in Wave 7	(3) Life Satisfaction in Wave 6	(4) Life Satisfaction in Wave 7
High SES Control	0.56***	0.44***	0.46*	0.65***
Low SES Treatment	0.48***	0.19	0.59***	0.38*
Age at Interview	-0.19	0.00	-0.30	-0.09
Baseline Life Satisfaction	0.22***	0.16***	0.19**	0.25**
Resident in Cologne	0.06	-0.00	0.09	0.03
Constant	8.39***	-1.57	10.27***	6.50**
Observations	415	408	415	415

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows the coefficients from a linear regression model of different measures of life satisfaction in the last waves. The coefficients were estimated using Ordinary Least Squares. We control for (mean) age at the time(s) of the interview(s), baseline life satisfaction to account for reporting behavior and city of residence. Low SES Control is the omitted category. The High SES Control group always reports statistically significantly higher life satisfaction. This holds irrespective of the variable we use to measure life satisfaction. The Treatment group also reports higher life satisfaction. Comparing Treatment and Low SES Control group yields p values of the two being equal of 0.01 , 0.11 , 0.00 , and 0.09 respectively. Effect sizes are in life satisfaction points on a zero to ten scale except for the Students' Life Satisfaction Score (SLSS) where effect sizes are in standard deviations.

Appendix 1.D Evidence on Treatment Effect Heterogeneity

This section provides the statistical analyses for the discussion of treatment effect heterogeneity and potential channels.

1.D.1 Treatment Effect Heterogeneity With Respect to Parental Life Satisfaction

First, we investigate whether children who are lacking a role model displaying high life satisfaction in their primary caregiver benefit especially from the intervention. Given that children often learn by imitation one possible driver of our treatment effect could be that children whose caregivers are less satisfied with their lives learn relevant skills and behaviors for later life satisfaction through imitation from their highly satisfied mentor and thus benefit especially from the intervention. On the other hand, life satisfaction has a strong genetic component (Bartels, 2015) and parental life satisfaction also reflects shared circumstances. We use parental life satisfaction at baseline to classify children to exclude possible treatment effects of the mentoring intervention on the primary caregiver and changes in the environment shared by main caregiver and child. Children who lack a role model of high life satisfaction in their primary caregiver appear to benefit more from the intervention, however the difference between treatment effects is not always statistically significant.

Table 1.D.1. Age Trends in Child Life Satisfaction on Treatment Status by Life Satisfaction of the Caregiver at Baseline

	(1) Life Satisfaction
Low SES Treatment	0.05
Age	-0.14***
Low SES Treatment x Age	-0.01
Unhappy Caregiver	0.67*
Low SES Treatment x Unhappy Caregiver	-0.89
Unhappy Caregiver x Age	-0.09**
Low SES Treatment x Unhappy Caregiver x Age	0.10*
Baseline Life Satisfaction	0.43***
Resident in Cologne	0.06
Constant	6.24***
Observations	2881

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from an unbalanced random effects regression model controlling for baseline life satisfaction and place of residence. Intercepts and age trends are estimated separately by treatment group and whether reported parental life satisfaction was below or above the median at baseline. For short we refer to the indicator of primary caregivers who reported below median life satisfaction as "unhappy".

The difference in age trends between Treatment and Low SES Control group in the group above the median is -0.01 (p=0.91) and 0.09 (p=0.00) below the median. The difference is 0.10 points per year (p=0.09).

Table 1.D.2. Cross Section Regression of Child Life Satisfaction on Treatment Status by Life Satisfaction of the Caregiver at Baseline

	(1) Mean Life Satisfaction in Waves 6 and 7
Low SES Treatment	0.24
Below	-0.51*
Low SES Treatment x Below	0.30
Age at Interview	-0.15
Baseline Life Satisfaction	0.31***
Resident in Cologne	0.06
Constant	7.35**
Observations	323

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from a linear regression model of life satisfaction. The coefficients were estimated using Ordinary Least Squares. We control for baseline life satisfaction and place of residence. To reduce measurement error we use the mean of children's life satisfaction reports in the last two waves. The treatment effect in the group above the median is 0.24 (p=0.53) and 0.54 (p=0.01) points below. The difference is 0.30 points (p=0.47).
Life satisfaction is measured as the average of answers at ages 14 and 15 on a zero to ten scale.

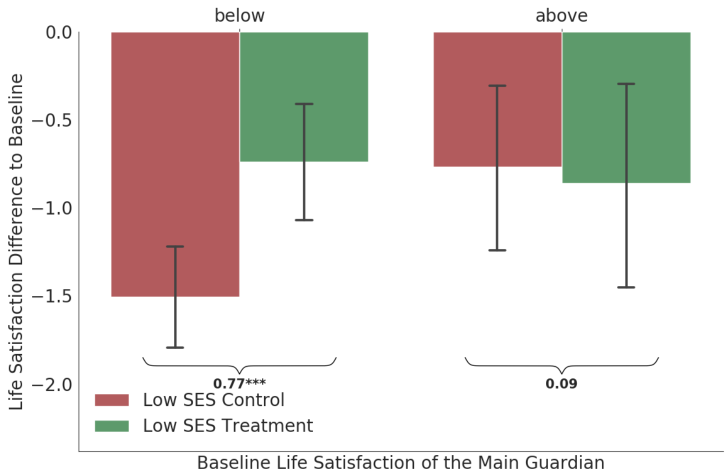


Figure 1.D.1. Treatment Effect by Baseline Life Satisfaction of the Caregiver

Note: Children whose caregiver reported a life satisfaction below the median at baseline benefit more from the intervention than children of caregivers who are more satisfied with their lives. There appears to be no treatment effect on life satisfaction in the later group. To reduce measurement error and control for the individual scaling we use the mean life satisfaction difference to baseline in the last two waves. Each bar shows the mean of the Low SES Control or Treatment group. The absolute treatment effect and its statistical significance level are included in each plot (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

1.D.2 Treatment Effect Heterogeneity With Respect to Number of Joint Activities

Secondly, we look at treatment effect heterogeneity with respect to the number of activities children are reported to do with their caregiver at baseline. As one main focus of the Baloo and You intervention is to expand the child's horizon through joint activities, children who are reported to have few joint engaging activities with their caregiver could be expected to benefit especially from the intervention if it is the experience of these activities that drive our treatment effect. We do not find such a pattern. The null results are robust to which activities we include in our measure of joint engaging activities.

Table 1.D.3. Age Trends in Life Satisfaction by Treatment Status and Joint Activities

	(1) Life Satisfaction
Low SES Treatment	-0.76**
Age	-0.23***
Low SES Treatment x Age	0.09**
Below	-0.35
Low SES Treatment x Below	0.30
Below x Age	0.04
Low SES Treatment x Below x Age	-0.04
Baseline Life Satisfaction	0.42***
Resident in Cologne	0.06
Constant	6.93***
Observations	2956

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from an unbalanced random effects regression model controlling for baseline life satisfaction and place of residence. Intercepts and age trends are estimated separately by treatment group and whether reported joint activities were below or above the median at baseline (above is the omitted category).
The difference in age trends between Treatment and Low SES Control group in the group above the median is 0.09 (p=0.02) and 0.05 (p=0.19) below the median. The difference is -0.04 points per year (p=0.40). Variables that went into the joint activity variable: conversations, eating meals together, outdoor activities, games (excluding computer games), reading stories in German and music making.

Table 1.D.4. Age Trends in Life Satisfaction by Treatment Status by an Alternative Measure of Engaging Activities

	(1)
	Life Satisfaction
Low SES Treatment	-0.61*
Age	-0.23***
Low SES Treatment x Age	0.06*
Below	-0.58*
Low SES Treatment x Below	0.02
Below x Age	0.06*
Low SES Treatment x Below x Age	0.01
Baseline Life Satisfaction	0.43***
Resident in Cologne	0.05
Constant	7.04***
Observations	2956

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from an unbalanced random effects regression model controlling for baseline life satisfaction and place of residence. Intercepts and age trends are estimated separately by treatment group and whether reported joint activities were below or above the median at baseline.
The difference in age trends between Treatment and Low SES Control group in the group above the median is 0.06 (p=0.09) and 0.07 (p=0.04) below the median. The difference is 0.01 points per year (p=0.87).
Variables that went into this broader activity variable: conversations, eating main meals together, outdoor activities, games (excluding computer games), reading stories in German and music making as well as play dates, doing crafts together, excursions and doing sports together.

Table 1.D.5. Cross Section Regression of Life Satisfaction on Treatment Status by Joint Activities

	(1) Mean Life Satisfaction in Waves 6 and 7
Low SES Treatment	0.66**
Below	0.16
Low SES Treatment x Below	-0.34
Age at Interview	-0.15
Baseline Life Satisfaction	0.30***
Resident in Cologne	0.05
Constant	6.98**
Observations	330

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from a linear regression model of life satisfaction. To reduce measurement error we use the mean of children's life satisfaction reports in the two latest waves. The coefficients were estimated using Ordinary Least Squares controlling for baseline life satisfaction and place of residence.
The treatment effect in the group above the median is 0.66 (p=0.01) and 0.32 (p=0.23) points below. The difference is -0.34 points (p=0.35).
Life satisfaction is measured as the average of answers at ages 14 and 15 on a zero to ten scale. Variables that went into the joint activity variable: conversations, eating meals together, outdoor activities, games (excluding computer games), reading stories in German and music making.

Table 1.D.6. Cross Section Regression of Life Satisfaction on Treatment Status by an Alternative Measure of Engaging Activities

	(1) Mean Life Satisfaction in Waves 6 and 7
Low SES Treatment	0.48*
Below	0.22
Low SES Treatment x Below	-0.01
Age at Interview	-0.12
Baseline Life Satisfaction	0.30***
Resident in Cologne	0.04
Constant	6.51**
Observations	330

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from a linear regression model of life satisfaction. To reduce measurement error we use the mean of children's life satisfaction reports in the two latest waves. The coefficients were estimated using Ordinary Least Squares controlling for baseline life satisfaction and place of residence.

The treatment effect in the group above the median is 0.48 (p=0.08) and 0.47 (p=0.05) points below. The difference is -0.01 points (p=0.97).

Life satisfaction is measured as the average of answers at ages 14 and 15 on a zero to ten scale.

Variables that went into this broader activity variable: conversations, eating main meals together, outdoor activities, games (excluding computer games), reading stories in German and music making as well as play dates, doing crafts together, excursions and doing sports together.

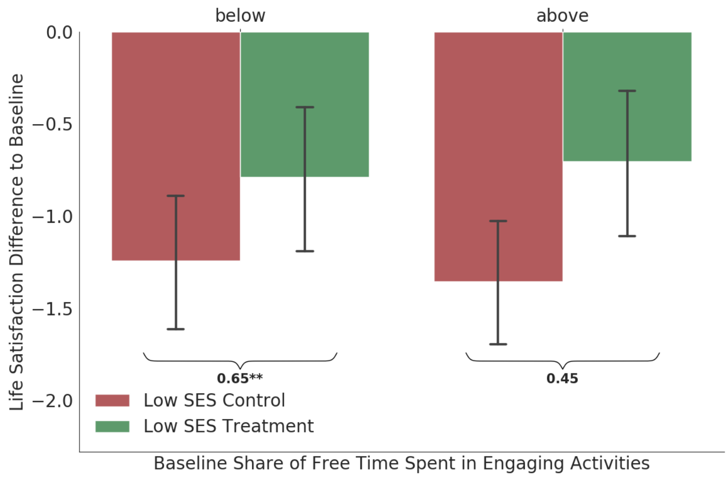


Figure 1.D.2. Treatment Effect Heterogeneity With Respect to Engaging Activities

Note: The figure shows the treatment effects of children who are reported to spend above median or below median shares of their time in joint activities with their main caregiver. To reduce measurement error and control for the individual scaling we use the mean life satisfaction difference to baseline in the last two waves. Each bar shows the mean of the Low SES Control or Treatment group. The absolute treatment effect and its statistical significance level are included in each plot (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

1.D.3 Treatment Effect Heterogeneity With Respect to Socio-Emotional Development

Another possible channel we explore is that of socio-emotional development (Frijters, Johnston, and Shields, 2014; Layard et al., 2014; Shek and Liu, 2014), measured by the *"Strengths and Difficulties Questionnaire"* (SDQ, Goodman (1997)). We use both the aggregate SDQ score and the conduct problems sub scale since behavioral problems in particular (Frijters, Johnston, and Shields, 2014) were associated with low life satisfaction. Looking at treatment effect heterogeneity both with respect to the aggregate SDQ score as well as the score for conduct problems in particular, we find very mixed results (Figure 1.D.3 and Tables 1.D.7, 1.D.8, 1.D.9 and 1.D.10). The differences in treatment effects across both dimensions are statistically insignificant and vary in sign and size.

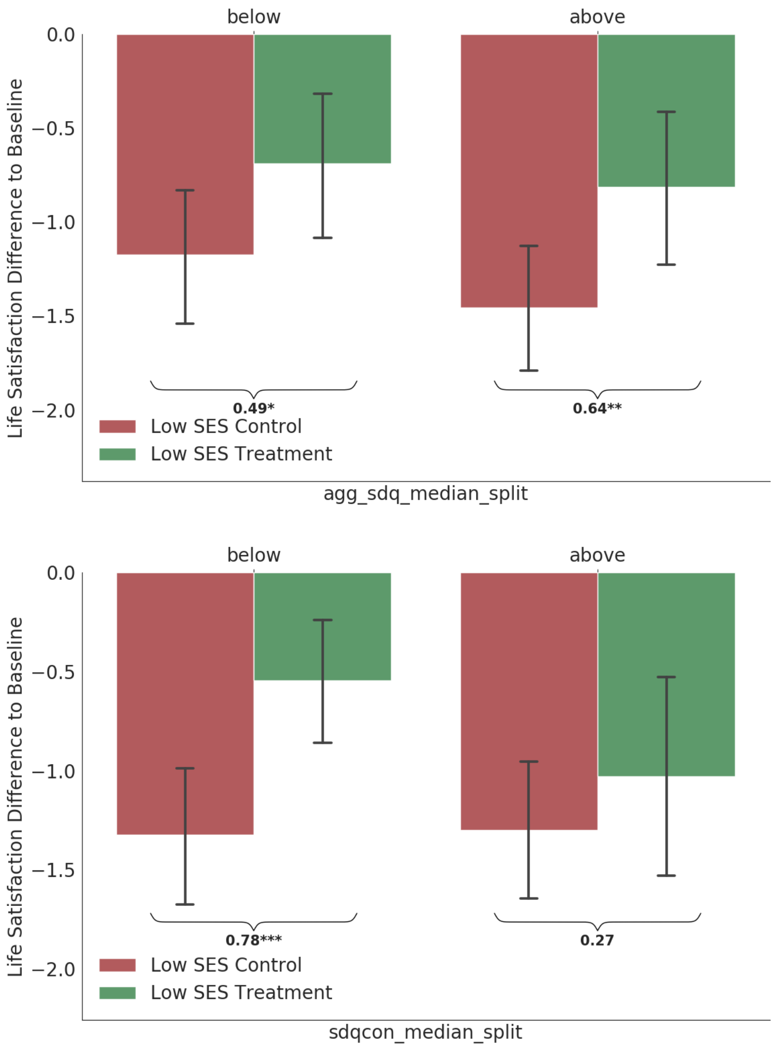


Figure 1.D.3. Treatment Effect Heterogeneity With Respect to Socio-Emotional Development

Note: The figure shows the treatment effects of children who are reported to have scores above or below the median on the aggregate scale (upper figure) or conduct problems scale (lower figure) of the *Strengths and Difficulties Questionnaire* (SDQ, Goodman (1997)). To reduce measurement error and control for the individual scaling we use the mean life satisfaction difference to baseline in the last two waves. Each bar shows the mean of the Low SES Control or Treatment group. The absolute treatment effect and its statistical significance level are included in each plot (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table 1.D.7. Panel Regression of Life Satisfaction on Treatment Status by Aggregate SDQ Score

	(1)
	Life Satisfaction
Low SES Treatment	-1.00***
Age	-0.24***
Low SES Treatment x Age	0.11***
Below	-0.63**
Low SES Treatment x Below	0.71
Below x Age	0.07**
Low SES Treatment x Below x Age	-0.08
Baseline Life Satisfaction	0.42***
Resident in Cologne	0.06
Constant	7.14***
Observations	2964

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from an unbalanced random effects regression model controlling for baseline life satisfaction and place of residence. Intercepts and age trends are estimated separately by treatment group and whether children scored a below or above median aggregate SDQ score at baseline. The difference in age trends between Treatment and Low SES Control group in the group above the median is 0.11 (p=0.00) and 0.03 (p=0.33) below the median. The difference is -0.08 points per year (p=0.14).

Table 1.D.8. Cross Section Regression of Life Satisfaction on Treatment Status by Aggregate SDQ Scores

	(1) Mean Life Satisfaction in Waves 6 and 7
Low SES Treatment	0.52**
Below	0.29
Low SES Treatment x Below	-0.05
Age at Interview	-0.14
Baseline Life Satisfaction	0.29***
Resident in Cologne	0.05
Constant	6.75**
Observations	332

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from a linear regression model of life satisfaction. Coefficients were estimated using Ordinary Least Squares controlling for baseline life satisfaction and place of residence.
The treatment effect in the group above the median is 0.52 (p=0.04) and 0.46 (p=0.09) points below. The difference is -0.05 points (p=0.88).
Life satisfaction is measured as the average of answers at ages 14 and 15 on a zero to ten scale.

Table 1.D.9. Panel Regression of Life Satisfaction on Treatment Status by SDQ Conduct Problems

	(1)
	Life Satisfaction
Low SES Treatment	-0.52
Age	-0.23***
Low SES Treatment x Age	0.06
Below	-0.26
Low SES Treatment x Below	-0.21
Below x Age	0.04
Low SES Treatment x Below x Age	0.02
Baseline Life Satisfaction	0.42***
Resident in Cologne	0.07
Constant	6.99***
Observations	2964

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Note: The table shows coefficients from an unbalanced random effects regression model controlling for baseline life satisfaction and place of residence. Intercepts and age trends are estimated separately by treatment group and whether reported score of SDQ conduct problems was below or above the median at baseline.

The difference in age trends between Treatment and Low SES Control group in the group above the median is 0.06 (p=0.17) and 0.08 (p=0.01) below the median. The difference is 0.02 points per year (p=0.66).

Table 1.D.10. Cross Section Regression of Life Satisfaction on Treatment Status by SDQ Conduct Problems

	(1) Mean Life Satisfaction in Waves 6 and 7
Low SES Treatment	0.27
Below	0.19
Low SES Treatment x Below	0.39
Age at Interview	-0.13
Baseline Life Satisfaction	0.29***
Resident in Cologne	0.09
Constant	6.71**
Observations	332

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

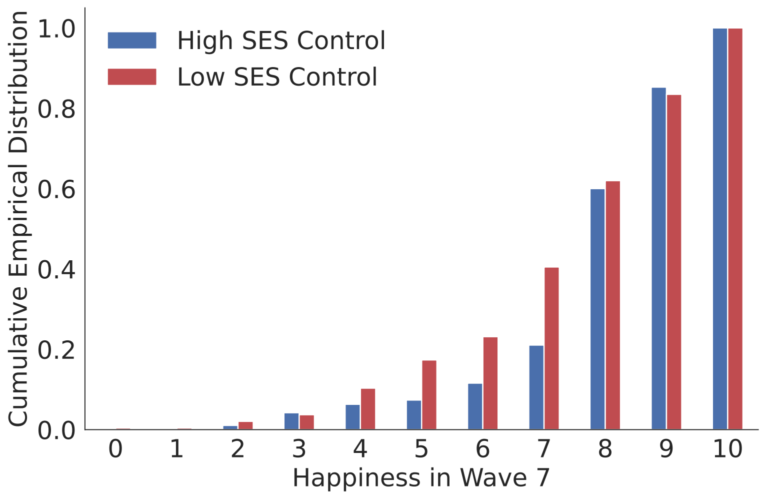
Note: The table shows coefficients from a linear regression model of mean life satisfaction of the last two waves. Coefficients were estimated using Ordinary Least Squares controlling for baseline life satisfaction and place of residence. The treatment effect in the group above the median is 0.27 (p=0.34) and 0.66 (p=0.01) points below.
The difference is 0.39 points (p=0.32). Life satisfaction is measured as the average of answers at ages 14 and 15 on a zero to ten scale.

Appendix 1.E Evidence Regarding First Order Stochastic Dominance

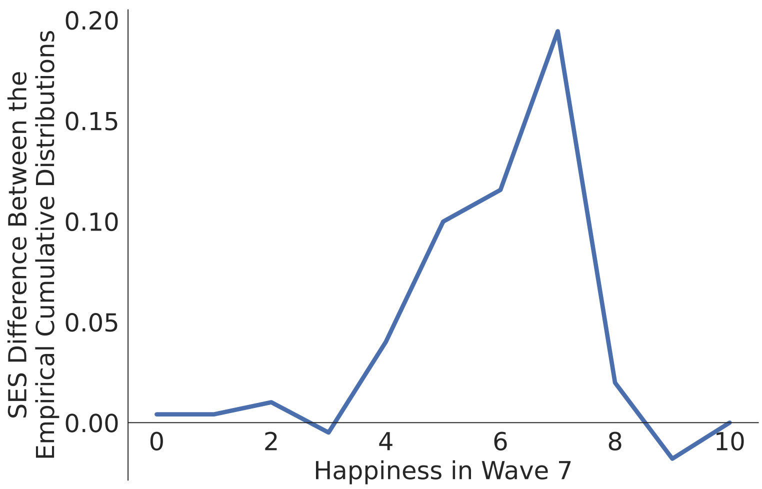
We now show that the distribution of the Low SES Control group is mostly dominated by that of the High SES Control group and that of the Treatment group. As our variable is ordinal, the choice to label the values zero to ten is arbitrary. While we have simply maintained the numerical values that were shown to the survey participants in our panel, any choice of numerical values that preserves the order would be a valid representation. This means that differences in means could be entirely driven by the scale choice. To be certain that the ordering of means holds irrespective of the scale, we would have to show that the distribution of one group first order stochastically dominates the later (Bond and Lang, 2019).

Figure 1.E.1 plots the distribution of life satisfaction in the latest wave between the Low SES Control group and the High SES Control group. The Low SES Control Group has more mass to the left of the distribution but there are small exceptions at the edges. On average, however, the Low SES Control group has four percentage points probability mass to the left of the High SES Control group. Thus, only very particular distortions to the scale would be able to reverse the sign of the SES gap we find.

The case is similar for the comparison between the Low SES Control group and the Treatment group in Figure 1.E.2. The Low SES Control Group has again more mass to the left of the distribution but there are fewer crossings. The plot of the difference between the two empirical distributions shows that the empirical distributions cross once at nine. Overall, the Low SES Control group has two percentage points probability mass to the left of the Treatment group. As with the SES Gap, only very particular distortions of the scale would be able to reverse the sign of our treatment effect.



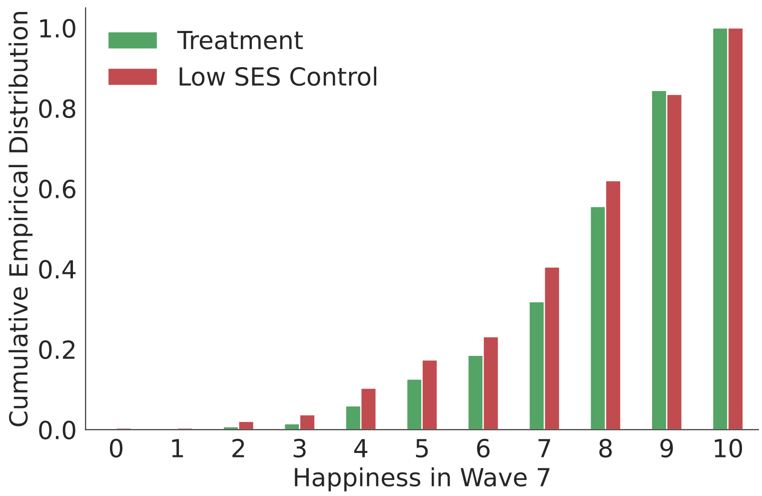
(a) Empirical Distribution of the Low and High SES Control Group



(b) Difference Between the Empirical Distributions of the Low and High SES Control Group

Figure 1.E.1. Comparison of the Empirical Distribution of Life Satisfaction in Wave 7 Between the Low and High SES Control Group

Note: The upper figure shows the cumulative empirical distributions of life satisfaction of the Low and High SES Control groups in wave 7. The figure below shows the difference between the cumulative distribution functions.



(a) Empirical Distribution of the Low SES Control and Treatment Group

(b) Difference Between the Empirical Distributions of the Low SES Control and Treatment Group

Figure 1.E.2. Comparison of the Empirical Distribution of Life Satisfaction in Wave 7 Between the Low SES Control and Treatment Group

Note: The upper figure shows the cumulative empirical distributions of life satisfaction of the Low SES Control and Treatment groups in wave 7. The lower figure shows the difference between the cumulative distribution functions.

References

Bartels, Meike. 2015. "Genetics of Wellbeing and Its Components Satisfaction with Life, Happiness, and Quality of Life: A Review and Meta-analysis of Heritability Studies." *Behavior Genetics* 45 (2): 137–56. [35]

Bond, Timothy N., and Kevin Lang. 2019. "The Sad Truth about Happiness Scales." *Journal of Political Economy* 127 (4): 1629–40. [49]

Frijters, Paul, David W. Johnston, and Michael A. Shields. 2014. "Does Childhood Predict Adult Life Satisfaction? Evidence from British Cohort Surveys." *Economic Journal* 124 (580): F688–F719. [43]

Goodman, Robert. 1997. "The Strengths and Difficulties Questionnaire: a research note." *Journal of child psychology and psychiatry* 38 (5): 581–86. [43, 44]

Layard, Richard, Andrew E. Clark, Francesca Cornaglia, Nattavudh Powdthavee, and James Vernoit. 2014. "What Predicts a Successful Life? A Life-Course Model of Well-Being." *Economic Journal* 124 (580): F720–F738. [43]

Shek, Daniel T. L., and Ting Ting Liu. 2014. "Life satisfaction in junior secondary school students in Hong Kong." *Social Indicators Research* 117 (3): 777–94. [30, 43]

Weber, Marco, Willibald Ruch, and E. Scott Huebner. 2013. "Adaptation and Initial Validation of the German Version of the Students' Life Satisfaction Scale (German SLSS)." *European Journal of Psychological Assessment* 29 (2): 105–12. [33]

Chapter 2

People Meet People – An Epidemiological Model for Evaluating Non-Pharmaceutical Interventions

Joint with Janoś Gabler, Tobias Raabe and Hans-Martin von Gaudecker

2.1 Introduction

Since early 2020, the CoViD-19 pandemic has presented an enormous challenge to humanity in many dimensions. The most frequent initial response were different combinations of non-pharmaceutical interventions (NPIs) to reduce contacts between individuals. While this has allowed some countries to sustain equilibria with very low infection numbers,[1] most have seen large fluctuations of infection rates over time. In addition, containment measures have become increasingly diverse and now include rapid testing, more nuanced NPIs and contact tracing.

Epidemiological models have not been designed to evaluate or predict the effects of such fine-grained policies. In particular, they usually lack detailed contact networks and a case detection model that takes test demand and test availability into account.

* The authors are grateful for support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866 – and through CRC-TR 224 (Projects A02 and C01), by the IZA Institute of Labor Economics and by the Google Cloud CoViD-19 research credits program.

1. See Contreras et al. (2021) for a theoretical equilibrium at low case numbers which is sustained with test-trace-and-isolate policies.

We develop a quantitative model incorporating these factors simultaneously. The framework allows to combine a wide variety of data and mechanisms in a timely fashion, making it useful to predict the effects of various interventions. In particular, our model is unique to include the following features: Firstly, we have a fine-grained representation of contact networks, distinguishing different contact types (such as school, work etc.) and recurrent and non-recurrent contacts, with a special focus on the education sector. We include vacations, teacher and student interactions and emergency care for students of parents working in systemically relevant sectors. Secondly, we have a sophisticated model of case detection: The number of available tests varies over time to fit the estimated share of known cases and tests can be triggered by both symptoms and positive rapid tests. Thirdly, we model the time-varying sensitivity of rapid tests in great detail. Behavioral reactions to symptoms or positive tests are explicitly taken into account.

This chapter proceeds as follows: We start with a literature review in Section 2.2, followed by an overview over the model (Section 2.3). After the overview, we describe each model aspect in turn. Firstly, we go into detail on how the number of contacts of each individual is determined (Section 2.4) and how they can be reduced by non-pharmaceutical interventions and endogenous contact reductions in Section 2.5. We then explain how we match individuals to get from the policy-adjusted number of contacts to actual contacts with potential infections in Section 2.6, followed by an explanation how seasonality modulates the infection probability (Section 2.7). Next, we show how the disease progresses for infected individuals in Section 2.8. Section 2.9 then describes how infected individuals can learn about their infection through PCR and rapid tests. Lastly, we explain how we solve the initial conditions problem (Section 2.10). Section 2.11 concludes.

2.2 Literature Review

A commonly used model class in epidemiology are agent-based simulation models. In a prototypical agent-based simulation model, individuals are simulated as moving particles. Infections take place when two particles come closer than a certain contact radius (e.g. Silva et al. (2020) and Cuevas (2020)). While the simulation approach makes it easy to incorporate heterogeneity in disease progression, it is hard to incorporate heterogeneity in meeting patterns. Moreover, policies are modeled as changes in the contact radius or momentum equation of the particles. The translation from real policies to corresponding model parameters is a hard task.

These shortcomings have motivated variations of agent-based simulation models where moving particles have been replaced by contact networks for households, work and random contacts. The OpenABM-Covid-19 model by Hinch et al. (2021) and the model by Aleta et al. (2020) are the closest in spirit to ours.

The OpenABM-Covid-19 model by Hinch et al. (2021) also uses detailed contact networks for workplaces, schools and households and can evaluate the effect of several NPIs. The main focus of their application are contact tracing policies (Abueg et al., 2021). Recently they have also added support for multiple virus strains and vaccinations.

Aleta et al. (2020) develop an agent-based simulation model with a very high geographical resolution by estimating contact networks from fine-grained mobility data for the Boston metropolitan area. They use this model to show how NPIs, contact tracing and PCR testing can influence the infection dynamics. However, they do not calibrate their model to match actual infection numbers which makes it more suitable to explore the general mechanics of different disease mitigation measures than for their quantitative evaluation.

Bicher et al. (2021) simulate the entire Austrian population. They use data from the first wave (February 21 to April 9, 2020) to calibrate their model and predict the effect of different NPIs and contact tracing policies until November 2020. They use the same data provided by Mossong et al. (2008) as we do to calibrate contact networks for households, workplaces and schools. The model focuses on analyzing the effect of different contact tracing strategies and not on modeling enacted Austrian policies over a long period of time.

Moreover, there are several working papers that develop agent-based simulation models with contact networks in conjunction with economic models. Examples are Basurto et al. (2020), Delli Gatti and Reissl (2020) and Mellacher (2020).

Our model combines elements from the above models and adds several others. To the best of our knowledge, our model is the only one with the following features:

1. The free model parameters have been estimated with the method of simulated moments (McFadden, 1989, see Chapter 3). Despite having few free parameters our model does an excellent job in explaining observed case numbers and the spread of the Alpha variant over more than nine months of data.

2. We have a fine-grained mechanistic representation of schools and preschools. We can thus easily model all schooling policies that have been implemented in Germany between March 2020 and November 2021. This includes complete school closures, phases where only those students whose parents could not find any private childcare arrangement could attend, split class approaches for some or all age groups and combinations thereof. Moreover, we can account for additional hygiene measures whose effect is estimated inside the model.

3. We model the evolution of the pandemic and all enacted policies since the start of the second wave. Since the vast majority of cases has occurred in that period of time and we also model unobserved infections, our simulations take into account that many people are already immune because they have recovered from an infection. Simulating most of the pandemic is important because immunity is not spread randomly across the population.

4. We have an extremely detailed model of PCR and rapid tests with a share of detected cases that varies over time and across age groups.

5. Our model is designed to combine information from many different data sources. Examples are surveys on rapid test demand (Betsch et al., 2021), reaction to test results (Betsch et al., 2021), contact diaries (Mossong et al., 2008), share of detected cases (Paul et al., 2020) and many more.

2.3 Model Overview

At the core of our agent-based model are physical contacts between heterogeneous agents (Figure 2.3.1). Each contact between an infectious individual and somebody susceptible to the disease bears the risk of transmitting the virus. Contacts can happen regularly (e.g. work colleagues) or occur at random. Empirical applications can take the population and household structure from census data and the network-specific frequencies of contacts from diary data measuring contacts before the pandemic (e.g. Mossong et al., 2008; Hoang et al., 2019). Within each network, meeting frequencies depend on age and geographical location (see Section 3.2.4). In our empirical application we distinguish four types of contacts: Within the household, at work, at school or in other settings (leisure activities, grocery shopping, medical appointments, etc.) but other contact networks are also possible.

The four contact networks are chosen so that common NPIs can be modeled in detail. NPIs affect the number of contacts or the risk of transmitting the disease upon having physical contact. The effect of different NPIs will generally vary across contact types. For example, a mandate to work from home will reduce the number of work contacts to zero for a fraction of the working population. Schools and daycare can be closed entirely, operate at reduced capacity (including an alternating schedule) or implement mitigation measures like face mask requirements or air filters (Lessler et al., 2021). Curfews may reduce the number of contacts in settings outside of work and school. In any setting, measures like face mask requirements would reduce the probability of infection associated with a contact (Cheng et al., 2021). The model is implemented in such a general way that it allows fine-grained and sophisticated

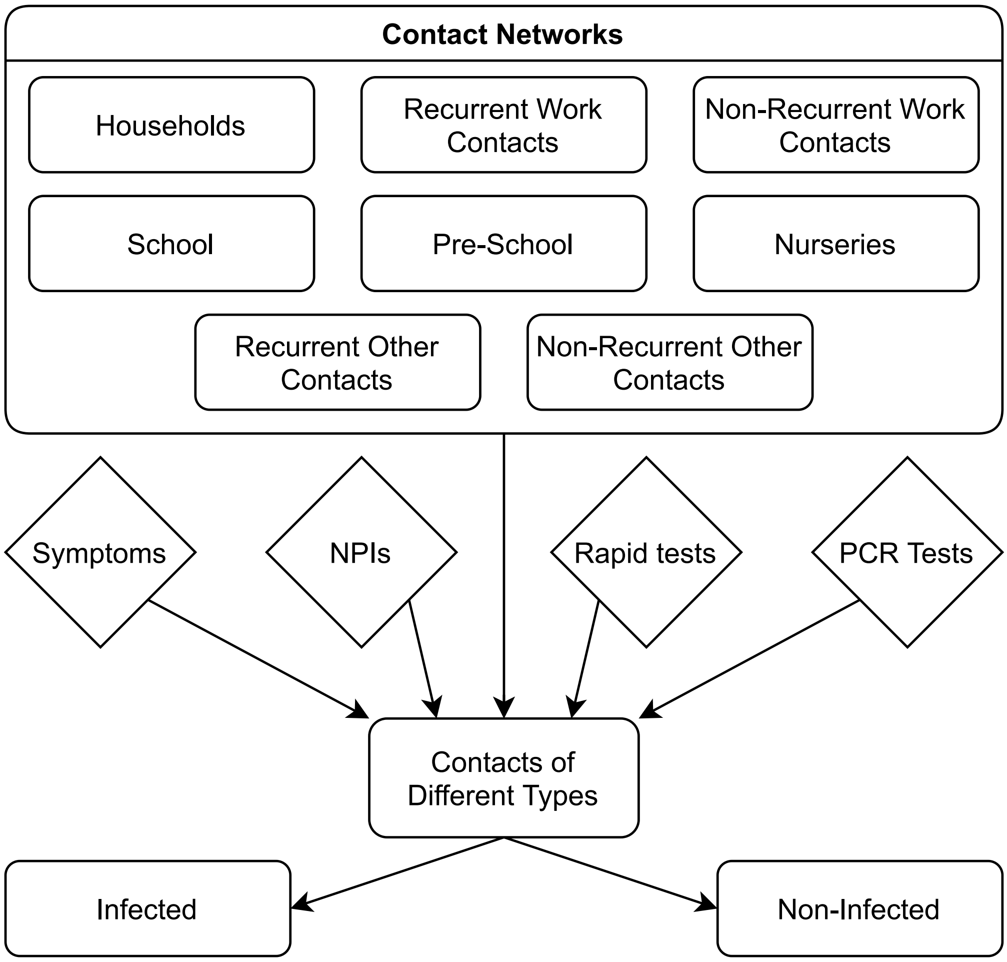


Figure 2.3.1. Model of Contacts and Infections

Note: The number of contacts an agent has depends on various variables, including the agent's characteristics. Demographic characteristics set the baseline number of contacts in different networks (η). The agents may reduce the number of contacts due to NPIs, showing symptoms or testing positively for SARS-CoV-2 (τ). Infections can occur when a susceptible and an infectious agent meet.

policies that condition the number of contacts on observable characteristics, risk contacts or health states.

The number of contacts is translated into infections by a matching algorithm. There are different matching algorithms for recurrent contacts (e.g. classmates, family members) and non-recurrent contacts (e.g. clients, contacts in supermarkets). All types of contacts can be assortative with respect to geographic and demographic characteristics.

When a contact involves an infectious agent, the other person can become infected. The infection probability varies with contact type, age of the susceptible person, the virus strain of the infected person and follows a seasonal pattern. The strength of the seasonality effect is higher for contacts that are easily moved to an outside location in summer (such as leisure contacts) and smaller for contacts that take place inside even in summer (e.g. work contacts).

Infections progress as shown in Figure 2.3.2. We differentiate between an initial period of infection without being infectious or showing symptoms, being infectious (presymptomatic or asymptomatic), showing symptoms, requiring intensive care, and recovery or death (similar to Aleta et al., 2020). The probabilities of transitioning between these states depend on age; their duration is random and calibrated to medical literature (for a detailed description, see Section 3.2.1). Conditional on the type of contact, infectiousness is independent of age (Jones et al., 2021).

We include two types of tests. Polymerase chain reaction (PCR) tests accurately reveal whether an individual is infected or not; there is no uncertainty to the result. PCR tests require at least one day to be processed and there are aggregate capacity constraints. In contrast, rapid antigen tests yield immediate results. Specificity and sensitivity of these tests is set according to data analyzed in Toptan et al. (2020), Brümmer et al. (2021), and Smith et al. (2021); sensitivity depends on the timing of the test relative to the onset of infectiousness.

In general, all information on all individuals – including their health states, time since last rapid and PCR test – can be used to specify which individuals demand a rapid test in each period. Figure 2.9.2 shows our specification for rapid test demand. Schools will require staff and students to be tested regularly. Rapid tests may be offered by employers to on-site workers. Individuals can demand tests for private reasons, which include having plans to meet other people, showing symptoms of CoViD-19, and a household member having tested positively for the virus. We endow each agent with an individual compliance parameter. This parameter determines when she takes up rapid tests. Positive rapid tests, just like symptoms, cause many individuals to reduce their number of contacts and demand a confirmatory PCR test.

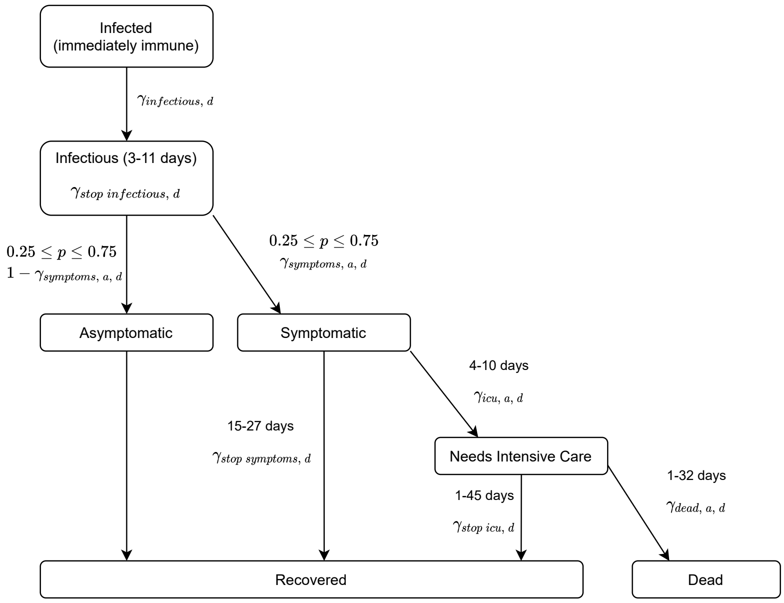


Figure 2.3.2. Disease Progression

Note: All infected individuals become infected within a few days of the transmission contact. Depending on their age, individuals face different probabilities to become symptomatic, require intensive care and die. Numbers of days at arrows give the possible durations to transition from one stage to the next. For example, symptomatic individuals that won't need intensive care show symptoms for 15-27 days before they recover. Probabilities are denoted by *p*.

Modeling PCR test demand is very important as only positive PCR tests are counted as official case numbers. Thus, it is very important for estimating the model parameters (see Section 3.3.1). Furthermore the share of cases that are recorded as official cases can be expected to vary over time and between age groups given CoViD-19's different rate of causing symptoms across age groups and a limited capacity for PCR tests. Our model supports modeling of PCR tests through demand, allocation and processing. However, given data limitations, we model recorded cases with a simpler model, which is depicted in Figure 2.9.1. We take mortality-based aggregate estimates of the share of detected cases and use data on the share of PCR tests administered because of CoViD-19 symptoms. As the share of asymptomatic individuals varies by age group, this gives us age-specific detection rates (the resulting shares of detected infections in our empirical application are shown in Figure 3.G.1). Once rapid tests become available demand for PCR tests can also stem from individuals wanting to confirm a positive rapid test.

The model includes several additional features, which are crucial to describe the evolution of the pandemic in 2020-2021. New virus strains with different infectiousness profiles may appear. Vaccines may become available. During the vaccination roll-out, priority may depend on age and occupation; vaccine hesitancy is modeled by some individuals refusing vaccination offers. With some probability, vaccinated agents become immune and do not transmit the virus (Hunter and Brainard, 2021; Levine-Tiefenbrun et al., 2021; Petter et al., 2021; Pritchard et al., 2021).

Modeling a population of agents according to actual demographic characteristics means that we can use a wide array of data to calibrate the model's many parameters (for an overview over the model parameters, see Appendix 2.A). The remaining parameters can be estimated by choosing them to match features of the data (see McFadden, 1989, for the general method). Both the calibration and estimation is discussed in detail in the next Chapter which presents the empirical application of the model.

The model has a modular structure and can easily be extended to distinguish more contact types, add more stages to the disease progression, implement new policies or test demand models. The main bottleneck is not complexity or computational cost but the availability of data to calibrate additional model features.

2.4 Numbers of Contacts

Consider a hypothetical population of 1,000 individuals in which 50 were infected with a novel infectious disease. From this alone, it is impossible to say whether: (1) Only those 50 people had contact with an infectious person and the disease has an infection probability per contact (β) of one or (2) Everyone met one infectious person but the disease has an infection probability of only five percent per contact. SEIR models do not distinguish between the number of contacts (η) and the infectiousness of each contact (β). Instead, they combine the two into one parameter that is not invariant to social distancing policies.

To model social distancing policies, we need to disentangle the effects of the number of contacts of each individual and the effect of mostly policy-invariant infection probabilities specific to each contact type. The number of types of contacts in our model can be easily extended. Each type of contacts is described by a function that maps individual characteristics, health states and the date into a number of planned contacts for each individual. This allows to model a wide range of contact types.

In our empirical application we distinguish the following contact types that are depicted in Figure 2.3.1 and can be further grouped into the categories household, work, education and others:

- **Households**: Each household member meets all other household members every day.
- **Recurrent work contacts**: These capture contacts with coworkers, repeating clients and superiors. Some of these recurrent contacts take place on every workday, others just once per week. The contacts are assortative in geographical location and age.
- **Non recurrent work contacts**: Working adults have contacts with randomly drawn other people, which are assortative in geographical location and age.
- **Schools**: Students meet all of their classmates every day. Class sizes are calibrated to be representative for Germany and students have the same age and mostly live in the same county. Schools are closed on weekends and during vacations, which vary by federal state. School classes also meet six teachers every day and some of the teachers meet each other.
- **Preschools**: Children who are between three and six years old attend preschool. Each group consists of nine children of mixed ages and two adults who live mostly in the same county. They all meet each other every work day when there are no vacations.
- **Nurseries**: Children younger than three years may attend a nursery and interact with one adult and other children there. The age of the children varies within groups but all live in the same county. They all meet each other every work day when there are no vacations.
- **Non recurrent other contacts**: Contacts with randomly drawn other people, which are assortative with respect to geographic location and and age group. This contact type reflects contacts during leisure activities, grocery shopping, medical appointments, etc.

- **Recurrent other contacts** representing contacts with friends, neighbours or family members who do not live in the same household. Some of these contacts happen daily, others only once per week. They are assortative in geographic location and age.

2.5 Non-Pharmaceutical Interventions (NPIs)

Our model makes it easy to model a wide range of NPIs, either in isolation or simultaneously. This is important for two reasons: Firstly, it allows to predict and quantify the effect of novel NPIs. Secondly, it allows to model the actually implemented policy environment in great detail, which is necessary to use the full time series of infections and fatality rates to estimate the model parameters.[2]

Instead of thinking of policies as completely replacing how many contacts people have, it is more helpful to think of them as measures that adjust the pre-pandemic number of contacts. Therefore, we implement policies as a step that happens after the pre-pandemic number of contacts is calculated but before individuals are matched.

On an abstract level, a policy is a function that modifies the number of contacts of one contact type.[3] This function can be random or deterministic. For example, school closures simply set all school contacts to zero. A work from home mandate leads to a share of workers staying home every day whereas those who cannot work from home are unaffected. Hygiene measures at work randomly reduce the number of infectious contacts for all workers who still go to work.

Policies can also interact. For example, school vacations are temporally reducing school contacts to zero while at the same time increasing other contacts to account for increased leisure activities and family visits during this time. This is important to reproduce the finding that school vacations do not reduce infection numbers even though schools lead to infections when open (Isphording, Lipfert, and Pestel, 2021).

The most complex policies are typically found in the education sector. Since spring 2020 schools have switched back and fourth between full closures, split class approaches with alternating schedules for some or all age groups and reopening while maintaining hygiene measures in Germany. On top of that there were different policies for allowing young students whose parents work full time to attend school even

2. See Avery et al. (2020) for an explanation why it can be harmful to use too long time series to estimate simple SEIR type models.

3. NPIs that reduce the infection risk such as face mask requirements are equivalent to reducing the number of contacts with risk of transmission.

on days where they normally would not. The detailed calibration is shown in Appendix 3.B.

Importantly, policies can depend on the health states of participating individuals. For example children rarely go to school when they have symptoms. It would even be possible to quarantine entire school classes if one student tested positive and to do many other forms of contact tracing. For an application of our model showcasing private contact tracing in the context of the Christmas holidays, see Gabler et al. (2020).

Not everything that reduces contacts compared to the pre-pandemic level is driven by NPIs. Therefore, we also model endogenous contact reductions that depend on the health state of individuals. Many other possible factors, such as the local incidence, are also possible. The extent to which contacts are reduced can be calibrated from surveys.

2.6 Matching Individuals

In the previous sections we only explained how we model the number of contacts each person has. In order to simulate transmissions of CoViD-19, the numbers of contacts have to be translated into actual meetings between people. This is achieved by a matching algorithm:

As described in section 2.4, some contact types are recurrent (i.e. the same people meet regularly), others are non-recurrent (i.e. it would only be by accident that two people meet twice). The matching process is different for recurrent and non recurrent contact models.

Recurrent contacts are described by two components: 1) A set of time invariant groups, such as school classes or groups of co-workers. Those groups are generated once at the beginning of the simulation. The groups can be sampled from empirical data or created by randomly matching simulated individuals into groups. 2) A deterministic or random function that takes the values 0 (non-participating) and 1 (participating) and can depend on the weekday, date and health states of the entire population. These can be used to account for things like vacations, weekends or symptomatic people who stay home.

Given these two components, the disease transmission for recurrent contacts is extremely simple: On each simulated day, every person who does not stay home meets all other group members who do not stay home. If there is a contact between individual i who is infected with virus variant v and infectious and individual j who is in age group a and susceptible, then j becomes infected with the following probability

$$P(infection) = \beta_c \cdot s_{c,t} \cdot \sigma_v \cdot \zeta_a \tag{2.6.1}$$

where β_c denotes the base infection probability of contact type c, $s_{c,t}$ is a seasonality factor between zero and one that depends on the contact type c and time t (see Equation 2.7.1), σ_v is the infectiousness factor of virus variant v and ζ_a is an age dependent susceptibility factor.

The assumption that all group members have contacts with all other group members is not fully realistic, but a good approximation to reality, especially in light of the role of aerosol transmission for CoViD-19 (Anderson et al., 2020; Morawska et al., 2020). Alternatively, the infection probability of recurrent contact types can be interpreted as being the product of a true infection probability and the probability that an actual contact takes place.

The matching of non-recurrent contact types is more difficult because the contact network is resampled every day. Moreover, it needs to allow for assortative matching. In our application, all random contacts are assortative with respect to age group a (it is usually more likely to meet people from the same age group) and county (it is more likely to meet people from the same county) but in principle any set of discrete variables can be used. This set of variables that influence matching probabilities induces a discrete partition of the population into groups.

Below, we first describe one iteration of a simplified matching algorithm that illustrates what we want to achieve. In practice, we approximate the result of this matching algorithm by a two stage sampling procedure that is computationally more efficient. The matching is done for each non-recurrent contact type c. The following step is repeated until no individual has unmatched contacts left. Let z be an iteration counter for the matching algorithm and i denote the individual whose unmatched contacts we are trying to match.

Let $K_{z,i,c}$ denote the number of unmatched contacts of individual i of contact type c before iteration z is completed. Note that $K_{z,i,c} \leq n_{ic}$ which is the total number of contacts individual i has of type c.

Let a_i denote i's age group and $county_i$ her county of residence.

We first draw individual j from the distribution defined by probability mass function F_z over individuals $j \neq i$ in the synthetic population where the probability f_{zj} is calculated as follows:

$$f_{zj} = \underbrace{\alpha_{c,a_i,a_j,county_i,county_j}}_{\text{Group Probability}} \cdot \underbrace{\frac{K_{z,j,c}}{\sum_{l=1}^{N} K_{z,l,c} \cdot \mathbb{I}_{county_l=county_j \wedge a_l=a_j}}}_{\text{Individual Probability}} \tag{2.6.2}$$

We then draw an individual j. If one of the two participants is susceptible and the other one is infectious, we sample whether an infection takes place. The success probability for this event is calculated as in Equation 2.6.1. Finally we update the remaining numbers of unmatched contacts by setting:

$$K_{z+1,i,c} = K_{z,i,c} - 1 \tag{2.6.3}$$
$$K_{z+1,j,c} = K_{z,j,c} - 1 \tag{2.6.4}$$

The runtime of this algorithm scales roughly cubic in the number N of simulated individuals. This is because the number of iterations is proportional to N, in each iteration we have to evaluate Equation 2.6.2 N times and each evaluation of that equation entails a sum over N individuals.

This makes it prohibitively expensive. We therefore replace the above algorithm by a two stage sampling procedure, where we first sample the group from which individual j will be drawn according to the group probabilities defined in Equation 2.6.2. Next we sample an individual from this group with the individual probabilities defined in Equation 2.6.2.

Thus, while the calculation of any given second stage probability entails exactly the same number of calculations as before we do not have to calculate a second stage probability for all simulated individuals but only for those who are members of the group that was sampled in the first stage.

It is easy to see that ex-ante the probability of being sampled are identical between the two stage sampling and the one stage sampling. The only drawback is that towards the end of the matching process it becomes possible to sample a group in which no unmatched contacts are left. In our empirical application this happens extremely rarely. This is so for two reasons: Firstly, the first stage sampling probabilities have been estimated from the same dataset as the numbers of contacts so there cannot be any mismatches such as for example a group that has a low probability of being sampled in the first stage but where all members have many contacts. Secondly, the group sizes are relatively large and we go over individuals in random order. Therefore, groups where no unmatched contacts remain only occur very late in the matching process.[4]

2.7 Seasonality

It is widely acknowledged that the transmission of SARS-CoV-2 is subject to seasonal influences. Infectiousness is increased in winter when most contacts take place inside and the immune system is weakened by low levels of vitamin D, dry air and large temperature swings. For a detailed overview of possible drivers see Kronfeld-Schor et al. (2021).

We follow Kühn et al. (2021) and Gavenčiak et al. (2021) in modeling seasonality in the transmission of SARS-CoV-2 as a multiplicative factor on the infection probability. The factor follows a sine curve that reaches its maximum at January 1st and its minimum on June 30.

For simplicity we normalize the factor to reach one at its maximum. Thus, the formula of the seasonality factor is given by:

$$s_{c,t} = 1 + 0.5\kappa_c sin\left(\pi\left(\frac{1}{2} + \frac{t}{182.5}\right)\right) - 0.5\kappa_c \tag{2.7.1}$$

Where κ_c is difference in the seasonality factor between peak infectiousness and lowest infectiousness.

The subscript c is needed because the strength of the seasonality effect differs across contact types: Work, household and school contacts are likely to take place inside even in summer. Thus, they are only subject to seasonality due to factors that influence the immune system. Other contacts (for example meeting friends and while doing leisure activities) are mostly happening outside in the summer. Therefore, transmission via those contacts should have a stronger seasonal pattern.

2.8 Course of Disease

The disease progression in the model is fairly standard. It is depicted in Figure 2.3.2.

4. If unmatched contacts were a concern one could simply use the fast two stage sampling process for a first pass over contacts and then match all remaining contacts with the slow algorithm.

First, infected individuals will become infectious after one to five days. Overall, about one third of people remain asymptomatic. The rest develop symptoms about one to two days after they become infectious. Modeling asymptomatic and pre-symptomatic cases is important because those people do not reduce their contacts nor do they have an elevated probability to demand a test – unless a test alerts them to their infection status. Thus, they can potentially infect many other people (Donsimoni et al., 2020). The probability to develop symptoms with CoViD-19 is highly age dependent with 75% of children not developing clinical symptoms (Davies et al., 2020).

A small share of symptomatic people will develop strong symptoms that require intensive care. The probability to require intensive care is age-dependent. An age-dependent share of intensive care unit (ICU) patients will die after spending up to 32 days in intensive care. Moreover, if the ICU capacity was reached, all patients who require intensive care but do not receive it die.

It would be easy to make the course of disease even more fine-grained. For example, we could model people who require hospitalization but not intensive care. We opted against this for two reasons: Firstly, only the intensive care capacities were feared to become a bottleneck in Germany. Secondly, there is little expected effect of such a more fine-grained disease progression on infections as we model that symptomatic individuals reduce their contacts and infectiousness usually lasts less than ten days.

We allow the progression of the disease to be stochastic in two ways: Firstly, state changes only occur with a certain probability (e.g. only a fraction of infected individuals develops symptoms). Secondly, the number of periods for which an individual remains in a state is drawn randomly. The parameters that govern these processes are taken from the literature and detailed in Section 3.2.1. For an overview of our disease progression parameters see Table 2.A.3.

2.9 Testing

Having a realistic model of PCR and rapid tests is crucial for two reasons: Firstly, only via a testing model can the simulated infections from the model be made comparable to official case numbers. Secondly, individuals with undetected or not yet detected infections are an important driver of the pandemic.

In principle, our modeling approach is flexible enough to incorporate mechanistic test demand, allocation and processing models. However, there is not enough data available to calibrate such a mechanistic model. Therefore, we build a simpler model of PCR and rapid tests. This model can be calibrated with available data on test demand and availability and – nevertheless – produce a share of detected cases that varies over time and across age groups. The simplified model also agrees with other estimates over the time periods where they are available. An overview of our testing model is shown in Figure 2.9.1.

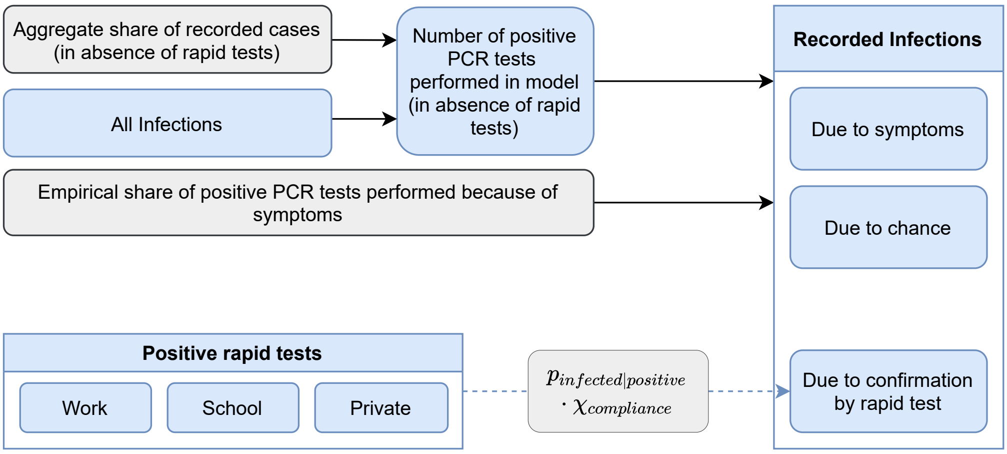


Figure 2.9.1. Translating Infections Into Official Cases

Note: The model relies on the endogenous total number of infections, the share of detected cases and positive rapid tests to translate all infections in the simulated data to age-specific detected infections. The model uses data on the aggregate share of detected cases (ψ), the share of positive PCR tests triggered by symptoms ($\chi_{symptom}$), and the false positive rate of rapid tests ($p_{positive|infected,\ i,\ t}$). The lower part of the graph is relevant only for periods where rapid tests are available.

PCR tests are modeled since the beginning of the simulation and determine whether an infection is officially recorded. Rapid tests are only added at the beginning of 2021. Positive rapid tests do not enter official case numbers directly but most people with a positive rapid test demand a confirmatory PCR test. However, positive rapid tests can have a strong effect on the infection dynamic because they trigger contact reductions and additional rapid tests.

During 2020 people can demand PCR tests either because they have symptoms or randomly. The probability that a PCR test is performed in each of the two situations depends on the number of new infections and the number of available tests. Thus, it varies strongly over time and is unknown.

To distribute the correct number of PCR tests among symptomatic and asymptomatic infections without knowing explicit test demand probabilities, we use the following approach: First, we calculate the total number of positive PCR tests by multiplying the number of newly infected individuals with an estimate of the share of detected cases from the Dunkelzifferradar project (Paul et al., 2020). Next, we determine how many of these tests should go to symptomatic and asymptomatic individuals from data by the RKI (Robert Koch-Institut, 2020). Then, we sample the individuals to which those tests are allocated from the pools of symptomatic and not (yet) symptomatic and untested infected individuals.

Sampling uniformly from the pool of symptomatic individuals ensures that age groups who are more likely to develop symptoms are also more likely to receive tests. Thus, the share of detected cases is much higher for the elderly than for children during times where many tests are done because of symptoms (see Figure 3.G.1 for the age specific shares of detected cases that arise endogenously in our empirical application).

At the beginning of 2021, two challenges arise: Firstly, the externally estimated share of detected cases from the Dunkelzifferradar project (Paul et al., 2020) can no longer be used because it is based on the case fatality rate which changes drastically once the elderly become increasingly vaccinated. Secondly, rapid tests become available at a large scale.

We solve the first challenge by assuming that the share of detected cases would have remained at the level it reached before Christmas if rapid tests had not become available. While this is only an approximation to reality, changes in the share of detected cases that would have happened without rapid tests are very likely to be small compared to the changes caused by rapid tests. In addition, the number of PCR tests per 1000 remains in the same 1.5 to 2.5 window without a visible trend over our estimation period (Roser et al., 2021) which is in line with less than 0.08% of the population receiving a positive rapid test on any given day in our model.

The second challenge is solved by mechanistic rapid test demand models for the workplace, schools and by private individuals. An overview of our rapid test demand model is given in Figure 2.9.2. The calibration of these models is described in Section 3.2.6.

In contrast to PCR tests, rapid tests are not perfect and can be falsely positive and falsely negative. While the specificity of rapid tests is calibrated at 99.4% (Brümmer et al., 2021), their sensitivity strongly depends on the timing of the rapid test relative to the start of infectiousness (Smith et al., 2021). Before the onset of infectiousness the sensitivity is low (35%). On the first day of infectiousness it is much higher (88%) but still lower than during the remaining infectious period (92%). After infectiousness stops, the sensitivity drops to 50%. Modeling the diagnostic gap before and at the beginning of infectiousness is very important to address concerns that rapid tests are too unreliable to serve as screening devices.

We do not distinguish between self administered rapid tests and those that are administered by medical personnel. While there were concerns that self administered tests are less reliable, Lindner et al. (2020) find no basis for this concern.

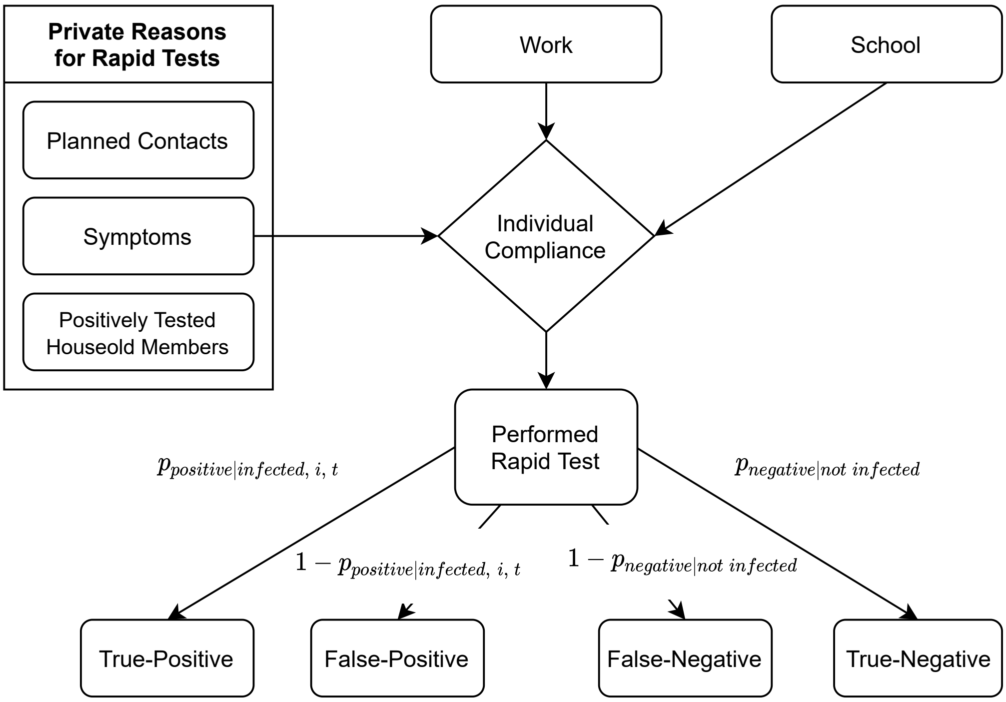


Figure 2.9.2. Model of Rapid Tests

Note: Individuals may perform rapid tests either in preparation to planned leisure contacts, in response to a household member that has tested positive, to own symptoms without access to a PCR test or because the individual plans to go to work or attend school. All reasons trigger a test only for a fraction of individuals depending on an individual compliance parameter; the thresholds for triggering test demand differ across reasons and they may depend on calendar time ($\pi_{c,t}$ and $\tau_{c,t}$). The i, t subscript of the probabilities denote that the sensitivity of the rapid tests depends on the timing relative to the start of infectiousness.

While rapid tests do not directly enter official case numbers, 82% ($\chi_{confirmation}$) of positively tested individuals seek a PCR test (Betsch et al., 2021). Importantly, those PCR tests are made in addition to the tests that would have been done otherwise. Appendix 3.G discusses the effect of rapid tests on the share of detected cases.

2.10 Initial Conditions

Consider a situation where one wants to simulate a period set amidst the pandemic. It means that several thousands of individuals should already have recovered from the disease, be infectious, symptomatic or in intensive care at the start of the simulation. Additionally, the sample of infectious people who will determine the course of the pandemic in the following periods is likely not representative of the whole population because of differences in behavior (number of contacts, assortativity), past policies (such as school closures), etc. The distribution of health states in the population at the beginning of the simulation is called initial conditions.

To come up with realistic initial conditions, we match reported infections from official data to simulated individuals by age group and county. We use one month of data to generate initial conditions with all possible health states. Meanwhile, health states evolve until the beginning of the simulation period without simulating infections by contacts. We also correct reported infections for a reporting lag and scale them up by the share of detected cases to arrive at the true number of infections.

Once vaccinations are being rolled out, these are also already distributed during the month before the actual simulation start to match the share of vaccinated individuals at the start of the simulation.

2.11 Conclusion

This chapter has presented an agent-based simulation model that is designed to predict and evaluate the effect of a variety of policies to contain the CoViD-19 pandemic, including school policies, work policies, vaccination campaigns and rapid test policies.

At the core of the model are physical contacts between heterogeneous agents who can vary in their age, the systemic relevance of their work, number of co-workers and weekly leisure contacts, degree of their compliance with rapid testing policies and many more. Contacts occur in different networks, such as school or work, and can be random or recurrent. The realistic implementation of different contact types allows an easy implementation of non-pharmaceutical interventions. Once rapid tests become available they can be used as entry tests for school, work and some other contacts. They can also serve as diagnostic tests in addition to the usual but slower PCR tests. Individuals can react to rapid tests – as to symptoms and PCR tests – with self-isolation and seek rapid tests in response to events such as a household member being tested positive.

Our model is unique in the sophisticated way in which it models case detection. This allows us to have different detection rates across age groups and over time. Despite its large number of parameters, the model is designed in such a way that most parameters can be calibrated from different literatures and surveys and some even directly from decrees and laws. Due to the ability to include different and changing non-pharmaceutical interventions and the detailed translation of total infections into observed cases the few remaining parameters can be estimated using long-running time series data of the number of infections. A smart matching algorithm makes such an estimation computationally feasible.

In the following chapter this model is applied to the situation in Germany from September 2020 until June 2021. It shows that the model can be estimated with empirical data and can closely recreate the evolution of the pandemic over a long period of time, including the emergence of a new variant, the introduction and expansion of rapid test policies and a vaccination campaign. We show the role the later two – as well as seasonality – played during the stark decrease in cases which Germany experienced in the spring of 2021.

References

Abueg, Matthew, Robert Hinch, Neo Wu, Luyang Liu, William Probert, Austin Wu, Paul Eastham, Yusef Shafi, Matt Rosencrantz, Michael Dikovsky, Zhao Cheng, Anel Nurtay, Lucie Abeler-Dörner, David Bonsall, Michael V. McConnell, Shawn O'Banion, and Christophe Fraser. 2021. "Modeling the effect of exposure notification and non-pharmaceutical interventions on COVID-19 transmission in Washington state." *npj Digital Medicine* 4 (1): [55]

Aleta, Alberto, David Martín-Corral, Ana Pastore y Piontti, Marco Ajelli, Maria Litvinova, Matteo Chinazzi, Natalie E. Dean, M. Elizabeth Halloran, Ira M. Longini Jr, Stefano Merler, Alex Pentland, Alessandro Vespignani, Esteban Moro, and Yamir Moreno. 2020. "Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19." *Nature Human Behaviour* 4 (9): 964–71. [55, 58]

Anderson, Elizabeth L., Paul Turnham, John R. Griffin, and Chester C. Clarke. 2020. "Consideration of the Aerosol Transmission for COVID-19 and Public Health." *Risk Analysis* 40 (5): 902–7. [64]

Avery, Christopher, William Bossert, Adam Clark, Glenn Ellison, and Sara Fisher Ellison. 2020. "An Economist's Guide to Epidemiology Models of Infectious Disease." *Journal of Economic Perspectives* 34 (4): 79–104. [62]

Basurto, Alessandro, Herbert Dawid, Philipp Harting, Jasper Hepp, and Dirk Kohlweyer. 2020. "Economic and Epidemic Implications of Virus Containment Policies: Insights from Agent-Based Simulations." *SSRN Electronic Journal*, [55]

Betsch, Cornelia, Lars Korn, Lisa Felgendreff, Sarah Eitze, Philipp Schmid, Philipp Sprengholz, Lothar Wieler, Patrick Schmich, Volker Stollorz, Michael Ramharter, Michael Bosnjak, Saad B. Omer, Heidrun Thaiss, Freia De Bock, and Ursula Von Rüden. 2021. "COVID-19 Snapshot Monitoring (COSMO Germany) - Wave 41." [56, 71]

Bicher, Martin, Claire Rippinger, Christoph Urach, Dominik Brunmeir, Uwe Siebert, and Niki Popper. 2021. "Evaluation of Contact-Tracing Policies against the Spread of SARS-CoV-2 in Austria: An Agent-Based Simulation." *Medical Decision Making*, 0272989X2110133. [55]

Brümmer, Lukas E., Stephan Katzenschlager, Mary Gaeddert, Christian Erdmann, Stephani Schmitz, Marc Bota, Maurizio Grilli, Jan Larmann, Markus A. Weigand, Nira R. Pollock, Sergio Carmona, Stefano Ongarello, Jilian Sacks, and Claudia M. Denkinger. 2021. "The accuracy of novel antigen rapid diagnostics for SARS-CoV-2: a living systematic review and meta-analysis." *medRxiv*, (3): [58, 69]

Cheng, Yafang, Nan Ma, Christian Witt, Steffen Rapp, Philipp S. Wild, Meinrat O. Andreae, Ulrich Pöschl, and Hang Su. 2021. "Face masks effectively limit the probability of SARS-CoV-2 transmission." *Science*, eprint: https://science.sciencemag.org/content/early/2021/05/19/science.abg6296.full.pdf. [56]

Contreras, Sebastian, Jonas Dehning, Sebastian B. Mohr, Simon Bauer, F. Paul Spitzner, and Viola Priesemann. 2021. "Low case numbers enable long-term stable pandemic control without lockdowns." *Science Advances* 7 (41): eabg2243. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.abg2243. [53]

Cuevas, Erik. 2020. "An agent-based model to evaluate the COVID-19 transmission risks in facilities." *Computers in Biology and Medicine* 121 (6): 103827. [54]

Davies, Nicholas G., Petra Klepac, Yang Liu, Kiesha Prem, Mark Jit, and Rosalind M. Eggo. 2020. "Age-dependent effects in the transmission and control of COVID-19 epidemics." *Nature Medicine* 26 (8): 1205–11. [67]

Delli Gatti, Domenico, and Severin Reissl. 2020. "ABC: An Agent Based Exploration of the Macroeconomic Effects of Covid-19." *CESifo Working Paper Series*, [55]

Donsimoni, Jean Roch, René Glawion, Bodo Plachter, and Klaus Wälde. 2020. "Projecting the spread of COVID-19 for Germany." *German Economic Review* 21 (2): 181–216. [67]

Gabler, Janoś, Tobias Raabe, Klara Röhrl, and Hans-Martin von Gaudecker. 2020. "Die Bedeutung individuellen Verhaltens über den Jahreswechsel für die Weiterentwicklung der Covid-19-Pandemie in Deutschland." Working paper. Institute of Labor Economics (IZA). [63]

Gavenčiak, Tomáš, Joshua Teperowski Monrad, Gavin Leech, Mrinank Sharma, Sören Mindermann, Jan Marcus Brauner, Samir Bhatt, and Jan Kulveit. 2021. "Seasonal variation in SARS-CoV-2 transmission in temperate climates." *medRxiv*, eprint: https://www.medrxiv.org/content/early/2021/06/13/2021.06.10.21258647.full.pdf. [66]

Hinch, Robert, William J. M. Probert, Anel Nurtay, Michelle Kendall, Chris Wymant, Matthew Hall, Katrina Lythgoe, Ana Bulas Cruz, Lele Zhao, Andrea Stewart, Luca Ferretti, Daniel Montero, James Warren, Nicole Mather, Matthew Abueg, Neo Wu, Olivier Legat, Katie Bentley, Thomas Mead, Kelvin Van-Vuuren, Dylan Feldner-Busztin, Tommaso Ristori, Anthony Finkelstein, David G. Bonsall, Lucie Abeler-Dörner, and Christophe Fraser. 2021. "OpenABM-Covid19—An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing." *PLOS Computational Biology* 17 (7): 1–26. [55]

Hoang, Thang, Pietro Coletti, Alessia Melegaro, Jacco Wallinga, Carlos G Grijalva, John W Edmunds, Philippe Beutels, and Niel Hens. 2019. “A systematic review of social contact surveys to inform transmission models of close-contact infections.” *Epidemiology* 30 (5): 723. [56]

Hunter, Paul R., and Julii Brainard. 2021. “Estimating the effectiveness of the Pfizer COVID-19 BNT162b2 vaccine after a single dose. A reanalysis of a study of real-world vaccination outcomes from Israel.” *medRxiv*, (2): [60]

Isphording, Ingo E., Marc Lipfert, and Nico Pestel. 2021. “Does re-opening schools contribute to the spread of SARS-CoV-2? Evidence from staggered summer breaks in Germany.” *Journal of Public Economics* 198 (6): 104426. [62]

Jones, Terry C., Guido Biele, Barbara Mühlemann, Talitha Veith, Julia Schneider, Jörn Beheim-Schwarzbach, Tobias Bleicker, Julia Tesch, Marie Luisa Schmidt, Leif Erik Sander, Florian Kurth, Peter Menzel, Rolf Schwarzer, Marta Zuchowski, Jörg Hofmann, Andi Krumbholz, Angela Stein, Anke Edelmann, Victor Max Corman, and Christian Drosten. 2021. “Estimating infectiousness throughout SARS-CoV-2 infection course.” *Science*, eprint: https://science.sciencemag.org/content/early/2021/05/24/science.abi5273.full.pdf. [58]

Kronfeld-Schor, N., T. J. Stevenson, S. Nickbakhsh, E. S. Schernhammer, X. C. Dopico, T. Dayan, M. Martinez, and B. Helm. 2021. “Drivers of Infectious Disease Seasonality: Potential Implications for COVID-19.” *Journal of Biological Rhythms* 36 (1): 35–54. [66]

Kühn, Martin J., Daniel Abele, Tanmay Mitra, Wadim Koslow, Majid Abedi, Kathrin Rack, Martin Siggel, Sahamoddin Khailaie, Margrit Klitz, Sebastian Binder, Luca Spataro, Jonas Gilg, Jan Kleinert, Matthias Häberle, Lena Plötzke, Christoph D. Spinner, Melanie Stecher, Xiao Xiang Zhu, Achim Basermann, and Michael Meyer-Hermann. 2021. “Assessment of effective mitigation and prediction of the spread of SARS-CoV-2 in Germany using demographic information and spatial resolution.” *Mathematical Biosciences* 339: 108648. [66]

Lessler, Justin, M. Kate Grabowski, Kyra H. Grantz, Elena Badillo-Goicoechea, C. Jessica E. Metcalf, Carly Lupton-Smith, Andrew S. Azman, and Elizabeth A. Stuart. 2021. “Household COVID-19 risk and in-person schooling.” *Science* 372 (6546): 1092–97. [56]

Levine-Tiefenbrun, Matan, Idan Yelin, Rachel Katz, Esma Herzel, Ziv Golan, Licita Schreiber, Tamar Wolf, Varda Nadler, Amir Ben-Tov, Jacob Kuint, Sivan Gazit, Tal Patalon, Gabriel Chodick, and Roy Kishony. 2021. “Initial report of decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine.” *Nature medicine* 27 (5): 790–92. [60]

Lindner, Andreas K., Olga Nikolai, Franka Kausch, Mia Wintel, Franziska Hommes, Maximilian Gertler, Lisa J. Krüger, Mary Gaeddert, Frank Tobian, Federica Lainati, Lisa Köppel, Joachim Seybold, Victor M. Corman, Christian Drosten, Jörg Hofmann, Jilian A. Sacks, Frank P. Mockenhaupt, and Claudia M. Denkinger. 2020. “Head-to-head comparison of SARS-CoV-2 antigen-detecting rapid test with self-collected nasal swab versus professional-collected nasopharyngeal swab.” *European Respiratory Journal* 57 (4): 2003961. [69]

McFadden, Daniel. 1989. “A method of simulated moments for estimation of discrete response models without numerical integration.” *Econometrica: Journal of the Econometric Society*, 995–1026. [55, 60]

Mellacher, Patrick. 2020. “COVID-Town: An Integrated Economic-Epidemiological Agent-Based Model.” MPRA Paper 103661. University Library of Munich, Germany. [55]

**Morawska, Lidia, Julian W. Tang, William Bahnfleth, Philomena M. Bluyssen, Atze Boerstra, Giorgio Buonanno, Junji Cao, Stephanie Dancer, Andres Floto, Francesco Franchimon, Charles Haworth, Jaap Hogeling, Christina Isaxon, Jose L. Jimenez, Jarek Kurnitski, Yuguo Li, Marcel Loomans, Guy Marks, Linsey C. Marr, Livio Mazzarella, Arsen Krikor Melikov, Shelly Miller, Donald K. Milton, William Nazaroff, Peter V. Nielsen, Catherine Noakes, Jordan Peccia, Xavier

Querol, Chandra Sekhar, Olli Seppänen, Shin-ichi Tanabe, Raymond Tellier, Kwok Wai Tham, Pawel Wargocki, Aneta Wierzbicka, and Maosheng Yao. 2020. "How can airborne transmission of COVID-19 indoors be minimised?" *Environment International* 142 (9): 105832. [64]

Mossong, Joël, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. 2008. "Social contacts and mixing patterns relevant to the spread of infectious diseases." *PLoS medicine* 5 (3): [55, 56]

Paul, Christian, Anna Steinberg, Christoph Schaller, Kathy Haemmerl, Leonie Thum, Marcel Parciak, Michael Knaf, Sarina Meyer, Tina Parciak, and Urs Grupp. 2020. *Dunkelzifferradar*. Edited by Prototype Fund. URL: https://prototypefund.de/project/dunkelzifferradar/. [56, 68, 69]

Petter, Ella, Orna Mor, Neta Zuckerman, Danit Oz-Levi, Asaf Younger, Dvir Aran, and Yaniv Erlich. 2021. "Initial real world evidence for lower viral load of individuals who have been vaccinated by BNT162b2." *medRxiv*, (2): [60]

Pritchard, Emma, Philippa C. Matthews, Nicole Stoesser, David W. Eyre, Owen Gethings, Karina-Doris Vihta, Joel Jones, Thomas House, Harper VanSteenHouse, Iain Bell, John I. Bell, John N. Newton, Jeremy Farrar, Ian Diamond, Emma Rourke, Ruth Studley, Derrick Crook, Tim E. A. Peto, A. Sarah Walker, and Koen B. Pouwels. 2021. "Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom." *Nature Medicine*, [60]

Robert Koch-Institut. 2020. *Laborbasierte Surveillance SARS-CoV-2*. URL: https://ars.rki.de/Content/COVID19/Main.aspx. [68]

Roser, Max, Hannah Ritchie, Esteban Ortiz-Ospina, and Joe Hasell. 2021. *Our World in Data: Daily now COVID-19 tests per 1000 people*. URL: https://bit.ly/2Y8AvVm. [69]

Silva, Petrônio C.L., Paulo V.C. Batista, Hélder S. Lima, Marcos A. Alves, Frederico G. Guimarães, and Rodrigo C.P. Silva. 2020. "COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions." *Chaos, Solitons & Fractals* 139 (10): 110088. [54]

Smith, Rebecca L, Laura L Gibson, Pamela P Martinez, Ruian Ke, Agha Mirza, Madison Conte, Nicholas Gallagher, Abigail Conte, Leyi Wang, Richard Fredrickson, Darci C Edmonson, Melinda E Baughman, Karen K Chiu, Hannah Choi, Tor W Jensen, Kevin R Scardina, Shannon Bradley, Stacy L Gloss, Crystal Reinhart, Jagadeesh Yedetore, Alyssa N Owens, John Broach, Bruce Barton, Peter Lazar, Darcy Henness, Todd Young, Alastair Dunnett, Matthew L Robinson, Heba H Mostafa, Andrew Pekosz, Yukari C Manabe, William J Heetderks, David D McManus, and Christopher B Brooke. 2021. "Longitudinal assessment of diagnostic test performance over the course of acute SARS-CoV-2 infection." *Journal of Infectious Diseases*, (6): eprint: https://academic.oup.com/jid/advance-article-pdf/doi/10.1093/infdis/jiab337/38842786/jiab337.pdf. jiab337. [58, 69]

Toptan, Tuna, Lisa Eckermann, Annika E. Pfeiffer, Sebstian Hoehl, Sandra Ciesek, Christian Drosten, and Victor M. Corman. 2020. "Evaluation of a SARS-CoV-2 rapid antigen test: Potential to help reduce community spread?" *J Clin Virol.* [58]

Appendix 2.A Overview Model Parameters

Table 2.A.1. Contacts, Matching and Policies

name	notation	time active	source	description
n_contacts	$\eta_{c,\,n}$	always	Mossong et al. (2008)	probability to have *n* contacts in contact type *c*. The number of non recurrent contacts is drawn for each individual every day. For recurrent contacts individuals are assigned to daily and weekly groups. See Section 3.2.3.
degree_of_assortativity	$\alpha_{c,\,a_i,\,a_j,\,county_i,\,county_j}$	always	Mossong et al. (2008)	probability in contact type *c* that an individual of age group a_i and county $county_i$ meets an individual of age group a_j and county $county_j$. Contacts can be assortative by age group, state and county. See Section 3.2.4 for details.
work_attend_multiplier	$\rho_{w,\,attend,\,t}$	always	Google, LLC (2021)	share of workers that continue to have work contacts on date *t*. We use the reduction in work mobility reported by Google, LLC (2021) as a proxy for the share of workers that work from home. See Section 3.2.5 for details.
hygiene_multiplier	$\rho_{hygiene}$	since Nov 2020	estimated	reduction in transmission during work and educational contacts due to stricter hygiene measures such as wearing face masks. See Section 3.3.1.
other_multiplier	$\rho_{other,\,t}$	always	estimated	reduction in the transmission during other contacts. This incorporates both hygiene measures as well as the reduction of physical meetings. There are nine breakpoints for the whole estimation period. See Section 3.3.1 for details.

Table 2.A.2. Infection Probabilities and Virus Variants

name	notation	time active	source	description
infection_prob	β_c	always	estimated	Base infection probability of contact type *c*. For each contact, this base probability is further adjusted by the seasonality, susceptibility of the contact and the virus strain. See Section 3.3.1 for details.
susceptibility	ζ_a	always	Davies, Klepac, et al. (2020)	Susceptibility to CoViD-19 depends on a person's age group. The higher the age the more easily people become infected. The susceptibility of the oldest age group is normalized to one.
seasonality	κ_c	always	Gavenčiak et al. (2021)	The probability to contract CoViD-19 when exposed depends on the seasonality. Since different contact types are more or less subject to seasonal variation (e.g. by moving contacts outdoors) the seasonality also depends on the contact type. Refer to Section 2.7 for an explanation.
variant_infectiousness	σ_v	always	Davies, Abbott, et al. (2021)	Variant *v*'s infectiousness relative to the wild type.
variant_introduction	$\omega_{v,t}$	time dependent	estimated	Number of infections per 100,000 individuals of variant *v* to be introduced on day *t*. See Section 3.3.1 for details.

Table 2.A.3. Disease and Vaccination Model

name	notation	time active	source	description
p_duration_immune	$\gamma_{immune,\,d}$	always	see Section 3.2.1	probability to stay immune for *d* days after having contracted CoViD-19
p_duration_until_infectious	$\gamma_{infectious,\,d}$	always	see Section 3.2.1	probability to become infectious *d* days after infection
p_duration_of_infectiousness	$\gamma_{stop\ infectious,\,d}$	always	see Section 3.2.1	probability that infectiousness lasts *d* days
p_duration_until_symptoms	$\gamma_{symptoms,\,a,\,d}$	always	see Section 3.2.1	probability for individuals of age group *a* to develop symptoms *d* days (possibly infinite) after becoming infectious
p_duration_of_symptoms	$\gamma_{stop\ symptoms,\,d}$	always	see Section 3.2.1	probability for individuals of age group *a* that symptoms last *d* days
p_duration_until_icu	$\gamma_{icu,\,a,\,d}$	always	see Section 3.2.1	probability for symptomatic individuals of age group *a* to require intensive care *d* days (possibly infinite) after symptom onset
p_duration_of_icu	$\gamma_{stop\ icu,\,a,\,d}$	always	see Section 3.2.1	probability to recover after *d* days from requiring intensive care
p_duration_until_death	$\gamma_{dead,\,a,\,d}$	always	see Section 3.2.1	probability for individuals of age group *a* in intensive care to die after *d* days (possibly infinite)
p_until_immune_by_vaccine	$\gamma_{vacc,\,d}$	2021	see Section 3.2.1	probability to develop immunity *d* days (possibly infinite) after being vaccinated
share_vaccine_refusers	ξ	2021	Frisch (2021)	share of individuals refusing to be vaccinated. See Section 3.2.2

Table 2.A.4. Rapid Testing

name	notation	time active	source	description
rapid_test_specificity	$p_{negative\|not\ infected}$	2021	Brümmer et al. (2021)	the probability of an uninfected person to receive a negative rapid test result. See Section 2.9.
rapid_test_sensitivity	$p_{positive\|infected,\ i,\ t}$	2021	Smith et al. (2021)	the probability of an infected person to receive a positive rapid test result. This depends on the timing of the test relative to the individual's onset of infectiousness. See Section 2.9.
share_accepting_work_rapid_test_offer	$\pi_{w,\ d}$	2021	Betsch et al. (2021)	share of workers that regularly pick up rapid test offers by their employers when they do not work from home. In our baseline specification this is time constant. See Section 3.2.6.
share_workers_receiving_rapid_test_offer	$\pi_{w,\ s,\ t}$	2021	see 3.2.6	share of workers that get a regular rapid test offer by their employer when they do not work from home on date *t*.
share_educ_workers_with_rapid_test	$\pi_{teacher,\ t}$	2021	see 3.2.6	share of educational workers who test themselves as part of their work on date *t*
share_students_with_rapid_test	$\pi_{students,\ t}$	2021	see 3.2.6	share of school pupils that do rapid tests at school on date *t*.
share_private_rapid_test_demand	$\pi_{private,\ t}$	2021	see 3.2.6	share of individuals that do a rapid test when any of the private reason events such as a household member testing positive occur on date *t*.
rapid_test_educ_freq	$\theta_{t,\ educ}$	2021	see 3.2.6	Frequency with which individuals test themselves in educational settings at time *t*. See Section 3.2.6.
rapid_test_work_freq	$\theta_{t,\ work}$	2021	see 3.2.6	Frequency with which complying workers test themselves at time *t*. See Section 3.2.6.

Table 2.A.5. PCR Testing, Case Detection and Behavioral Response

name	notation	time active	source	description
p_duration_until_test_result	$\gamma_{PCR,\, d}$	always	Robert Koch-Institut (2020)	probability that it takes *d* days between the performance of a PCR test and receiving the result. See Section 3.2.7.
share_of_tests_for_symptomatics	$\chi_{symptom,\, t}$	always	calibrated from RKI	share of positive PCR tests that are performed on individuals because of CoViD-19 symptoms. See Section 3.2.7.
share_w_positive_rapid_test_requesting_pcr	$\chi_{confirmation}$	2021	Betsch et al. (2021)	share of individuals with positive rapid test that seek a PCR test. See Section 3.2.7.
share_known_cases_without_rapid_tests	ψ_t	always	Paul et al. (2020)	share of cases that would be detected in the absence of rapid tests (see Section 2.9)
symptomatic_multiplier	$\tau_{symptoms,\, c}$	always	see 3.2.8	share of symptomatic individuals that still have contacts of type *c*.
positive_pcr_multiplier	$\tau_{positive\ PCR,\, c}$	always	see 3.2.8	share of individuals with a positive PCR test that still have contacts of type *c*.
positive_rapid_test_multiplier	$\tau_{positive\ rapid\ test,\, c}$	2021	Betsch et al. (2021)	share of individuals with a recent positive rapid test that still have contacts of type *c*. See Section 3.2.8

Appendix 2.B Reproducibility

The source code used for this paper is open source and available under the MIT License. It is split into two parts

- The source code for the model can be found at https://github.com/covid-19-impact-lab/sid/ and its documentation at https://sid-dev.readthedocs.io.
- The source code for the application to Germany can be found at https://github.com/covid-19-impact-lab/sid-germany/ with a shorter documentation at https://sid-germany.readthedocs.io.

We are grateful to the authors and contributors of the following software packages upon which our software is built: conda (Anaconda, 2016), conda-forge (conda-forge community, 2015) dask (Rocklin, 2015), estimagic (Gabler, 2020), holoviews (Stevens, Rudiger, and Bednar, 2015), matplotlib (Hunter, 2007), numba (Lam, Pitrou, and Seibert, 2015), numpy (Harris et al., 2020), pandas (McKinney, 2010; The pandas development team, 2020), pytask (Raabe, 2020), Python (Van Rossum and Drake Jr, 1995), scipy (Virtanen et al., 2020), and seaborn (Waskom, 2021).

References

Anaconda. 2016. “Anaconda Software Distribution.” Version Vers. 2-2.4.0. [81]

Betsch, Cornelia, Lars Korn, Lisa Felgendreff, Sarah Eitze, Philipp Schmid, Philipp Sprengholz, Lothar Wieler, Patrick Schmich, Volker Stollorz, Michael Ramharter, Michael Bosnjak, Saad B. Omer, Heidrun Thaiss, Freia De Bock, and Ursula Von Rüden. 2021. “COVID-19 Snapshot Monitoring (COSMO Germany) - Wave 41.” [79, 80]

Brümmer, Lukas E., Stephan Katzenschlager, Mary Gaeddert, Christian Erdmann, Stephani Schmitz, Marc Bota, Maurizio Grilli, Jan Larmann, Markus A. Weigand, Nira R. Pollock, Sergio Carmona, Stefano Ongarello, Jilian Sacks, and Claudia M. Denkinger. 2021. “The accuracy of novel antigen rapid diagnostics for SARS-CoV-2: a living systematic review and meta-analysis.” *medRxiv*, (3): [79]

conda-forge community. 2015. “The conda-forge Project: Community-based Software Distribution Built on the conda Package Format and Ecosystem.” [81]

Davies, Nicholas G., Sam Abbott, Rosanna C. Barnard, Christopher I. Jarvis, Adam J. Kucharski, James D. Munday, Carl A. B. Pearson, Timothy W. Russell, Damien C. Tully, Alex D. Washburne, Tom Wenseleers, Amy Gimma, William Waites, Kerry L. M. Wong, Kevin van Zandvoort, Justin D. Silverman, Karla Diaz-Ordaz, Ruth Keogh, Rosalind M. Eggo, Sebastian Funk, Mark Jit, Katherine E. Atkins, and W. John Edmunds. 2021. “Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England.” *Science*, (3): [77]

Davies, Nicholas G., Petra Klepac, Yang Liu, Kiesha Prem, Mark Jit, and Rosalind M. Eggo. 2020. “Age-dependent effects in the transmission and control of COVID-19 epidemics.” *Nature Medicine* 26 (8): 1205–11. [77]

Frisch, Julia. 2, 2021. *Deutsche mit hoher Corona-Impfbereitschaft – aber Wissenslücken*. URL: https://www.aerztezeitung.de/Politik/Deutsche-mit-hoher-Corona-Impfbereitschaft-aber-Wissensluecken-417569.html. [78]

Gabler, Janos. 2020. “estimagic – A Python Tool for the Estimation of (Structural) Econometric Models.” [81]

Gavenčiak, Tomáš, Joshua Teperowski Monrad, Gavin Leech, Mrinank Sharma, Sören Mindermann, Jan Marcus Brauner, Samir Bhatt, and Jan Kulveit. 2021. “Seasonal variation in SARS-CoV-2 transmission in temperate climates.” *medRxiv*, eprint: https://www.medrxiv.org/content/early/2021/06/13/2021.06.10.21258647.full.pdf. [77]

Google, LLC. 2021. “Google COVID-19 Community Mobility Reports.” Working paper. [76]

Harris, Charles R., K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. 2020. “Array programming with NumPy.” *Nature* 585 (7825): 357–62. [81]

Hunter, John D. 2007. “Matplotlib: A 2D graphics environment.” *Computing in science & engineering* 9 (03): 90–95. [81]

Lam, Siu Kwan, Antoine Pitrou, and Stanley Seibert. 2015. “Numba: A llvm-based python jit compiler.” In *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, 1–6. [81]

McKinney, Wes. 2010. "Data Structures for Statistical Computing in Python." In *Proceedings of the 9th Python in Science Conference*. Edited by Stéfan van der Walt and Jarrod Millman, 56–61. [81]

Mossong, Joël, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. 2008. "Social contacts and mixing patterns relevant to the spread of infectious diseases." *PLoS medicine* 5 (3): [76]

Paul, Christian, Anna Steinberg, Christoph Schaller, Kathy Haemmerl, Leonie Thum, Marcel Parciak, Michael Knaf, Sarina Meyer, Tina Parciak, and Urs Grupp. 2020. *Dunkelzifferradar*. Edited by Prototype Fund. URL: https://prototypefund.de/project/dunkelzifferradar/. [80]

Raabe, Tobias. 2020. "pytask – A Python tool for managing scientific workflows." [81]

Robert Koch-Institut. 2020. *Laborbasierte Surveillance SARS-CoV-2*. URL: https://ars.rki.de/Content/COVID19/Main.aspx. [80]

Rocklin, Matthew. 2015. "Dask: Parallel computation with blocked algorithms and task scheduling." In *Proceedings of the 14th python in science conference*. Vol. 130, Citeseer, 136. [81]

Smith, Rebecca L, Laura L Gibson, Pamela P Martinez, Ruian Ke, Agha Mirza, Madison Conte, Nicholas Gallagher, Abigail Conte, Leyi Wang, Richard Fredrickson, Darci C Edmonson, Melinda E Baughman, Karen K Chiu, Hannah Choi, Tor W Jensen, Kevin R Scardina, Shannon Bradley, Stacy L Gloss, Crystal Reinhart, Jagadeesh Yedetore, Alyssa N Owens, John Broach, Bruce Barton, Peter Lazar, Darcy Henness, Todd Young, Alastair Dunnett, Matthew L Robinson, Heba H Mostafa, Andrew Pekosz, Yukari C Manabe, William J Heetderks, David D McManus, and Christopher B Brooke. 2021. "Longitudinal assessment of diagnostic test performance over the course of acute SARS-CoV-2 infection." *Journal of Infectious Diseases*, (6): eprint: https://academic.oup.com/jid/advance-article-pdf/doi/10.1093/infdis/jiab337/38842786/jiab337.pdf. jiab337. [79]

Stevens, Jean-Luc R, Philipp Rudiger, and J Bednar. 2015. "HoloViews: Building complex visualizations easily for reproducible science." In *Proceedings of the 14th Python in Science Conference*. Citeseer, 61–69. [81]

The pandas development team. 2020. *pandas-dev/pandas: Pandas*. Version latest. [81]

Van Rossum, Guido, and Fred L Drake Jr. 1995. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam. [81]

Virtanen, Pauli, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. 2020. "SciPy 1.0: fundamental algorithms for scientific computing in Python." *Nature methods* 17 (3): 261–72. [81]

Waskom, Michael L. 2021. "Seaborn: statistical data visualization." *Journal of Open Source Software* 6 (60): 3021. [81]

Chapter 3

How To Beat SARS-CoV-2? The Role of Rapid Tests, Vaccinations, and NPIs to Contain CoViD-19

Joint with Janoś Gabler, Tobias Raabe and Hans-Martin von Gaudecker

3.1 Introduction

Since early 2020, the CoViD-19 pandemic has presented an enormous challenge to humanity. To contain the spread of the disease and avoid an overburdening of the health care system, most countries implemented non-pharmaceutical interventions (NPIs) aimed at reducing contacts between individuals. These have been updated over time and new ones – such as rapid testing requirements – were introduced as they became available. Neither the effects of these policies nor the effect of seasonal patterns are well understood in quantitative terms.

Given the high speed with which NPIs were implemented and the often joint adjustments of NPIs, empirical evaluations of NPIs are rare.[1] Thus, many debates about

* The authors are grateful for support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany´s Excellence Strategy – EXC 2126/1– 390838866 – and through CRC-TR 224 (Projects A02 and C01), by the IZA Institute of Labor Economics, and by the Google Cloud CoViD-19 research credits program.

1. See Isphording, Lipfert, and Pestel (2021) and Vlachos, Hertegård, and B. Svaleryd (2021) for exceptions in the school context and Berger, Fritz, and Kauermann (2021) and Pavelka et al. (2021) for rapid tests.

policies – such as the degree to which rapid testing can help to stop infections and may even be able to substitute other, more costly, NPIs – had to take place with very little empirical evidence.

In this chapter we evaluate the role of Germany's vaccination campaign and rapid test policies during spring 2021. We do this accounting for and comparing their effects to the effect seasonality played during that time. To do so, we use the agent-based simulation model presented in Chapter 2.

Modeling a population of agents according to actual demographic characteristics means that we can use many different data sources to identify and calibrate the model's many parameters. For example: Contact diaries yield pre-pandemic distributions of contacts for different contact types and their assortativity by age group. Mobility data is used to model the evolution of work contacts. School and daycare policies can be incorporated directly from official directives. Administrative records on the number of tests, vaccinations by age and region and the prevalence of virus strains are generally available. Surveys may ask about test offers, propensities to take them up and past tests. The estimates of other studies concerning the seasonality of infections can be incorporated directly.

The remaining parameters – most notably these include infection probabilities by contact type and the effects of some NPIs – will be chosen numerically so that the model matches features of the data (see McFadden, 1989, for the general method). In our application, we keep the number of free parameters low in order to avoid overfitting. The data features to be matched include official case numbers for each age group and region, deaths and the share of the Alpha variant. We fit the second and third wave of the CoViD-19 pandemic in Germany, covering the period mid-September 2020 to the end of May 2021.

We find that the German vaccination campaign, which prioritized the elderly, only had a small effect on the incidence. Only 16% of the joint decrease caused by rapid tests, seasonality and vaccinations are attributed to vaccinations. Instead our analysis shows that, aside from seasonality, frequent and large-scale rapid testing caused the bulk of the decrease in the spring of 2021. According to our simulations over 40% of the joint effect are due to rapid tests.

Even though we model the imperfections of rapid tests, they are very cost effective. Unless incidences are very low, there are only few false positives per detected case. Because of their effectiveness, we are able to show that stringent rapid testing can serve as substitute for more costly NPIs such as schooling restrictions and work from home mandates. Given the German rapid testing policy of April 2021, opening schools completely would have had only a small effect on case numbers. The same applies to relaxing work from home mandates. However, requiring stringent rapid testing at work – as in schools – could have saved Germany up to 50 cases per million per day.

We now first describe how we calibrate most of the parameters and how we construct our synthetic population in Section 3.2. Section 3.3 then explains which parameters we estimate and how we estimate them and presents the model fit. Section 3.4 presents our main results: the role rapid tests, vaccinations and seasonality have played in the decline of Germany's CoViD-19 cases in the spring of 2021. It then investigates the usefulness of rapid tests as a screening device and lastly whether rapid tests can substitute some more costly NPIs. Section 3.5 concludes.

3.2 Data And Calibrated Parameters

This section discusses all the parameters that we calibrate through surveys and other data sources outside our model. See Appendix 2.A for an overview over all model parameters.

3.2.1 Course of Disease

We start with the calibration of the parameters governing the infection probability and the course of disease. See Figure 2.3.2 for a summary of our disease progression model.[2]

The first stage of any disease is the infection. As detailed in Equation 2.6.1 the infection probability depends on the contact type, the calendar date to determine the seasonality, the age group of the susceptible person and the variant the infectious person is carrying. The base infection probabilities (β_c) are estimated inside our model (Section 3.3.1). We calibrate the strong and weak seasonality parameters as follows: κ_{strong} is set to 0.42 and κ_{weak} to 0.21. This is in line with Gavenčiak et al. (2021) and Kühn et al. (2021). The two resulting seasonality curves are shown in Figure 3.2.1.

For the susceptibility of each age group (ζ_a) we take the estimates of Davies, Klepac, et al. (2020, Extended Data Fig. 4) and for the infectiousness factor of the Alpha variant relative to the wild type (σ_{Alpha}) the estimated 1.67 of Davies, Abbott, et al. (2021).

2. An even more detailed description of our disease progression parameters can be found in our online documentation

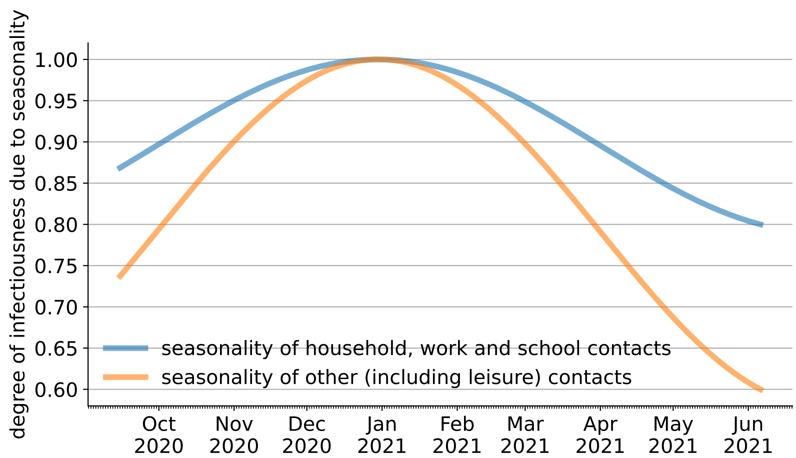


Figure 3.2.1. Seasonality by Type of Contact

Note: We model seasonality as a factor that reduces the probability of infection of all encounters. The factor depends on the day and is calculated from a sinus shaped function with its maximum on January 1st. Since seasonality can affect the transmission both through physical conditions such as temperature and humidity as well as through the numbers of contacts that take place outside we assume two seasonality factors. One for other contacts which we expect to be strongly affected by fairer weather with a maximum reduction of 42% in the infection probability. The other seasonality only makes contacts up to 21% less infectious and is applied to household, work and school contacts.

Some time after infection a person becomes infectious. This time span is called the latent period, which we denote by $\gamma_{infectious}$. Zhao et al. (2021) estimate the latent period to last 3.3 days (95% CI: 0.2, 7.9) on average. In line with this estimate our latent period lasts one to five days.

Once individuals become infectious, a share of them goes on to develop symptoms while others remain asymptomatic. We rely on data by Davies, Klepac, et al. (2020) for the age-dependent probability to develop symptoms. It varies from 25% for children and young adults to nearly 70% for the elderly. Similar to Peak et al. (2020) and in line with He et al. (2020) we set the length of the presymptomatic stage after the onset of infectiousness ($\gamma_{symptoms,a}$) to be one or two days for all age groups. The probability to become symptomatic for every age group a is split equally between one and two days. This combined with our latency period leads to an incubation period that is in line with the meta analysis by McAloon et al. (2020).

We assume that the duration of infectiousness ($\gamma_{stop\ infectious}$) is the same for both symptomatic and asymptomatic individuals as evidence suggests little differences in the transmission rates between symptomatic and asymptomatic patients (Yin and Jin, 2020) and that the viral load between symptomatic and asymptomatic individuals are similar (Byrne et al., 2020; Singanayagam et al., 2020; Zou et al., 2020). Our distribution of the duration of infectiousness is based on Byrne et al. (2020):

For symptomatic cases they arrive at zero to five days before symptom onset (see their figure 2) and three to eight days of infectiousness afterwards.[3] Excluding the most extreme combinations, we arrive at three to eleven days as the duration of infectiousness.

We use the duration to recovery of mild and moderate cases reported by Bi et al. (2020, Figure S3, Panel 2) for the duration of symptoms for non-ICU requiring symptomatic cases ($\gamma_{stop\,symptoms}$). We only disaggregate by age how likely individuals are to require intensive care.

For the time from symptom onset until need for intensive care ($\gamma_{icu,\,a}$) we rely on data by Stokes et al. (2020)) and Hinch et al. (2021). For those who will require intensive care we follow Chen et al. (2020) who estimate the time from symptom onset to ICU admission to be 8.5 ± 4 days. This aligns well with numbers reported for the time from first symptoms to hospitalization: Gaythorpe et al. (2020) report a mean of 5.76 days with a standard deviation of four days. We assume that the time between symptom onset and ICU takes four, six, eight or ten days with equal probabilities.

We take the survival probabilities and time to death and time until recovery ($\gamma_{stop\ icu\ a}$ and $\gamma_{dead,\,a}$) from intensive care from Hinch et al. (2021). They report time until death to have a mean of 11.74 days and a standard deviation of 8.79 days. To match this we discretize that 41% of individuals who will die from CoViD-19 do so after one day in intensive care, 22% die after twelve days, 29% after 20 days and 7% after 32 days. Again, we rescale this for every age group among those that will not survive. For survivors Hinch et al. (2021) reports a mean duration of 18.8 days until recovery and a standard deviation of 12.21 days. We discretize this such that 22% of survivors recover after one day, 30% after 15 days, 28% after 25 days and 18% after 45 days.

Individuals can become immune either through infection (γ_{immune}) or vaccination ($\gamma_{vacc,\,d}$). As reinfections are very rare (Abu-Raddad et al., 2020), we set the immunity period to one year with probability one, i.e. everyone that has been infected enjoys immunity for the rest of the simulation period.

The second route to immunity is through vaccination. Germany has mostly relied on the Pfizer-BioNTech BNT162b2 and Oxford-AstraZeneca ChAdOx1-S vaccines with smaller shares of the Moderna and Johnson&Johnson vaccines (impfdashboard.de, 2021). As Pritchard et al. (2021) and Harris et al. (2021) find no difference in

3. Viral loads may be detected much later but eight days seems to be the time after which most people are culture negative, as also reported by Singanayagam et al. (2020).

the effectiveness between the two most common vaccines, we do not distinguish between vaccines.

Immunity is binary in our model, i.e. individuals achieve either sterile immunity or remain susceptible. Thus, we cannot simply use the reported effectiveness but must also include the risk of asymptomatic and sub-clinical reinfection among the vaccinated in our probability to become immune upon vaccination. This is important as there is ample evidence by now that vaccinated individuals can still get infected with SARS-CoV-2 and transmit the disease (Harris et al., 2021; Levine-Tiefenbrun et al., 2021; Petter et al., 2021).

The reported effectiveness for BNT162b2 is estimated to be 90% 21 days after the first shot (Hunter and Brainard, 2021). The effectiveness does not increase much through the booster shot as Thompson et al. (2021) report 90% (95% CI = 68%–97%) effectiveness against PCR-confirmed infections after two doses for mRNA vaccines in general. We therefore do not distinguish between the first and the booster shot.

On the other hand, Lipsitch and Kahn (2021) report a lower bound on transmission for the very similar Moderna vaccine of 61%. To strike a middle ground we assume that 75% of individuals achieve sterile immunity after vaccination. This is split into 35% reaching immunity after 14 days after the first shot and 40% reaching immunity after 21 days. This is also supported by more recent evidence by Bernal et al. (2021). They report vaccine effectiveness of approx. 50% against Alpha after the first and 88% after the second dose. Given large numbers of both individuals with one or two vaccines in our study period, our assumption of 75% seems a realistic approximation even though this latest evidence does show important differences in the effectiveness of first and second dose for the Alpha and Delta variant.[4]

3.2.2 The Synthetic Population

We build a synthetic population based on the German microcensus (Forschungsdatenzentren Der Statistischen Ämter Des Bundes Und Der Länder, 2018). We only use private households, i.e. exclude living arrangements such as nursing homes as non-private households vary widely in size and it is very difficult to know which contacts take place in such living arrangements.

We sample households to build our synthetic population of over one million households keeping for each of the 2.3 million individuals their age, gender, occupation

4. Note that Delta is not part of our simulations as it only started to appear in Germany at the end of our simulation period.

and whether they work on Saturdays and Sundays. For each household we draw its county and set the corresponding federal state.

We randomly assign 35% of children below three to attend a nursery (Destatis, 2020). For children between three and six years old, we assume all go to preschool.[5] Children that attend a nursery meet in groups of four (Bertelsmann Stiftung, 2019) plus one adult care taker every weekday when there are no school vacations. Preschool children meet in groups of nine (Bertelsmann Stiftung, 2019) with two adult care takers. These groups are mixed with respect to age but all belong to the same state and most to the same county.

Every child that goes to school is part of a school class. Each school class meets three times per weekday, each time with a different set of two teachers, unless there are vacations or policies that suspend schools.[6] Each class consists of approximately 23 students (OECD, 2013). All students in a class are of the same age and live in the same state and mostly also in the same county. In addition, each child gets assigned a value that captures his or her need to attend nursery, preschool or school. This allows us to capture various degrees of emergency care that can be granted while educational facilities are closed or are on some kind of rotating schedule.

Workers are assigned to a work group that meets daily. The group sizes vary to match the number of daily repeating work contacts reported by working individuals in Mossong et al. (2008). These groups only consist of workers that work in the same county. For a distribution of the number of daily recurring work contacts see Figure 3.2.3b. To match the number of reported weekly work contacts shown in Figure 3.2.3d we pair each worker with up to 14 other workers. Each pair is assigned a weekday on which they meet potentially. 80% of these contacts are individuals from the same county. In the same way children have an educational priority determining if they are entitled to emergency care, workers are assigned a work contact priority that captures how necessary their work is and to which degree they can work from home. This means that it is always the same individuals that continue to have work contacts when work from home mandates of a certain strictness are in place.

In addition to creating groups for educational facilities and work we also have other recurring contacts to represent things like groups of friends or sports teams which meet regularly. Both daily and weekly groups are created analogously to the work groups but matching the numbers in Figure 3.2.3a and Figure 3.2.3c. In addition, since leisure contacts are highly assortative by age all individuals that have a daily leisure contact are matched with a person not only from the same county but also from the same age group.

5. According to Destatis (2020) the share is 92.5%.
6. We implement vacations on the federal state level.

The individuals in our population can react to events such as developing symptoms that are typical of CoViD-19, a positive PCR test or a positive rapid test by reducing their contacts. To determine who would reduce their contacts in such a situation or demand a rapid test we introduce a quarantine compliance parameter. Similarly, we introduce a rapid test compliance parameter that determines in which order individuals start demanding rapid tests when rapid tests become increasingly available. This makes sure that when for example only 10% of workers get tested, it is the same workers that have access to tests every week.

Lastly, for the distribution of vaccinations every individual is assigned a vaccination group and a vaccination rank from that group that creates a complete vaccination queue over the population including a share that refuses to be vaccinated (ξ) which we calibrate to 15% (Robert Koch-Institut, 2021b). The vaccination groups are created to match the recommendations by the Ständige Impfkommission (Vygen-Bonnet et al., 2020).[7] An individual's vaccination priority depends on her work contact priority, her age group and a random component to capture preconditions like diabetes. To cover that the Pfizer-BioNTech vaccine was later approved for younger age groups we put adolescents and children into two groups that follow after the adult population. These groups do not become eligible within our simulation frame until June.

The resulting vaccination shares by age group over time are shown in Figure 3.2.2. One can see that the first vaccinations go to some workers with very high work contact priority and to the 80 to 100 year olds followed by the 60 to 79 year olds. Both groups are saturated with vaccinations by mid March and start of May, respectively. By June a third of the younger adults have received the vaccination but these groups still remain far from herd immunity thresholds.

7. We cover that teachers were prioritized more than recommended by the commission.

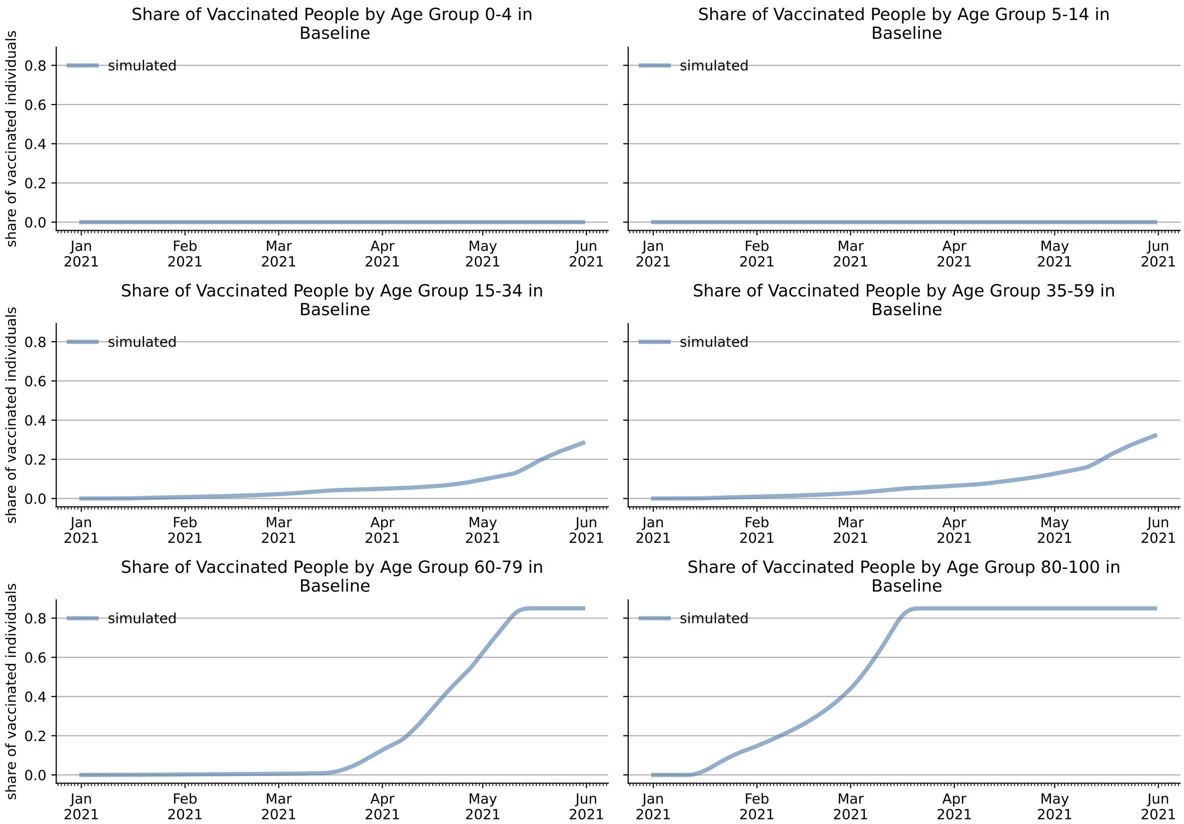
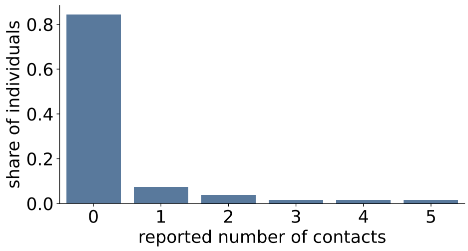


Figure 3.2.2. Vaccination Rates by Age Group

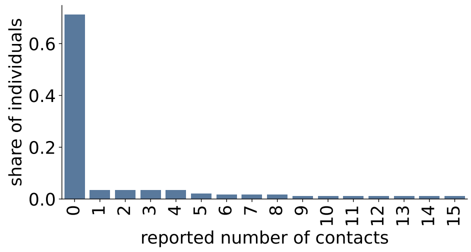
Note: Each figure shows the share of individuals who are vaccinated in the respective age group from January 2021, when vaccinations started in Germany to June 2021 where our simulations end. The flattening for the oldest age groups shows that these groups are saturated and only vaccination refusers remain.

3.2.3 Number of Contacts

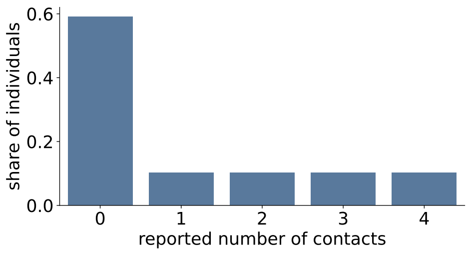
We calibrate the parameters for the predicted numbers of contacts from contact diaries of over 2000 individuals from Germany, Belgium, the Netherlands and Luxembourg (Mossong et al., 2008). Each contact diary contains all contacts an individual had throughout one day, including information on the other person (such as age and gender) and information on the contact. Importantly, for each contact individuals entered of which type the contact (school, leisure, work etc.) was and how frequent the contact with the other person is. Binning the number of contacts for very high numbers, we arrive at the distributions of the numbers of contacts by type of contact (η_c) as shown in Figure 3.2.3.



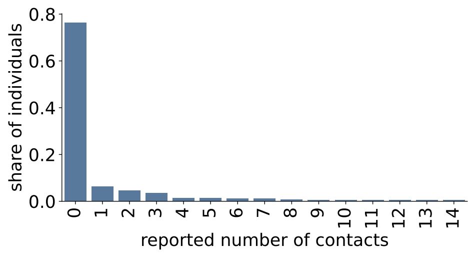
(a) Number of Daily Recurrent Other Contacts



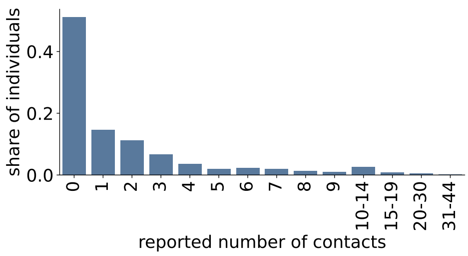
(b) Number of Daily Recurrent Work Contacts



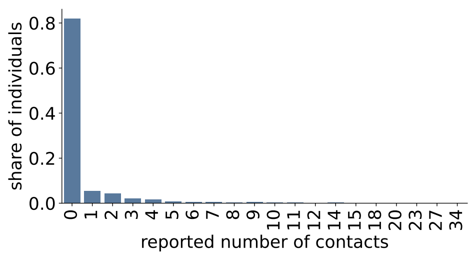
(c) Number of Weekly Recurrent Other Contacts



(d) Number of Weekly Recurrent Work Contacts



(e) Number of Non Recurrent Other Contacts



(f) Number of Non Recurrent Work Contacts

Figure 3.2.3. Number of Contacts of the Different Contact Types

Note: This figure shows the pre-pandemic number of contacts individuals report of different contact types (η_c). In the model it is sampled every day which of the numbers of non recurrent contacts a person is planned to have. Note that the contact diaries include such high values that super spreading events are well possible in our model through non recurrent models. For recurrent contacts individuals are put into groups that meet either every day or on a particular week day every day. The pre-pandemic number of contacts with transmission potential is reduced by policies (ρ), seasonality (κ) and individual responses (τ) to events such as receiving a positive rapid test to the number of actual contacts with transmission potential. The left column shows the distribution of the number of other contacts individuals report (η_{other}). Other contacts include all contacts that are not household members, school contacts or work contacts, for example leisure contacts. We assume that individuals in households with children or teachers or retired individuals have additional non recurrent other contacts during school vacations to cover things like family visits or travel during vacations. The right column shows the distribution of the different types of work contacts (η_{work}). Work contacts only take place between working individuals.

An exception where we do not rely on the data by Mossong et al. (2008) are household contacts. Since households are included in the the German microcensus (Forschungsdatenzentren Der Statistischen Ämter Des Bundes Und Der Länder, 2018) on which we build our synthetic population we simply assume for the household contacts that individuals meet all other household members every day. The distribution of household contacts is shown in Figure 3.2.4.

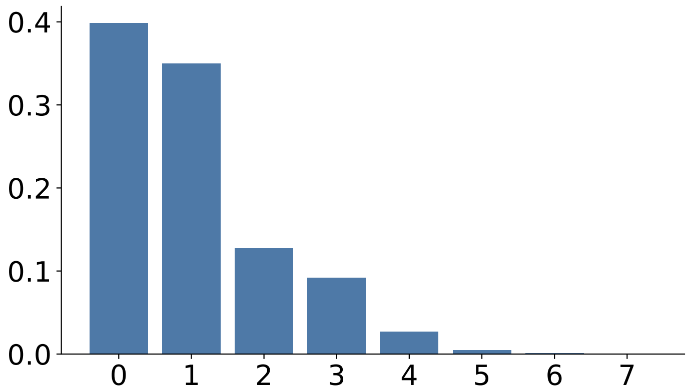


Figure 3.2.4. Number of Household Contacts

Note: Every individual meets all other household members every day. The German microcensus sampled full households such that our synthetic population automatically fits population characteristics such as size and age distribution.

3.2.4 Assortativity

As explained in section 2.6, the probability that two individuals are matched can depend on background characteristics. In particular, we allow this probability to depend on age and county of residence (α). While we do not have good data on geographical assortativity and set it such that 80% of contacts are within the same county, we can calibrate the assortativity by age from Mossong et al. (2008).

Figure 3.2.5 shows that assortativity of the other non-recurrent contacts by age is especially strong for children and adolescents. For older people, the pattern becomes more dispersed around their own age group, but within-age-group contacts are still the most common contacts. Figure 3.2.6 shows that assortativity by age is also important among non-recurrent work contacts. These probabilities over age groups are used in the matching algorithm when it is drawn which individuals meet.

For recurrent contacts, we construct groups to have the following features: Recurrent work contacts are not assortative by age. Daily work groups are always of the same county and weekly work contacts are to 80% with workers from the same county.

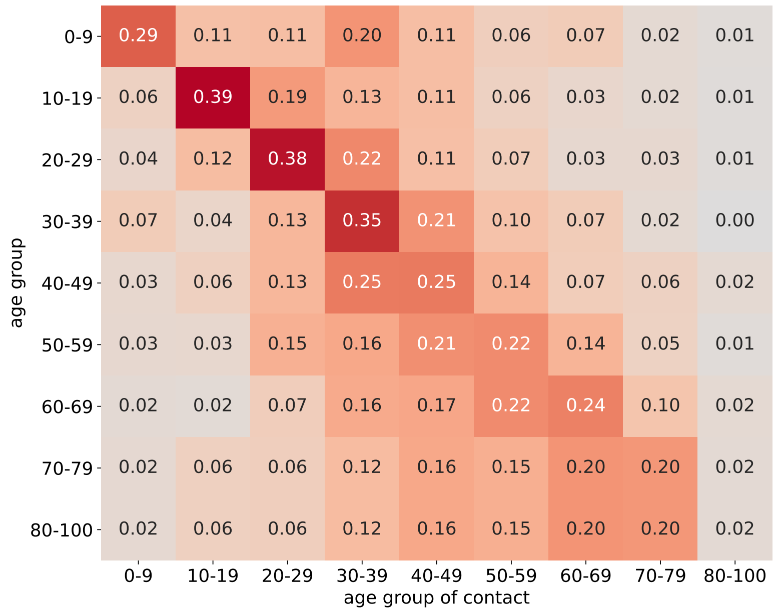


Figure 3.2.5. Distribution of Non-Recurrent Other Contacts by Age Group

Note: The figure shows the distribution of non recurrent contacts by age group for other contacts. A row shows the share of contacts a certain age group has with all other age groups. Higher values are colored in darker red tones. The diagonal represents the share of contacts with individuals from the same age group. The 80-100 age group for other contacts was so small that we assumed for them to have the same contact distribution as the 70-79 year olds.

Other recurrent contacts are constructed the same way but we impose for daily contacts that they are always with individuals from the same age group. School classes are groups where the same children of the mostly same age and county meet with teachers every day. Nurseries and preschools mix children by age but match them to come mostly from the same county. Household age composition follows directly from the German microcensus data we use to construct our synthetic population (see Section 3.2.2 for details).

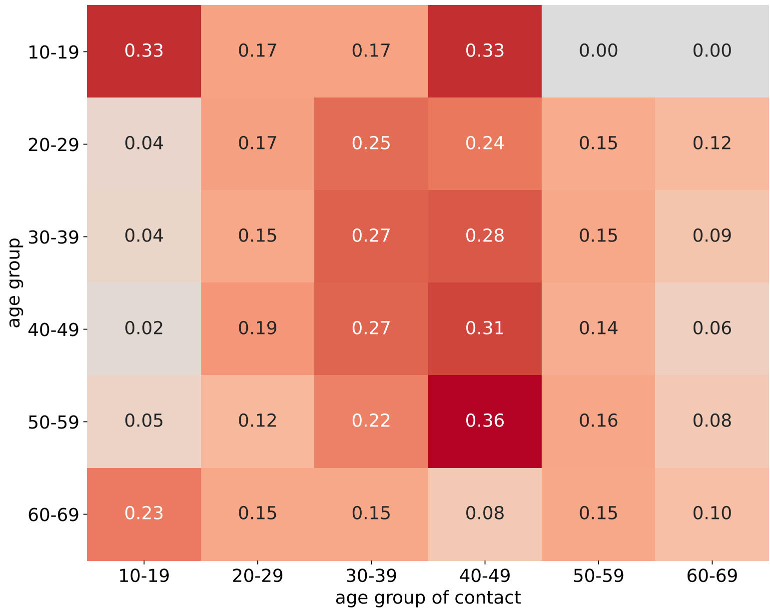


Figure 3.2.6. Distribution of Non Recurrent Work Contacts by Age Group

Note: The figure shows the distribution of work contacts. A row shows the share of contacts a certain age group has with all other age groups. Higher values are colored in darker red tones. The diagonal represents the share of contacts with individuals from the same age group. We only show age groups that have a significant fraction of working individuals.

3.2.5 NPIs

For each of the three contact types education, work and other, we have NPIs in place that reduce the number of contacts with transmission potential that people have for most or the entire simulation period, each with several changes over time. Germany had no policies limiting contacts within households so there are no policies on them.[8]

For nurseries, preschools and schools we implement vacations as announced by the German federal states as well as school closures, emergency care and rotating schedules where only one half of students attends every other week or day. An approximation of the share of contacts still taking place with the different school regulations can be found in Figure 3.2.7. Schooling policies differed between states and usually involved rules based on local incidences. We simplify these rules to one federal policy based on the federal incidence and the policies of the three most populous federal

8. Household contacts can, however, be reduced when individuals quarantine themselves after developing symptoms, for example. This happens to a lesser degree than other contacts to capture difficulties in isolation within the home.

states (North Rhine-Westphalia, Bavaria and Baden-Württemberg). For a detailed description of the education policies we refer the reader to Appendix 3.B.

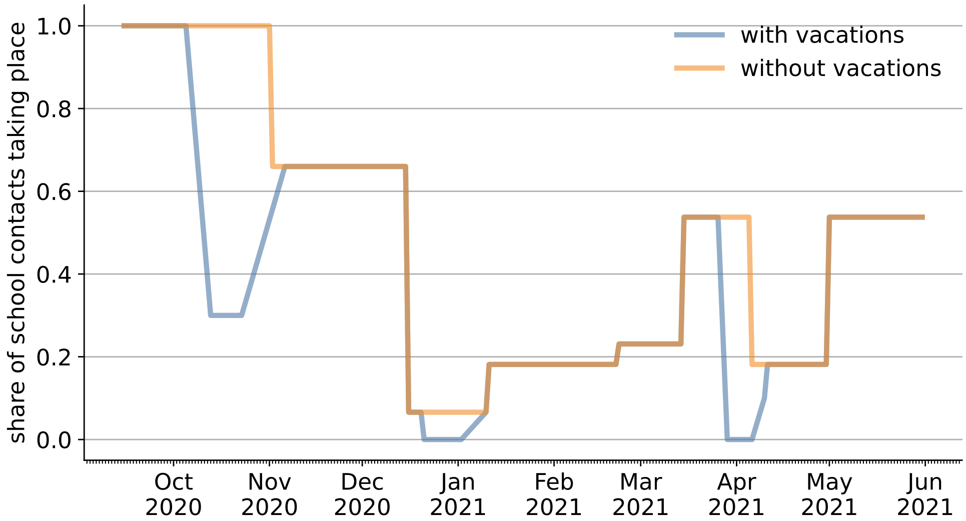


Figure 3.2.7. The Contact Reduction Effects of School Attendance Policies

Note: The figure shows the approximate share of school contacts taking place with and without vacations factored in. In contrast to other policies, school policies are not implemented via multipliers but as mechanistic models (e.g. split classes with emergency care). For the above plot we assigned approximate multipliers to those policies. Thus, the figure is only an illustration that shows the approximate share of contacts taking place compared to the pre-pandemic level with and without vacations.

Our main work policy is home office usage. We use the reductions in work mobility reported by the Google Mobility Data (Google, LLC, 2021) as a proxy for the share of individuals in home office, $\rho_{w, attend, t}$. The resulting reductions in work contacts are not random but governed through a work contact priority which goes from zero to one and is fixed over time for each individual. For example, if the Google Mobility Data indicate that work mobility was reduced by 30% in February 2021, workers whose work contact priority is below 0.30 do not have work contacts. Thus, the Google Mobility data gives us for each day and federal state the threshold for the work contact priority below which workers stay home. Figure 3.2.8 shows the share of workers that go to work over time at the German level. Using the data on the state level allows us to account for local holidays and differences in state regulations.

In addition, for both work and school contacts we assume that hygiene measures (such as masks, ventilation and hand washing) became more strict and more conscientiously observed in November 2020, leading to an estimated reduction of 33% in the number of contacts with the potential to transmit CoViD-19 ($\rho_{hygiene}$). Lastly, for the other contacts category ($\rho_{other, t}$) we could also not calibrate the policies from

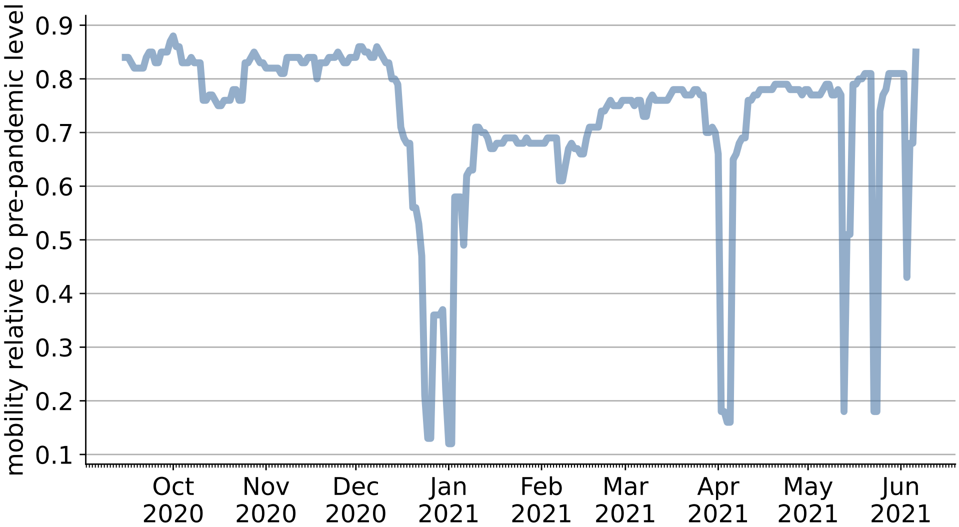


Figure 3.2.8. The Contact Reduction Effects of Work Attendance Policies

Note: The figure shows the work mobility as reported by the Google Mobility Data (Google, LLC, 2021). We take this as a proxy of the share of workers who still have physical work contacts ($\rho_{w,\,attend,\,t}$). The figure interpolates over weekends as we handle weekend effects through information on work on weekends in the German census data. The figure shows the share for Germany as a whole. To capture the effect that local policies, local holidays, etc. have on work contacts we use the data on the state level to determine which workers go to work depending on the state they live in.

data but had to estimate the policy effects. The estimation and values are detailed in Section 3.3.1 and Figure 3.3.2.

3.2.6 Rapid Test Demand

In our model, there are five potential reasons why rapid tests are done:

1. the individual plans to have work contacts
2. the individual plans to attend school or to work in an education facility
3. the individual has a household member who has tested positive or developed symptoms
4. the individual has developed symptoms but has not received a PCR test
5. the individual plans to participate in weekly meeting of the other contact type (such as a choir rehearsal)

For work contacts, we know from the COSMO study (Betsch, Korn, Felgendreff, Eitze, Schmid, Sprengholz, Wieler, Schmich, Stollorz, Ramharter, Bosnjak, Omer, Thaiss, De Bock, and Von Rüden, 2021) that 60% of workers who receive a test offer by their employer regularly use it ($\pi_{w,d}$). We assume this share to be time constant. In addition, there are some surveys that allow us to trace the expansion of employers who offer tests to their employees ($\pi_{w,s,t}$), see Appendix 3.C.1 for details. We interpolate between these points linearly, arriving at the offer shares shown in Figure 3.2.9. In addition, we increase the frequency of testing ($\theta_{t,work}$) from weekly to twice weekly during April.

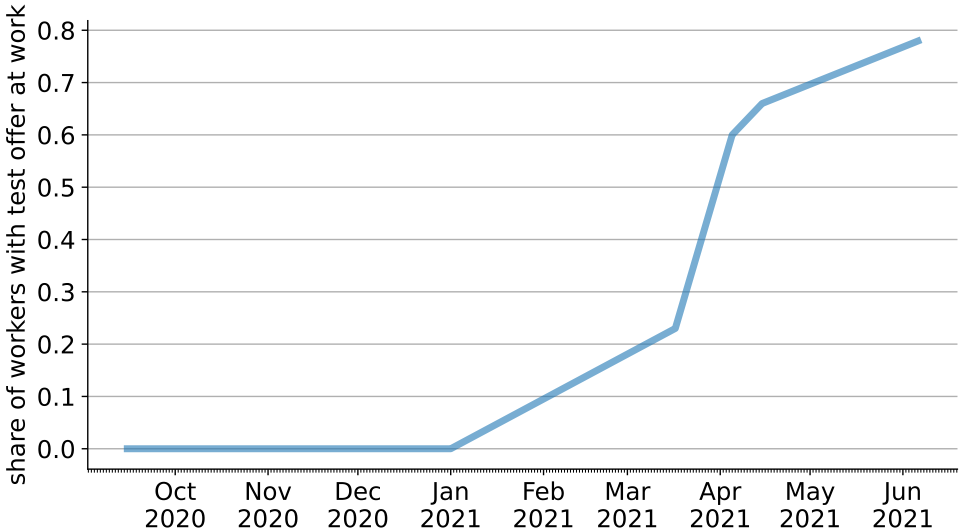


Figure 3.2.9. Share of Individuals Being Offered a Rapid Test at Work

Note: Employees were never subject to mandatory tests. According to surveys only 60% of workers regularly used rapid test offers by their employers (Betsch, Korn, Felgendreff, Eitze, Schmid, Sprengholz, Wieler, Schmich, Stollorz, Ramharter, Bosnjak, Omer, Thaiss, De Bock, and Von Rüden, 2021).

For educators and school students we rely on decrees and data provided by the education ministries of the federal states. Again, we focus on the three most populous states to create one educational rapid testing policy that is applied nationwide. The shares are shown in Figure 3.2.10. The purple line shows the share for educators and the red line school students. Educators have access to tests earlier than school students. By mid April, both school students and educators are tested nearly universally twice per week. There are no tests in preschools and nurseries. For the detailed construction of our shares see Appendix 3.C.2.

There is hardly any data on private rapid test demand, which covers the last three reasons above. To limit our degrees of freedom, we only have one parameter that governs how many individuals do a rapid test because of any of these private demand reasons ($\pi_{private,t}$). We assume that there is no private rapid test demand until March when both the free citizens' tests and rapid tests for lay people started to become available (Bundesanzeiger, 2021a; Presse- und Informationsamt der Bundesregierung, 2021) and other access to rapid tests was very limited.

According to the COSMO study (Betsch, Korn, Felgendreff, Eitze, Schmid, Sprengholz, Wieler, Schmich, Stollorz, Ramharter, Bosnjak, Omer, Thaiss, De Bock, and Rüden, 2021) 63% would have been willing to take a test in the round of 23rd of February 2021 when an acquaintance would have tested positive. Since this is only asking for willingness not actual behavior, we take this as the upper bound of

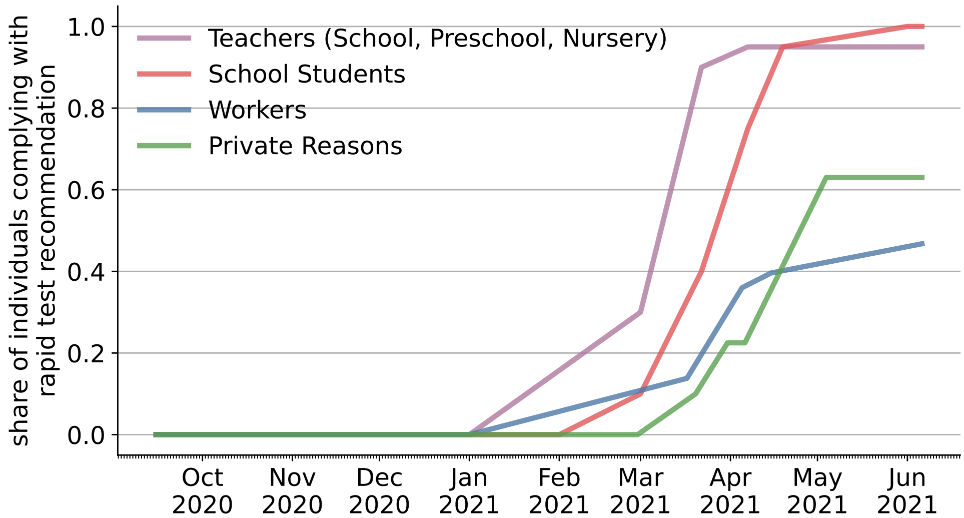


Figure 3.2.10. Share of Individuals Doing a Rapid Test

Note: Rapid test demand can be triggered by individuals planning to have education contacts ($\pi_{educator,\,t}$ or $\pi_{students,\,t}$), work contacts ($\pi_{w,\,d}$ and $\pi_{w,\,s,\,t}$), developing symptoms without access to a PCR test, having a household member with a positive test or symptoms or planning to attend a weekly other meeting ($\pi_{private,\,t}$). In each case whether a rapid test is done depends on how long it has been since the individual's last rapid test and her individual compliance parameter. As an example, take an educator in early February. At that time 20% of educators have access to rapid tests at their educational facility. Thus, all educators whose (time-constant) compliance parameter belongs to the upper 20% of the population test themselves if it has been more than seven days since their last rapid test.

private rapid test demand which we estimate in our model to be reached in the beginning of May. To cover that many people are likely to have sought and done their first rapid test before the Easter holidays, we add another point that we estimate for the rapid test demand around Easter. Similarly, we estimate one point in mid March when tests started to become available in grocery stores and pharmacies. The resulting share of private rapid test demand is shown as the green line in Figure 3.2.10. See Section 3.3.1 for details on the estimation.

3.2.7 PCR Testing

This section describes the parameters needed to specify our model of PCR tests, namely the share of detected cases, the share of tests that go to symptomatic individuals, the time it takes to process a PCR test and the share of individuals with a positive rapid test that request a PCR test. Refer to Section 2.9 for a description of the full testing model.

At the core of our case detection model is the estimate for the share of cases that is detected in the absence of rapid tests (ψ_t). For this, we rely on the Dunkelzifferradar Project (Paul et al., 2020) which uses estimates of the case fatality rate to estimate the number of total cases given the number of CoViD-19 deaths which are assumed to be perfectly observable. For 2020, we follow their reported share of detected cases quite closely. One exception is the phase of November 2020 where we interpolate to maintain monotonicity during the fall as there was no reason why the share of detected cases should have risen in that time.[9]

Since vaccinations started after Christmas 2020 and these were predominantly given to the elderly in the beginning and other vulnerable groups in spring, we expect the relationship between deaths and the number of total infections to change rapidly in 2021. This is why we stop using the share of detected cases estimated by the Dunkelzifferradar after Christmas. Instead, we assume that the share of detected cases would have stayed the same in the absence of rapid tests. Thus, we achieve an increase in the share of detected cases that is driven from inside our model through increased rapid testing which leads follow-up PCR tests when individuals test positive. In Appendix 3.G we show the resulting share of detected cases in our simulations by age group over time.

Furthermore, we model reductions in the share of detected cases due to the two major holidays in our simulation period, Christmas and Easter. During both holidays many laboratories did not process tests and most physicians' offices were closed, leading to less PCR tests and short and large drops in the share of detected cases. The resulting share of detected cases in the absence of rapid tests is shown in Figure 3.2.11.

9. The testing policy changed in November (Robert Koch Institute, 2020). However, this only moved the rare PCR tests more towards vulnerable groups.

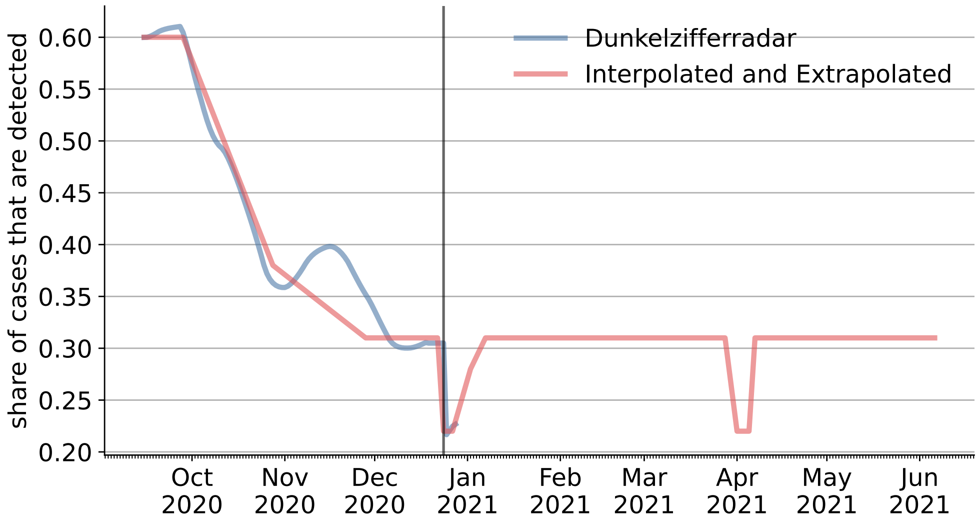


Figure 3.2.11. Share of Detected Cases in the Absence of Rapid Tests

Note: The figure shows the share of cases that is reported as an official case via PCR confirmation. We use the overall share of detected cases that was estimated through the case fatality ratio by the Dunkelzifferradar Project (Paul et al., 2020) for all of 2020 and then assume it to be constant as vaccinations of the elderly strongly affect the case fatality rate which the project does not account for. Starting in 2021 in addition to the overall numbers of detected cases through symptoms and a random component, cases are additionally detected through confirmation of positive rapid tests which happens endogenously inside the model. For the public holidays of Christmas and Easter we lower the share of detected cases as fewer PCR tests are available during public holidays. See Figure 3.G.1 for how the share of detected cases develops in our model for each age group.

From the share of detected cases (in the absence of rapid tests) and the number of infections we arrive at the number of positive PCR tests in our model (excluding those triggered by rapid tests). A share of these positive tests goes to symptomatic individuals ($\chi_{symptom,t}$). This share is calibrated from German data on case characteristics (Robert Koch Instititue, 2021) and shown in Figure 3.2.12. We keep $\chi_{symptom,t}$ constant after Christmas because the RKI data does not include if a PCR test was done to confirm a positive rapid test and this share is used for PCR test demand without prior rapid test indication.

PCR tests take one to four days until their result is revealed to the individual ($\gamma_{PCR,d}$). Relying on the ARS data (Robert Koch-Institut, 2020) we calculate that 33% of individuals receive the test result after one day, 50% after two days, 10% after three days and 7% after four days.

To model the demand for PCR tests through rapid tests, we only need the share of individuals that seek a PCR test to confirm a positive rapid test result ($\chi_{confirmation}$). We calibrate this from the COSMO study (Betsch, Korn, Felgendreff, Eitze, Schmid,

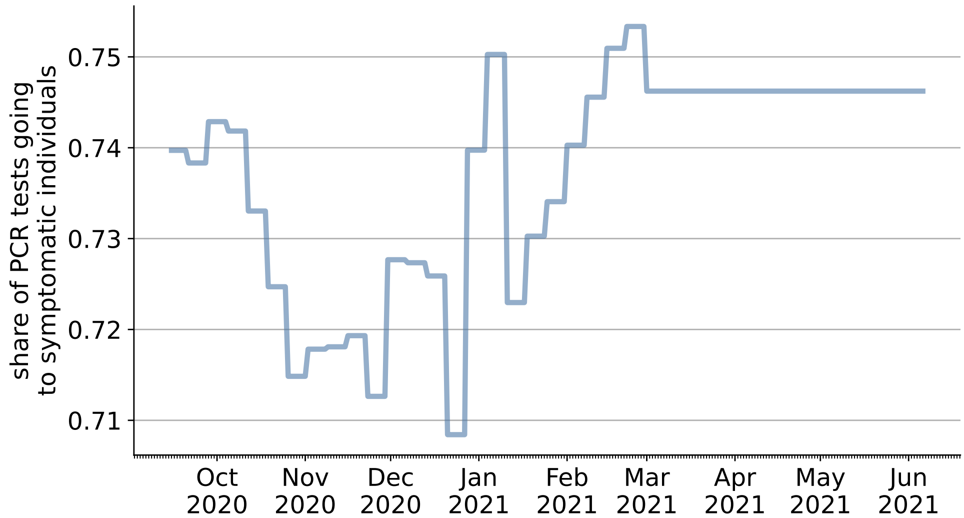


Figure 3.2.12. Share of Positive PCR Tests Administered to Symptomatic Individuals

Note: The share of positive PCR tests that are administered to symptomatic individuals ($\chi_{symptom,\,t}$). Since it was not recorded for every case if the person was symptomatic or not we take the midpoint between the upper and lower bound. We keep the share constant after Christmas because the RKI data does not include if a PCR test was done to confirm a positive rapid test and this share is used for PCR test demand without prior rapid test indication.

Sprengholz, Wieler, Schmich, Stollorz, Ramharter, Bosnjak, Omer, Thaiss, De Bock, and Von Rüden, 2021) who asked this as a hypothetical question in March of 2021. There 82% of Germans reported that they would follow up on a positive rapid test with a PCR test.

3.2.8 Behavioral Response

Lastly, we need to set the parameters that decide how individuals reduce their contacts after certain events, τ. We distinguish between the reduction in household contacts (which are harder to avoid) and non household contacts. There are three events which trigger potential contact reductions: showing symptoms of CoViD-19, having received a positive rapid test and having received a positive PCR test. The only survey data we are aware of on this is again the COSMO survey (Betsch, Korn, Felgendreff, Eitze, Schmid, Sprengholz, Wieler, Schmich, Stollorz, Ramharter, Bosnjak, Omer, Thaiss, De Bock, and Von Rüden, 2021) where 85% of individuals claimed they would isolate and restrict their contacts after a positive rapid test. We assume this reduction for non household contacts. As household contacts are much more difficult to avoid, we assume that they are only reduced by 30%. We assume the same behavior for individuals that develop symptoms. Lastly, we assume the response to a positive PCR test to be stronger than in the other two cases and set the reduction of non household contacts to 95% and the reduction of household contacts to 50%. These endogenous contact reductions are also implemented using a compliance parameter.

3.3 Estimation and Fit

3.3.1 Estimated Parameters

We estimate parameters that cannot be calibrated outside of the model with the method of simulated moments (McFadden, 1989) by minimizing the distance between simulated and observed infection rates (aggregated and disaggregated by federal state and age groups), fatality rates and virus variant shares. Using a sample of over 2.3 million agents from the population structure in September 2020 we simulate recorded infection rates until the end of May 2021 with different parameter values. Since our model includes a lot of randomness, we average simulated infection rates over several model runs for evaluating the model fit. All estimated parameters are described in Table 3.3.1.

Table 3.3.1. Estimated Parameters

notation	estimate	note
Infection Probabilities		
$\beta_{household}$	0.1	base probability of getting infected by an infectious household member
β_{school}	0.012	base probability of getting infected by an infectious classmate or teacher
$\beta_{young\ educ}$	0.005	base probability of getting infected by an infectious classmate or teacher
β_{work}	0.1475	base infection probability for work contacts
β_{other}	0.15875	base infection probability for other contacts
Policy Parameters		
$\rho_{hygiene}$	0.66	reduces infectiousness of work and education contacts from November to end of simulation
$\rho_{other,\ before\ Oct\ 1}$	0.75	before October
$\rho_{other,\ Oct\ 1\ to\ Oct\ 20}$	1.00	high activity due to reopenings and fall vacations
$\rho_{other,\ Oct\ 21\ to\ Nov\ 1}$	0.75	anticipation of a lockdown and precaution due to high incidenes
$\rho_{other,\ Nov\ 2\ to\ Dec\ 1}$	0.52	"lockdown light"
$\rho_{other,\ Dec\ 2\ to\ Dec\ 23}$	0.57	"lockdown light" with lockdown fatigue and holiday shopping
$\rho_{other,\ Dec\ 24\ to\ Dec\ 26}$	0.65	Christmas holidays
$\rho_{other,\ Dec\ 27\ to\ Feb\ 10}$	0.35	hard lockdown after Christmas
$\rho_{other,\ Feb\ 11\ to\ Feb\ 28}$	0.50	lower precaution due to low incidences and lockdown fatigue
$\rho_{other,\ after\ Feb\ 28}$	0.515	many contact reducing policies are lifted
Introduction of the Alpha Strain		
$\omega_{Alpha,\ Jan\ 31}$	0.986	number of Alpha cases per 100 000 individuals to import on January 31st. Imported Alpha cases rise from January 1st where 0 cases are imported. No cases are imported in other months.
Rapid Test Introduction		
$\pi_{private,t}$	Figure 3.2.10	the private rapid tests levels in mid March and at Easter as well as the date at which full availability of private rapid tests is reached are fit to the data. Values between those levels are interpolated linearly. The remaining rapid tests demands are calibrated from surveys. See Section 3.2.6

We fit our model to data for Germany from mid September 2020 until June 2021. We do not use earlier periods for three reasons. Firstly, in the beginning PCR tests were very scarce and the reported case numbers unreliable. Secondly, during the summer the case numbers were very low. This could lead to the epidemic going extinct in our simulation. Thirdly, over the summer, imported cases from touristic travel were likely important for the infection dynamic but there is not enough data to include them into our model.

To avoid over-fitting and simplify the numerical optimization problem, we only allow for five different infection probabilities: 1) for contacts in schools 2) for contacts in preschools and nurseries. 3) for work contacts. 4) for households. 5) for other contacts.

Since the infectiousness of a contact between an infectious and a susceptible person depends on many things, the numerical values of the infection probabilities in Table 3.3.1 only reflect a base probability. This base probability is modified by a seasonality factor, an age specific susceptiblity factor and an infectiousness factor that depends on the virus variant of the infected person. The base infection probability is only equal to the actual infection probability when all of those factors are one. This would be the case for a contact between an 80+ year old susceptible person with a person who is infected with the Alpha variant of the virus on January 1st.

It is not possible to rank different types of contacts according to their infectiousness just from the numerical values of the infection probabilities. There are two reasons for this: Firstly, for computational reasons the seasonality factor is normalized such that it reaches one at its peak. It has thus a lower average for contact types with strong seasonality (e.g. other contacts) than for contact types with weak seasonality (e.g. work contacts). Secondly, for household and school contacts we do not have data on whether people meet physically. Thus, the infection probabilities for those contact types are the product of the probability to have physical contact on a given day and the infection probability of that contact.

In order to get a feeling for the infectiousness of each contact type it is more informative to look at how many infections were caused by each contact type. This is depicted in Figure 3.3.1. We can see that work and other contacts are the main drivers of the pandemic, followed by infections in households. Schools and preschools contribute fewer infections which is to be expected given that there are much fewer students than working adults in the German population. Nevertheless, Figure 3.4.10 shows that schools do have a notable effect on the infection dynamic in the long run.

We also estimate a parameter that reflects the effect of hygiene measures at work and in educational facilities. This parameter becomes active in November 2020 when

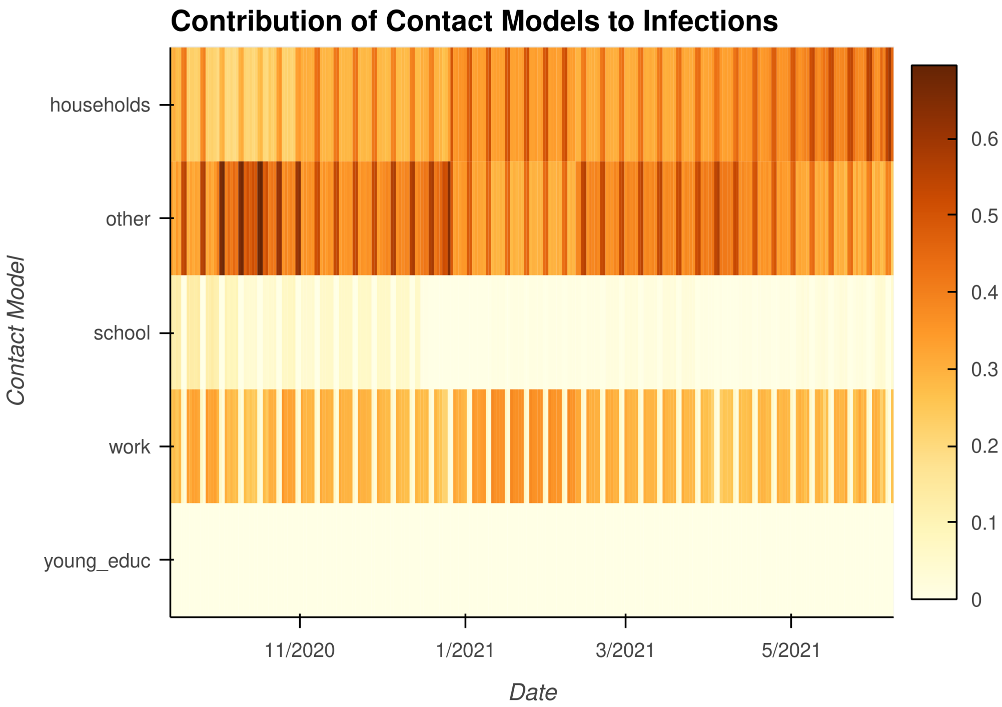


Figure 3.3.1. Daily share of infections by contact type

Note: Daily share of infections that were contributed by each contact type. Darker colors mean that a larger share of infections were contributed by that contact type. The majority of infections take place in the workplace, in households and via other contacts. Schools and preschools contribute less infections, especially after hygiene measures have been introduced.

stricter mask mandates and distancing rules were introduced. It is estimated to reduce infectiousness of contacts by one third.

Moreover, we estimate nine different multipliers that reflect how strongly other contacts are reduced over time. The dates at which we switch between the multipliers usually coincide with policy changes and is not determined from the case numbers. The only exception to this are slight adjustments to parameters to incorporate lockdown fatigue (towards the end of a lockdown period) or precautionary contact reductions (in times of high incidences right before a lockdown is enacted). The estimated other multipliers are depicted in Figure 3.3.2.

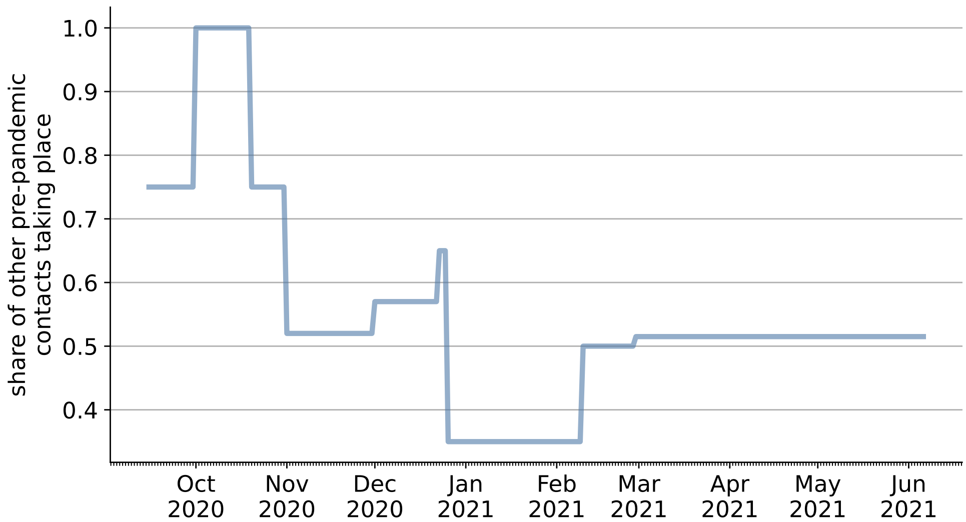


Figure 3.3.2. Share of Pre-Pandemic Other Contacts Taking Place with Infection Potential

Note: Values of the other multiplier. All values are estimated via the method of simulated moments. The rationale behind each switching point is described in Table 3.3.1

While we estimate nine different values for the other contact multiplier, they are not estimated completely freely. In particular, we ensure that the ordering of the parameter values is consistent with the stringency of policies. For example, the strongest contact reduction was estimated for January 2021 during which very strict measures and curfews were in place, whereas the weakest contact reduction was in October 2020 where policies were very lenient.

Since we do not have good data on the reduction of other contacts, it is not possible to separately estimate parameters for contact reduction and the effect of hygiene measures. The reported other multipliers in Figure 3.3.2 are thus a combination of contact reductions and hygiene measures.

Finally, we estimate one parameter that governs the introduction of the Alpha variant in January 2021. This parameter implies that at the end of January roughly one case per 100 000 individuals per day is imported. After January we do not model imported cases of Alpha anymore because they are negligible compared to the endogenous growth of that virus variant.

While a formal identification argument is beyond the scope of this paper, we give a rough intuition which features of the data help us to estimate each parameter below.

The different infection probabilities can be separately identified because the degree to which each contact type is active varies over time (e.g. school closures, vacations and different work from home policies)[10] and they affect different subgroups of the population differently (e.g. β_{school} most strongly affects kids whereas β_{work} has the strongest effect on adults in working age and β_{other} affects all age groups equally). The hygiene and other multipliers can be identified because they are only active in certain time periods. However, it is necessary to normalize one other multiplier to 1 because there is no period without any contact reduction in our data. The introduction parameter for the Alpha mutation can be identified from the share of that virus variant in the population. The rapid test demand parameters are identified because rapid tests first lead to a very steep increase in observed cases and then to a sudden decrease – in a time where almost all other things in the model would not cause a change in trend.

3.3.2 Model Fit

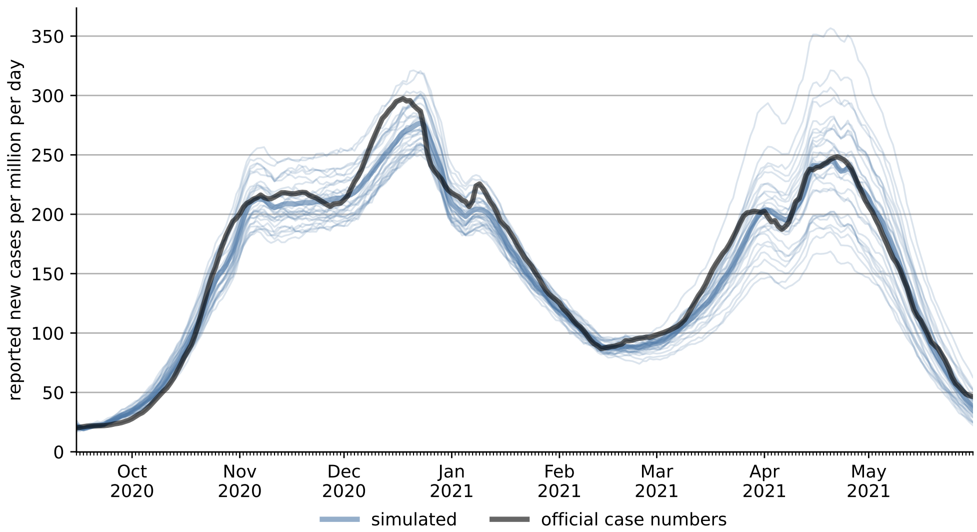
This section compares simulated data from our model with empirical data from Germany. We look at observed infections (overall as well as by age group and federal state), the effective replication number, the spread of Alpha and vaccinations. Overall, our model achieves an excellent fit of the two waves of infections with few free parameters (Figure 3.3.3a). As a result the effective replication number R_t also closely follows that reported by the RKI (see Figure 3.3.3b). We also achieve an excellent fit for most age groups in Germany. The fit is also good for many German federal states. Despite the fact that the number of performed rapid tests and their distribution in the population are determined endogenously in our model, we fit the share of the population with at least a weekly rapid test very well. For the share of individuals who have ever done a rapid test we err on the side of too few tests.

Our fit of the infection rates in Germany between October 2020 and June 2021 is excellent. The incidence in our model matches both the levels and the shape of the

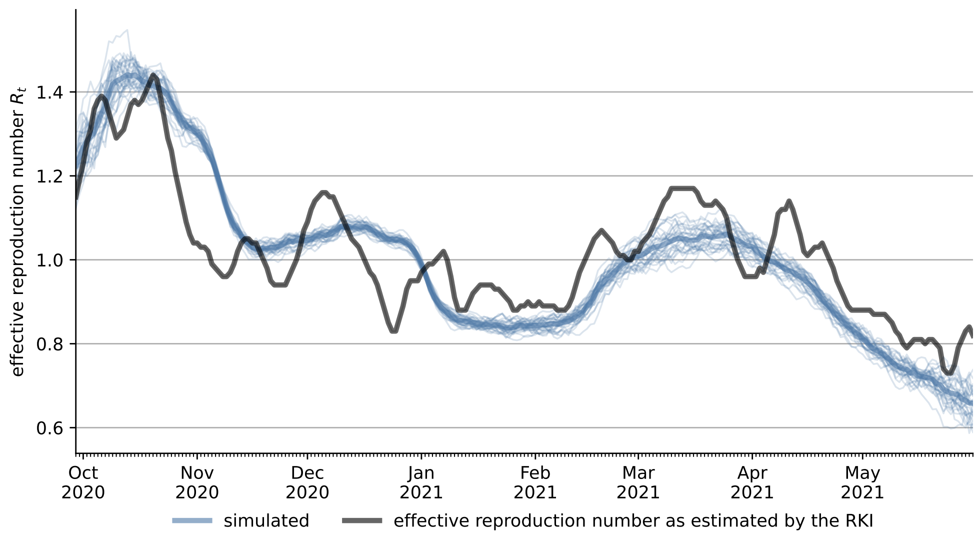
10. see Appendix 3.A for a history of Germany's response to the CoViD-19 pandemic and the resulting variation in policies.

reported incidence almost perfectly. When the prevalence of the virus is high and especially after explosive growth phases, the effect of random events on the incidence is large. Therefore, all reported simulations average over at least 30 simulation runs which is enough to reduce the sampling uncertainty to a negligible level.

Our fit of the effective replication number R_t closely follows the values reported by the RKI (see Figure 3.3.3b) even though we calculate R_t on all infected individuals not just the detected cases. This explains why the R_t in our simulations is higher during phases where the share of detected cases (ψ_t) falls. This is the case in the fall of 2020 (see Figure 3.2.11) where the RKI underestimated the effective replication number due to observing a falling share of cases. Analogously, the R_t in our simulations is lower than the R_t reported by the RKI in spring where the share of known cases increased due to increased rapid testing.



(a) Observed Incidence in the Model and as Reported by the RKI



(b) Effective Replication Number R_t in the Model and as Reported by the RKI

Figure 3.3.3. Model Fit of the Reported Cases and the Effective Replication Number

Note: Both figures show averages and single runs. The average is the thick line. Single runs are shown as lighter and thinner lines. We averaged and show 30 simulation runs. The upper figure shows the daily incidence rate per million for the simulated reported infection rates. The official case numbers as reported by the RKI are plotted in black. The fit is overall very good. The lower figure shows the effective replication number (R_t) as reported by the RKI and as calculated in our model. The R_t gives the average number of new infections caused by one infected individual. The R_t in our model broadly follows the R_t reported by the RKI. Two differences stand out. Firstly, the RKI's R_t drops faster in November. This is likely due to a decline in the estimated overall share of detected cases (ψ_t) when the second wave hit Germany. The second difference is from mid February to mid March where the RKI's reported R_t increased more rapidly than that in our model. Here the opposite effect can be expected. During this time rapid tests increased strongly leading to more cases being detected. In the short term this leads an R_t estimation that is based on detected cases to overestimate the replication number. For legibility reasons, all lines are rolling 7-day averages.

Zooming into the different age groups in Figure 3.3.4, we can see that our model is also able to reproduce the infection rates on this level. The only major deviation from this pattern is that our model predicts too few infections for the 80 to 100 year olds. This was to be expected because our synthetic population does not include inhabitants of nursing homes. Outbreaks in nursing homes led to a large number of infections among the oldest during the second wave of the pandemic in Germany. Moreover, the model predicts too few observed infections for the 15 to 34 years old at the end of 2020 and the 5 to 14 year olds in April and May 2021. The former is likely due to the fact that this age group has a very active social life which is not fully captured by our contact networks. The latter probably comes from a too conservative model of school reopenings.

it is important to note that this fit is only possible due to our age specific and time variant shares of detected cases. As can be seen in Appendix 3.G, these shares vary strongly over time and between age groups and especially once rapid tests enter the picture the differences between detection rates change quite substantively, with strong effects on how infections in our model are translated into observed infections.

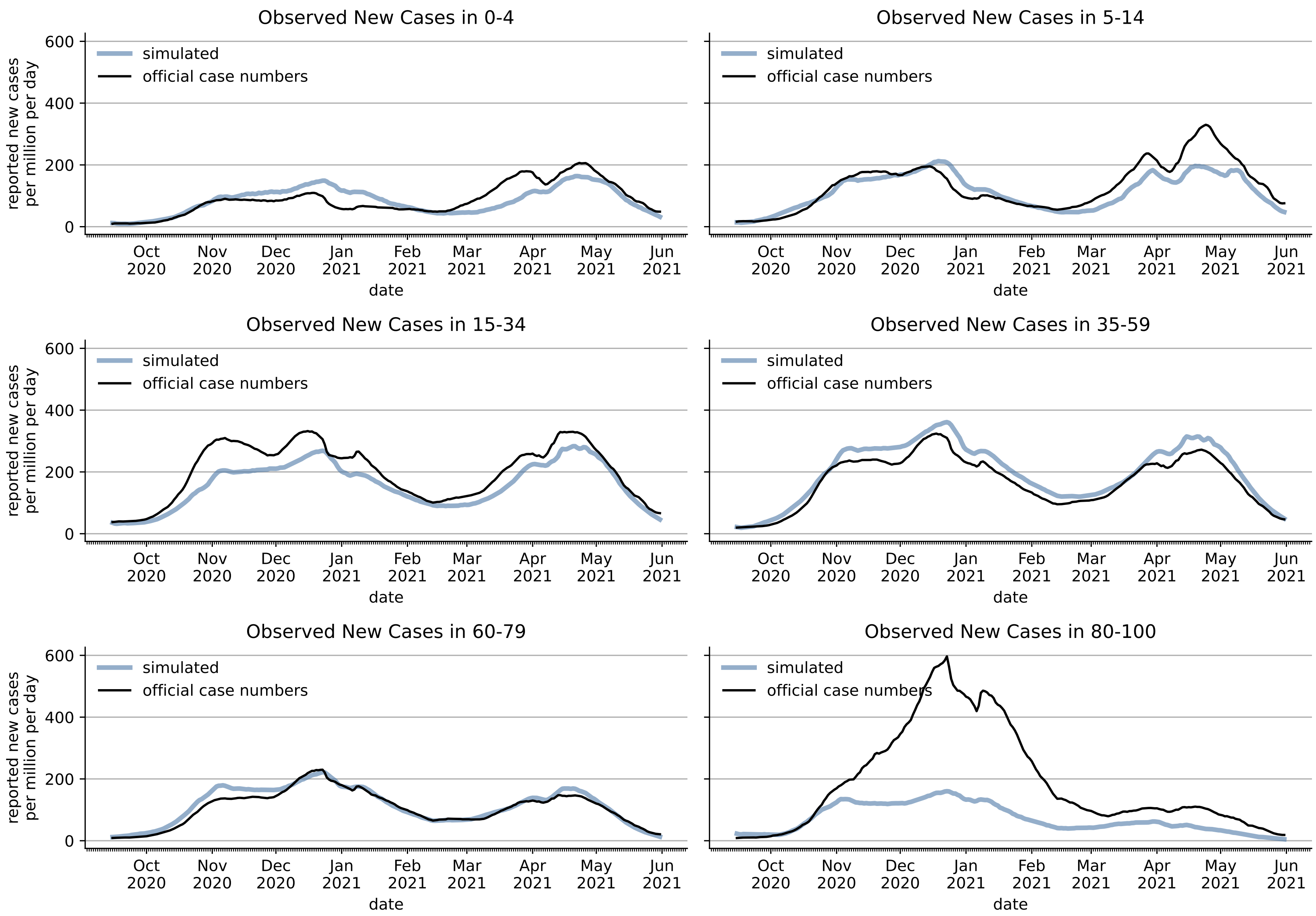


Figure 3.3.4. Simulated and Empirical Infections by Age Group

Note: The figure shows the number of reported versus simulated cases per one million people per day for different age groups. The age group of individuals above 80 needs to be interpreted with caution because our synthetic population only includes private households, i.e. nursing homes are not represented in our model. They accounted for many cases and deaths in the winter of 2020 and many 80 to 100 year olds live in these facilities. We average over 30 simulation runs. For legibility reasons, all lines are rolling 7-day averages.

Our model fit is also very good for the different German federal states (see Figures 3.3.5 and 3.3.6). This holds not only for the large states such as North Rhine-Westphalia or Bavaria but also for many smaller states such as Hessen or Rhineland-Palatinate. This shows that using school vacations dates and work mobility reductions by Google, LLC (2021) at the state level combined with county and age group specific initial conditions (see Section 2.10) and county level assortativity of contacts is sufficient to represent many local differences. The fit is especially good given that our model does not aim to have a high local resolution. For example we abstract from population density and cross-border travel. It is, thus, unsurprising that there are states that we do not match well, such as very thinly populated Mecklenburg-Vorpommern and Schleswig-Holstein or Saxony with its large border to the Czech Republic that had a much higher incidence than Germany at times.

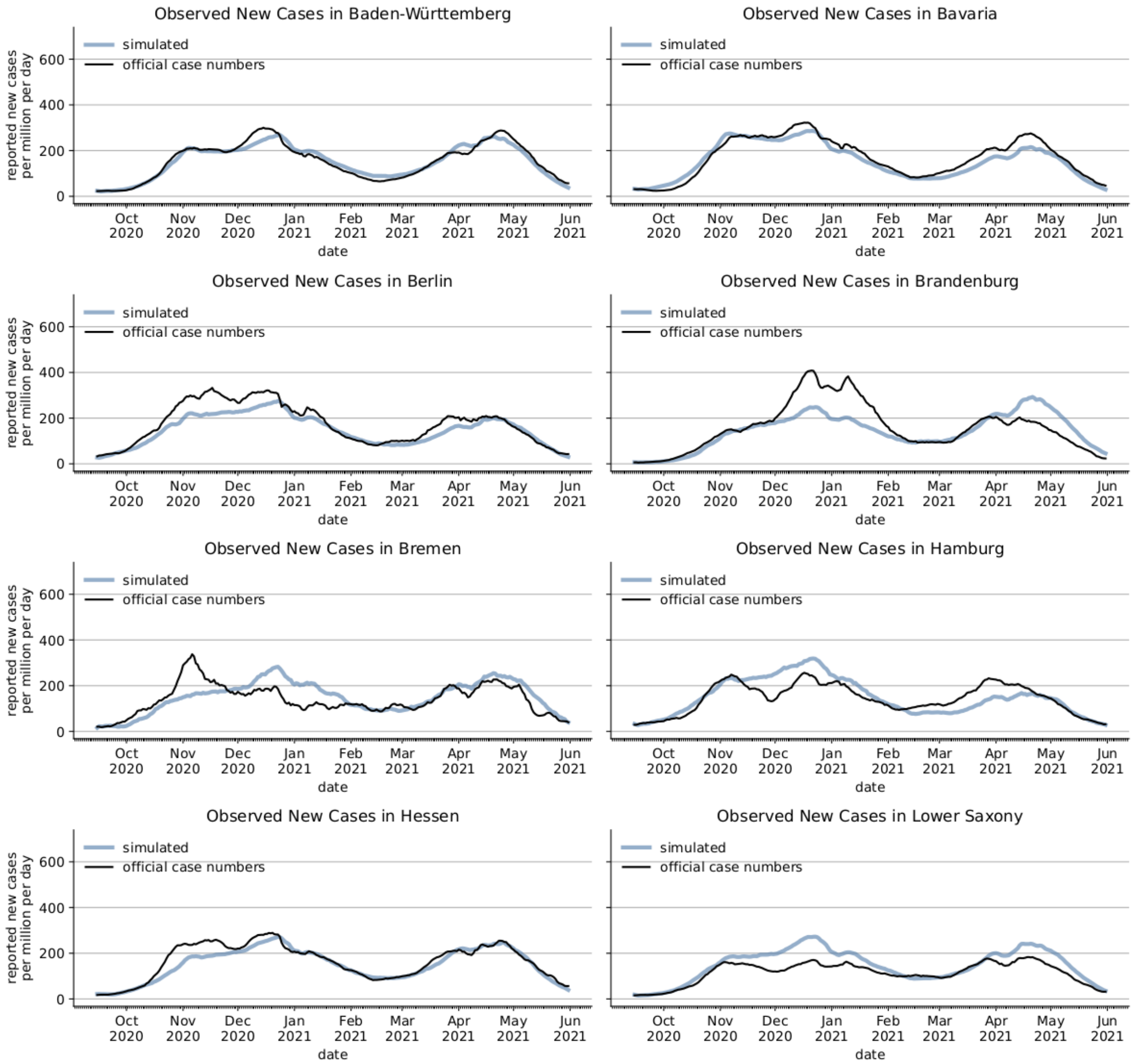


Figure 3.3.5. Simulated and Empirical Infections by Federal State (1)

Note: The figure shows the number of reported versus simulated cases per one million people per day for different federal states. We averaged over 30 simulation runs. For legibility reasons, all lines are rolling 7-day averages.

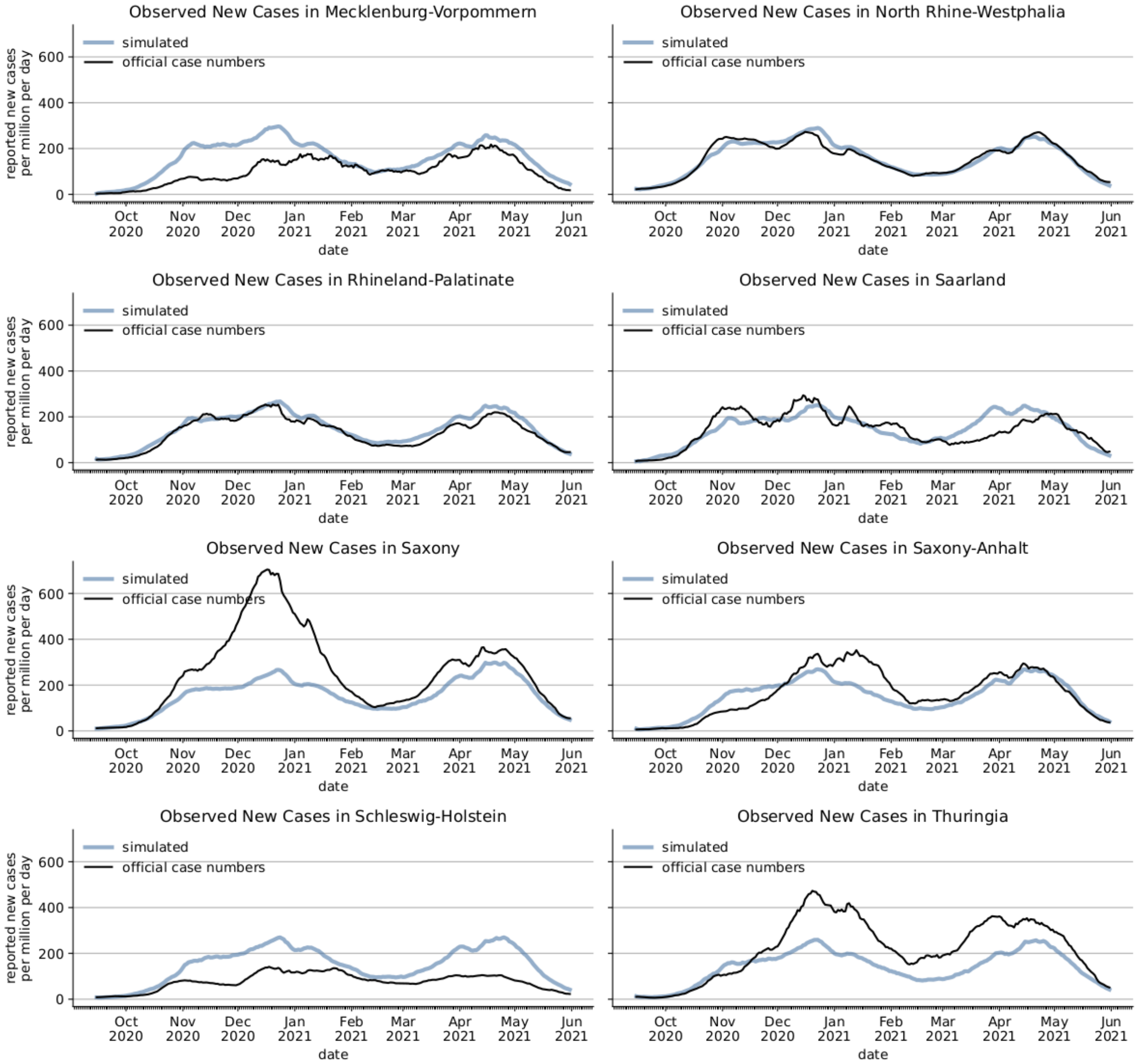


Figure 3.3.6. Simulated and Empirical Infections by Federal State (2)

Note: The figure shows the number of reported versus simulated cases per one million people per day for different federal states. We averaged over 30 simulation runs. For legibility reasons, all lines are rolling 7-day averages.

We fit the proliferation of the Alpha variant quite exactly despite only introducing a few cases in January ($\omega_{Alpha,t}$) as can be seen in Figure 3.3.7. Since we only model Alpha and do not include other variants, Alpha reaches a share of nearly 100% by May while the true rate plateaued at 90%. By the end of May Delta gained traction in Germany. However, given that it made up less than 5% at the end of our simulation period, we did not include it in our model.

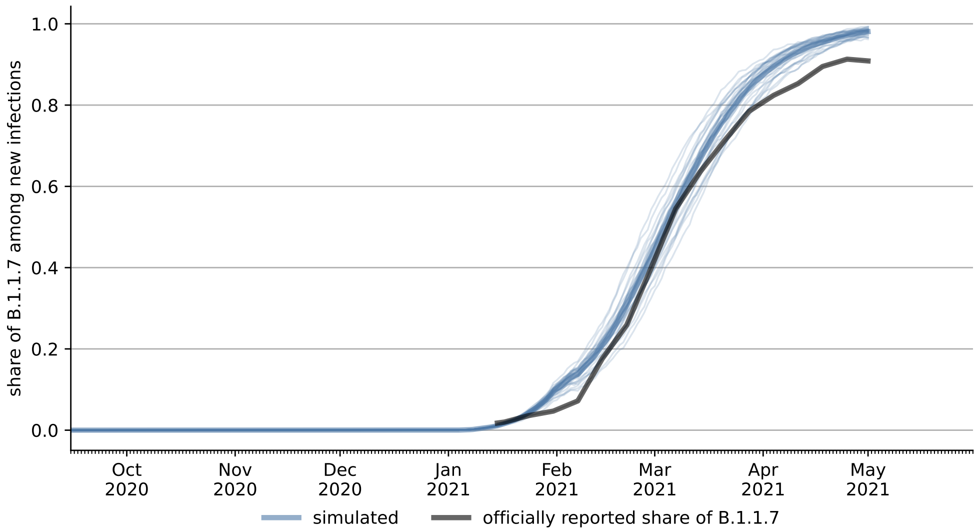


Figure 3.3.7. Share of the Alpha Variant (B.1.1.7) in the Model and as Reported by the RKI

Note: The figure shows the share of Alpha as reported by the RKI and as it arises in our simulations. We only introduce a few cases over the course of January. From then Alpha takes over endogenously through its increased infectiousness (σ_{Alpha}).

The fit of the share of vaccinated individuals can be seen in Figure 3.3.8. In Germany, vaccines were rolled out according to four priority groups. The first vaccines were mostly reserved for nursing homes and some selected professions such as first responders. Since we do not have nursing home inhabitants in our model, we subtract the first percent of vaccinations which is equivalent to the share of Germans living in nursing homes. Afterwards, the share of vaccinated individuals in the population follows the German increase exactly. We took great care to model the prioritization of older individuals and professions that cannot reduce physical contact easily such as teachers or medical staff (see Section 3.2.2 and Figure 3.2.2 for the vaccination rates in our model by age group).

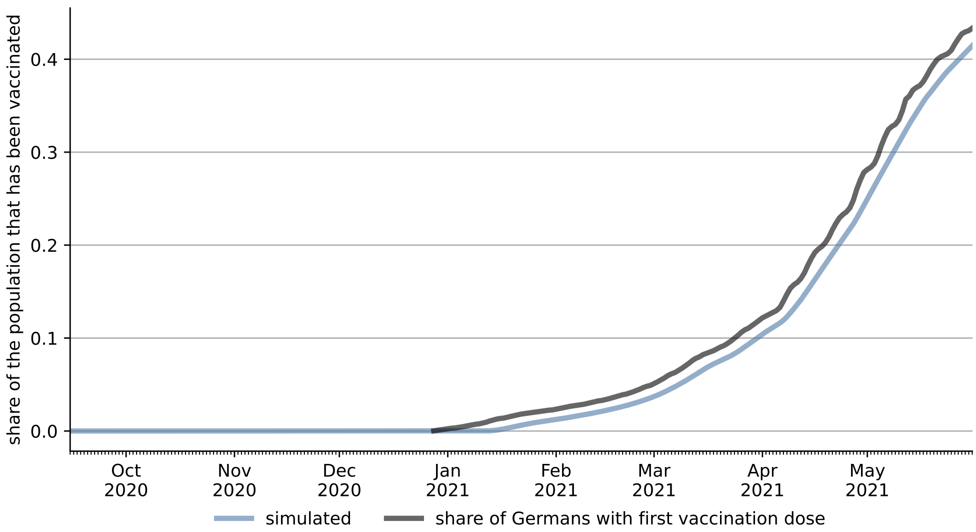


Figure 3.3.8. Share of Vaccinated Individuals in the Model and the German Population

Note: The figure shows the rate of individuals that are vaccinated in our synthetic population versus in the general German population.

The most difficult moment to match in our model is the rapid test demand. This is because we have five different channels through which individuals demand rapid tests and many of the demand curves are at least partially calibrated through survey data. It is therefore very reassuring that we fit the share of individuals that do weekly rapid tests almost perfectly (Figure 3.3.9).

For the share of individuals that have ever done a rapid test (Figure 3.3.10) our model is conservative. There are two reasons for this: Firstly, we do not model people who have done rapid tests out of curiosity once they became available. Secondly, in the model, the decision to take a rapid test is based on a time invariant individual specific compliance factor without any additional random components. While this captures important features of rapid test demand it abstracts from people who turn down rapid tests most of the time but accept them sometimes.

Appendix 3.E shows that our results are quite robust to some variation in the parameters which are not known ex ante, like the work mobility multiplier. This provides evidence that the model is robust enough to be suitable for predictions.

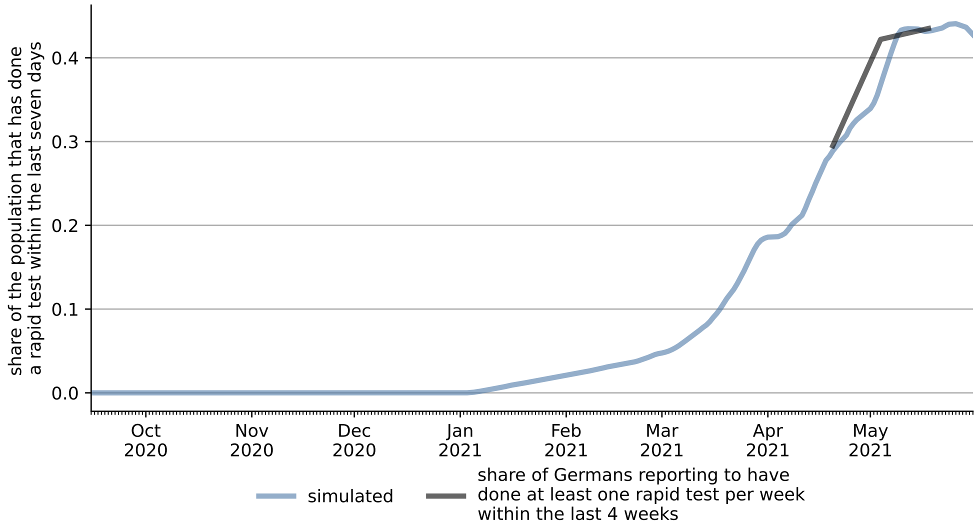


Figure 3.3.9. Share of the Population That Did a Rapid Test in the Last Week

Note: compares the share of individuals who have done a rapid test within the last seven days in our simulation compared to the share reporting to have done at least weekly rapid tests in the last four weeks in the COSMO survey (Betsch, Wieler, et al., 2021). For legibility reasons, all simulated lines are rolling 7-day averages.

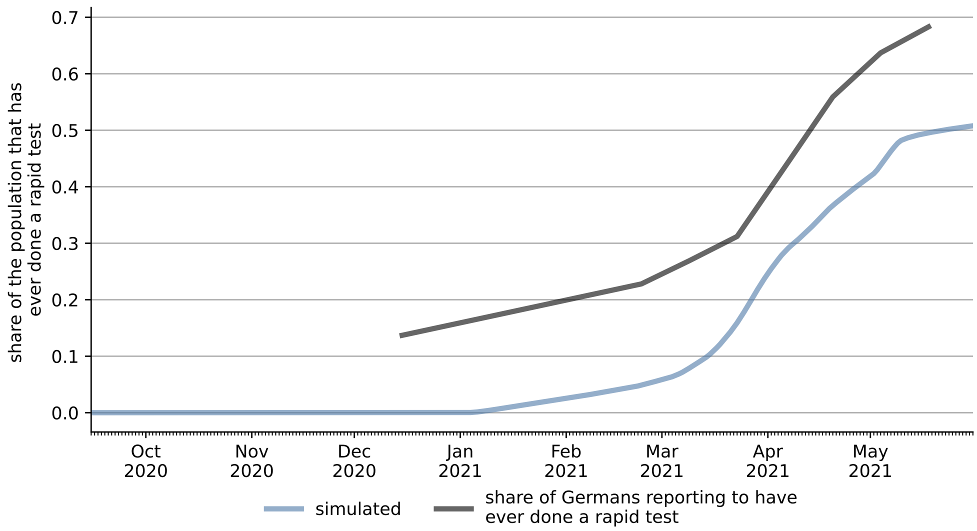


Figure 3.3.10. Share of the Population That Has Ever Done a Rapid Test

Note: This figure compares the empirical and simulated share of individuals that have ever done a rapid test. The empirical data comes from Betsch, Wieler, et al. (2021). For legibility reasons, all simulated lines are rolling 7-day averages.

3.4 Results

3.4.1 The Role of Rapid Tests, Vaccinations and Seasonality

In order to better understand the contributions of rapid tests, vaccinations, and seasonality on the evolution of infections in 2021, we consider various scenarios. Figure 3.4.1 shows our baseline (the blue line, same as in Figure 3.3.3a), a scenario without any of the three factors (red line), and three scenarios turning each of these factors on individually. NPIs are always held constant at their values in the baseline scenario. Figure 3.4.2 does the same for total infections in the model. Figure 3.4.3 employs Shapley values (Shapley, 2016) to decompose the difference in total infections between the scenario without any of the three factors and our main specification.[11]

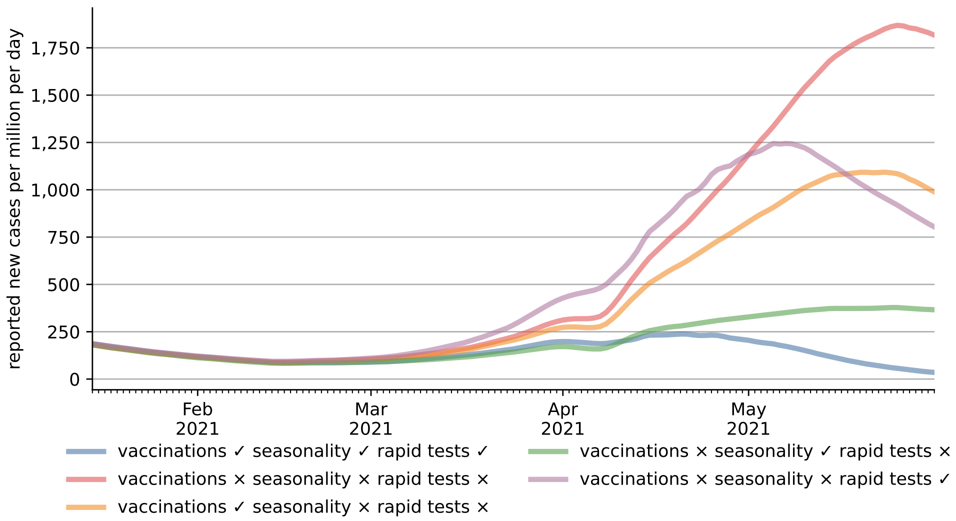


Figure 3.4.1. Recorded Cases: 2021 Scenarios

Note: The blue line shows the number of detected cases in our baseline scenario. The red line refers to a situation where NPIs evolve as in the baseline scenario and the Alpha variant is introduced in the same way but vaccinations, rapid tests, and seasonality remain at their January levels. The other scenarios turn each of these three factors on individually. For legibility reasons, all lines are rolling 7-day averages.

Until mid-March, there is no visible difference between the different scenarios. Seasonality hardly changes, and only few vaccinations and rapid tests were administered. Even thereafter, the effect of the vaccination campaign is surprisingly small at first sight. Whether considering recorded or total infections with only one channel active, the final level is always the highest in case of the vaccination campaign

11. An explanation of Shapley values can be found in Appendix 3.D.

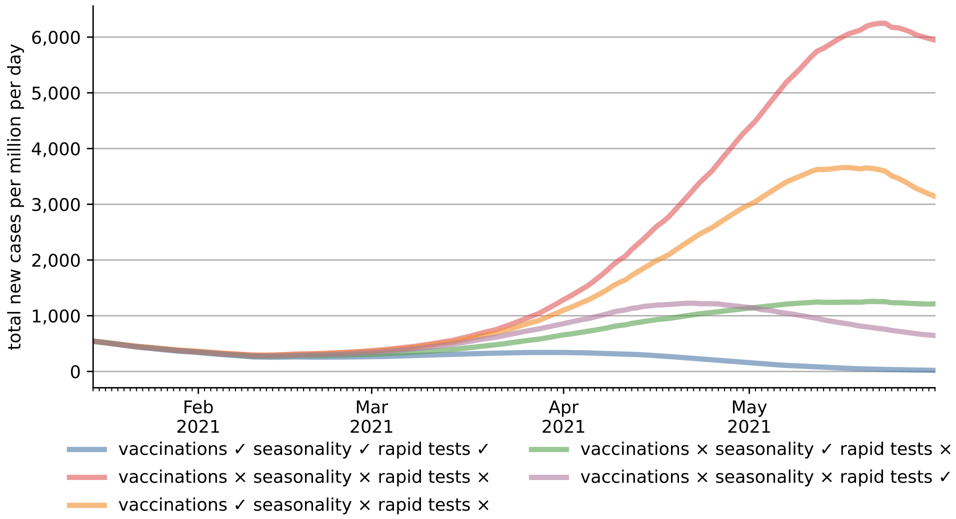


Figure 3.4.2. Total Cases: 2021 Scenarios

Note: The blue line shows the number of total infections in our baseline scenario. The red line refers to a situation where NPIs evolve as in the baseline scenario and the Alpha variant is introduced in the same way but vaccinations, rapid tests, and seasonality remain at their January levels. The other scenarios turn each of these three factors on individually. For legibility reasons, all lines are rolling 7-day averages.

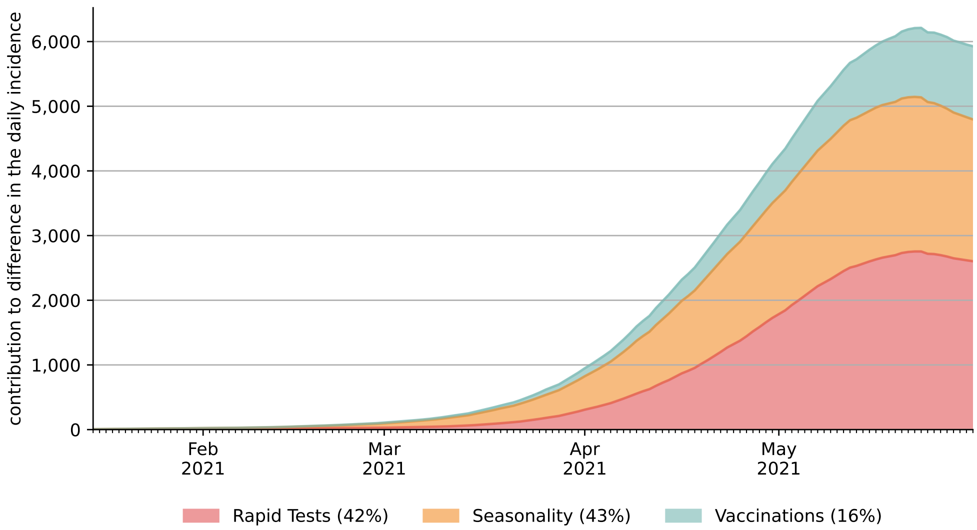


Figure 3.4.3. Decomposition of the Difference in Total Cases Between the Scenario Without Any of the Three Factors and the Baseline Scenario

Note: Each area shows the contribution – measured as the Shapley value – of each factor to avoided infections relative to a scenario without seasonality, vaccinations and rapid tests. The avoided infections are the difference between the red and the blue line in Figure 3.4.2. Shapley values are explained in Appendix 3.D.

(orange lines). The Shapley value decomposition shows that vaccinations contribute 16% to the cumulative difference between scenarios. Reasons for the low share are the slow start – it took until March 24th until 10% of the population had received their first vaccination, the 20% mark was reached on April 19th – and the focus on older individuals. These groups contribute less to the spread of the disease than others due to a lower number of contacts. By the end of our study period, when first-dose vaccination rates reached 43% of the population, the number of new cases would have started to decline. It is important to note that the initial focus of the campaign was to prevent deaths and severe disease. Indeed, the case fatality rate was considerably lower during the third wave when compared to the second – 4.4% between October and February and 1.4% between March and the end of May, (Robert Koch-Institut, 2021c).

Seasonality has a large effect in slowing the spread of SARS-CoV-2. By May 31, both observed and total cases are only a fourth of those in the scenario with seasonality compared to the scenario without vaccinations, rapid tests and seasonality. However, in this period, cases would have kept on rising throughout, just at a slower pace (this is in line with results in Gavenčiak et al., 2021, which our seasonality measure is based on). Nevertheless, we estimate seasonality to be a quantitatively important factor determining the evolution of the pandemic, explaining most of the early changes and 43% of the cumulative difference by the end of May.

A similar-sized effect – 42% in the decomposition – comes from rapid testing.[12] Here, it is crucial to differentiate between recorded cases and actual cases. Additional testing means that otherwise undetected infections are additionally recorded.[13] Figure 3.4.1 shows that this effect is large and may persist for some time. Until late April, recorded cases are higher in the scenario with rapid testing alone when compared to the setting where none of the three mechanisms are turned on. The effect on total cases, however, is visible immediately in Figure 3.4.2. Despite the fact that only 10% of the population performed weekly rapid tests in March on average, new infections in early April would have been reduced by a half relative to the scenario without vaccinations, rapid tests, or seasonality.

So why is rapid testing so effective? In order to shed more light on this question, Figure 3.4.4 decomposes the difference in the scenario without rapid tests and the main specification into the three demand channels for rapid tests. Tests at schools have the smallest effect, which is largely explained by schools not operating at full

12. For a comparison of our effect size to other studies see Appendix 3.F.

13. See Appendix 3.G for how the share of detected cases develops over time for different age groups.

capacity during our period of study and the relatively small number of students.[14] Almost 40% come from tests at the workplace. Despite the fact that rapid tests for private reasons are phased in only in mid-March, they make up for more than half of the total effect. The reason lies in the fact that a substantial share of these tests is driven by an elevated probability to carry the virus, i.e., showing symptoms of CoViD-19 or following up on a positive test of a household member. The latter is essentially a form of contact tracing, which has been shown to be very effective (Kretzschmar et al., 2020; Contreras et al., 2021; Fetzer and Graeber, 2021). Indeed, a deeper analysis in Appendix 3.H shows that the same amount of rapid tests administered randomly in the population would not have been nearly as effective.

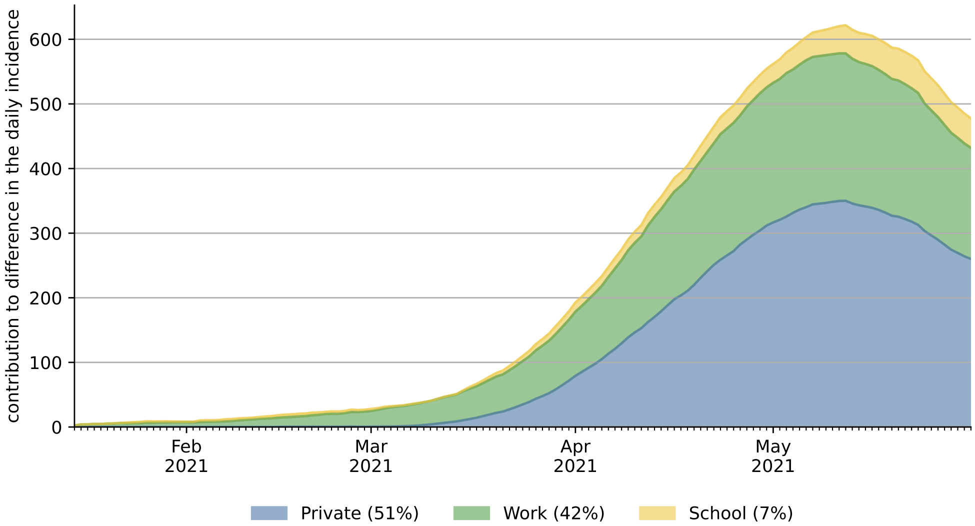


Figure 3.4.4. Decomposition of the Difference in Total Cases Between the Scenario Without Rapid Tests and the Baseline Scenario

Note: Each area shows the contribution – measured as the Shapley value – of each rapid test chanel to avoided infections relative to a scenario without any rapid tests. Shapley values are explained in Appendix 3.D.

3.4.2 The Suitability of Rapid Tests as a Screening Device

In this section, we show how the usage of rapid tests expands in our simulations over time. To make the most out of little available data, we model actual rapid test usage as the result of an interplay between time and policy invariant demand preferences with a changing epidemiological environment as well as a changing supply

14. 18% of our population are in the education sector (pupils, teachers, etc.); 46% are workers outside the education sector.

curve. As a result, the number of performed rapid tests is only indirectly determined through our calibrated parameters and the simulated number of rapid tests becomes an interesting result of the model.

The share of the population doing a rapid test and receiving a positive rapid test over time by the channel through which the test was demanded is shown in Figures 3.4.5. Overall, the share of the population getting a rapid test on a given day increases from 2% in mid March to over 10% by May. The rapid tests at the workplace are a little ragged because of public holidays. For rapid tests in the education sector both vacations (first half of April) as well as the opening of schools in May are very visible in the rapid test demand. Overall, work tests make up the largest fraction of rapid tests. The share of private rapid tests roughly aligns with numbers reported by Germany's federal states on the demand for the free citizens' tests (Land Nordrhein-Westfalen, 2021).[15]

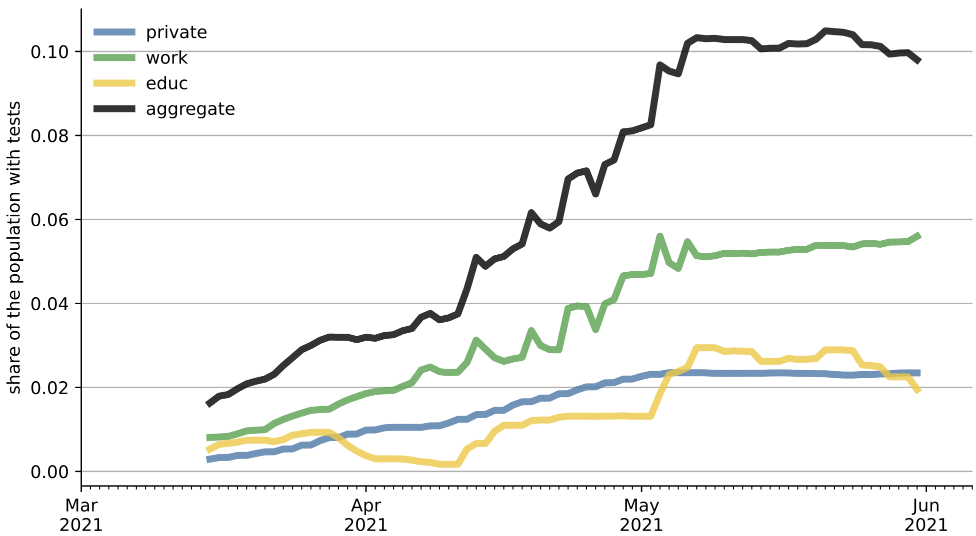


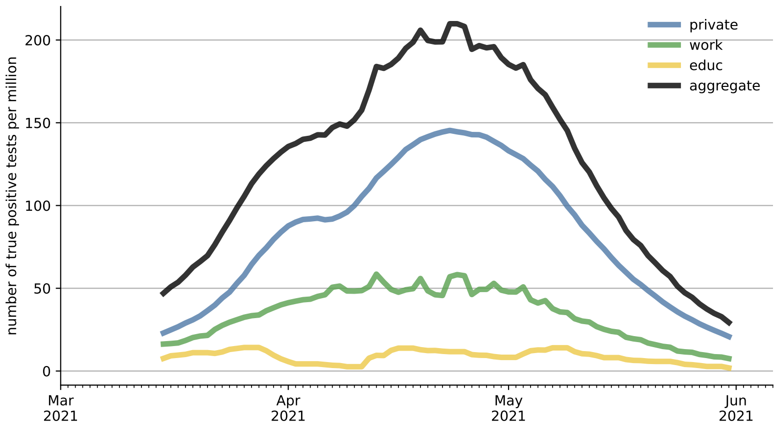
Figure 3.4.5. Share of the Population Doing a Rapid Test on a Given Day by Channel

Note: Each line shows the share of the population doing a rapid test on a given day through one of the three channels (education in yellow, work in green and private in blue) as well as the aggregate (black) share. Rapid tests in the education sector are demanded by teachers (nursery, preschool and school) as well as pupils. After Easter the required frequency of tests is increased from once per week to twice per week. Rapid tests at the workplace are demanded by individuals that still have work contacts, i.e. do not work from home. The share of employers offering rapid tests increases over time and the frequency of testing is also increased. Private rapid tests are demanded by individuals for three potential reasons: having developed symptoms without access to a PCR test, having a household member that has tested positive or developed symptoms or planing to participate in a weekly other contact type meeting (such as weekly soccer training).

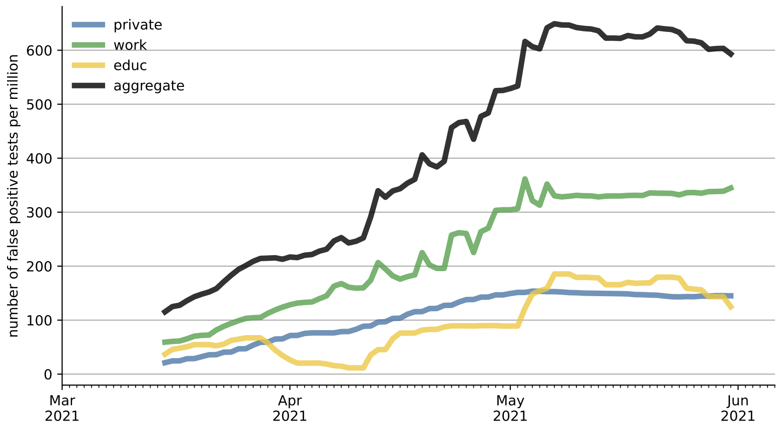
15. Note that our private test demand includes self-administered tests which are not included in the statistics on the free citizens' tests.

A frequently raised concern about rapid tests is that they are likely to miss latent infections and sometimes sound false alarms. Both factors raised doubts how effectively they can curb infections (The Brussels Times, 2020; Rehrmann, 2021). Since we model these imperfections very carefully, we can now evaluate how well rapid tests work as a screening device. To do so we show the number of tests split by whether they are true positive, false positive, true negative or false negative (see Figure 3.4.6) in numbers per million individuals to make the metric comparable to incidences.

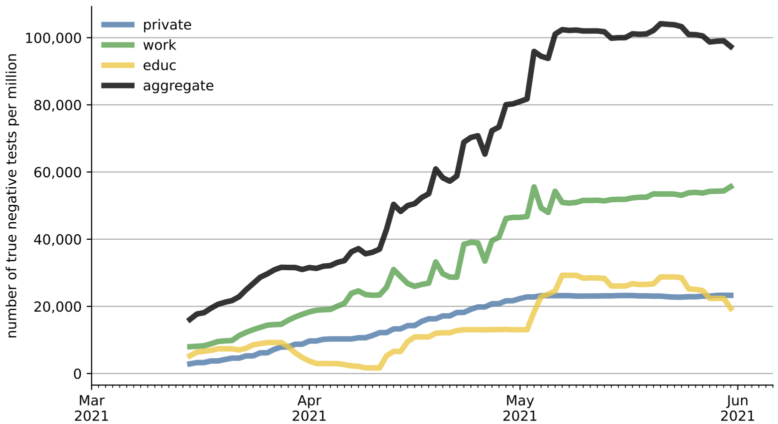
The number of true positives (Figure 3.4.6a) rapidly increases and peaks at the end of April with over 200 cases per million detected through rapid tests per day. This means that – according to our model – Germany was able to detect up to 16,600 cases per day that would have likely gone undetected otherwise. The most powerful tool for detecting cases are the private rapid tests. This is because a large share of them are targeted, i.e. triggered by events in the household. However, this does not mean that rapid tests in the workplace or at school are less important. It is rather the combination of large scale screening at work and in schools and very efficient follow up tests whenever those screening tests uncover a case. The Shapley values (Figure 3.4.4) take this into account and assign about 50% of the overall reduction of case numbers via rapid tests to private rapid tests with work and school rapid tests accounting for 40% and 7%, respectively.



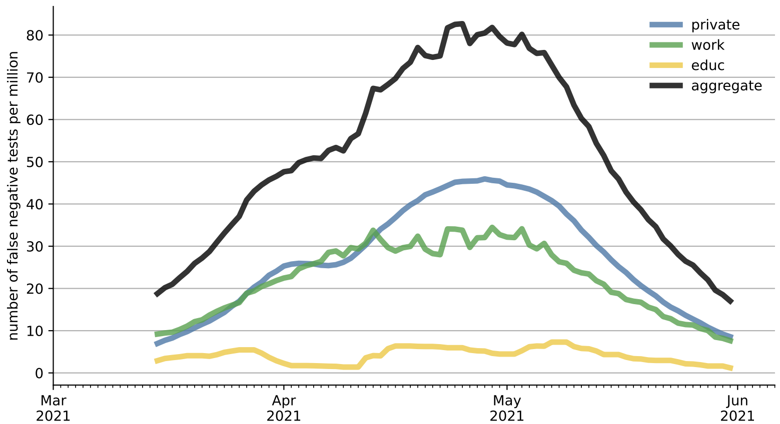
(a) Number of Discovered Cases Due to Rapid Tests by Channel



(b) Number of False Positive Rapid Tests by Channel



(c) Number of True Negative Rapid Tests by Channel



(d) Number of False Negative Rapid Tests by Channel

Figure 3.4.6. Rapid Test Statistics

Note: Each panel shows the number of rapid tests per million inhabitants that fall into the respective category. Private rapid tests are especially good at detecting cases but since they are often triggered by rapid tests from other channels, the other groups of tests, especially rapid tests at the workplace, also play an important role for containing the pandemic. All results are averaged over 30 simulation runs. For legibility reasons, all lines are rolling 7-day averages.

Such a large effect of rapid tests seems to be at odds with the general perception that they are not very reliable. However, one has to differentiate between the reliability of one test in isolation and the effect imperfect tests can have when employed at a large scale. On average our tests have a sensitivity of slightly more than 70%. This means they miss almost 30% of infections among the tested. Of course perfect tests would have an even larger effect. Then up to an additional 80 cases per million would have been detected per day as the incidence of false negatives (Figure 3.4.6d) shows. However, the relevant number to compare is that up to 200 cases per million are detected by rapid tests every day which would have otherwise gone undetected.

Another common argument against rapid tests as a screening device are the costs that false positive rapid tests cause. Falsely positive individuals – hopefully – self-isolate until a PCR test refutes their rapid test result. So how many individuals would have had to self-isolate and wait for a PCR test despite not being infected for each detected case in our model? This is shown in Figure 3.4.7. Until May the number is always below five. When the incidence is very high as in January, only up to 2.5 individuals erroneously receive a positive rapid test result for each detected case. When the incidence falls that number strongly increases. By end of May, for each detected case there are 20 false positive tests.

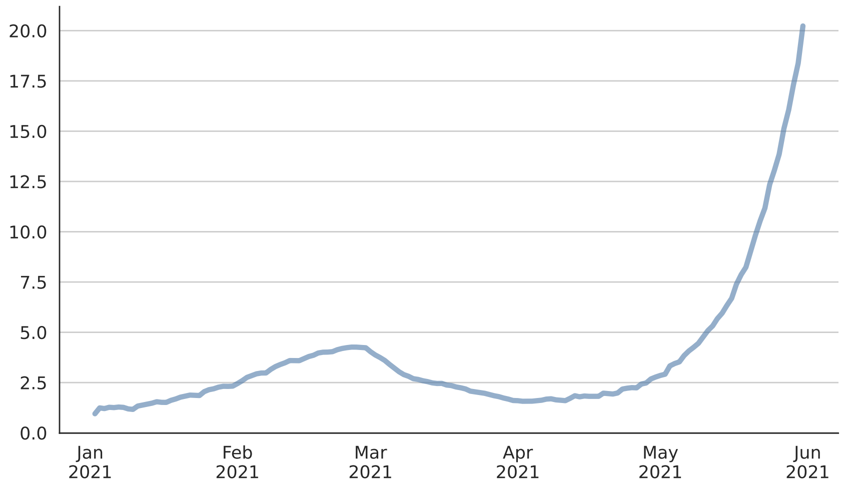


Figure 3.4.7. Number of False Positives per Detected Case

Note: The line shows how many individuals receive an erroneously positive rapid test per detected case over time. Rapid tests are imperfectly specific and their sensitivity varies over the course of the infection. All results are averaged over 30 simulation runs. For legibility reasons, all lines are rolling 7-day averages.

Another way to present the costs of false positive tests is the false positive rate, i.e. the share of positive tests that go to people who are not infected. Figure 3.4.8 shows that the false positive rate is very high. On average 60% to 93% of positive tests are received by individuals that are not infected which is broadly in line with rates later reported in the German media (Süddeutsche Zeitung, 2021). The false positive rate increases over time. This is due to the falling prevalence of infections in the population. Again, private rapid tests are an exception with a much lower false positive rate because those tests are primarily demanded when there is a high likelihood of being infected. The false negative rate of 0.2% looks very low. As discussed above this is deceiving and just a mechanical consequence of a very low prevalence of the disease and the many rapid tests done by non-infected people.

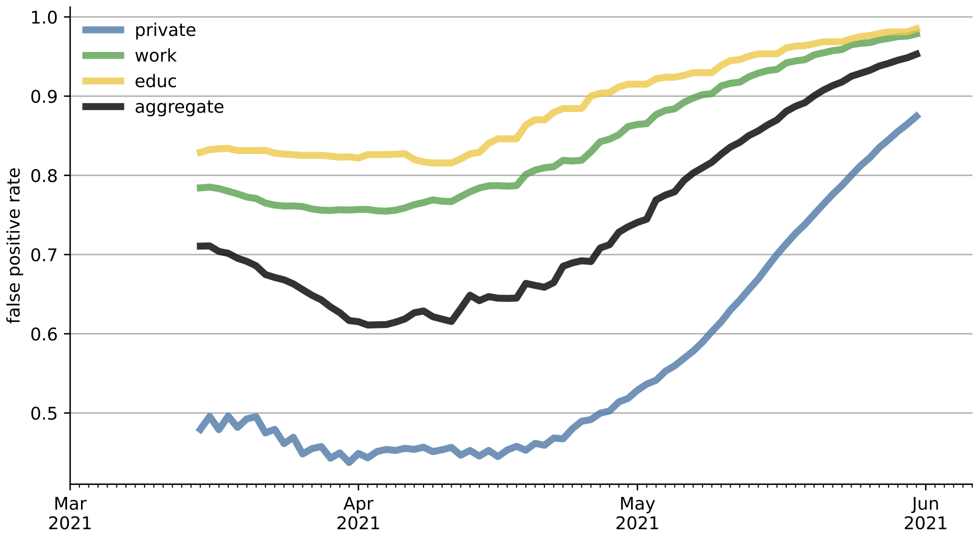


Figure 3.4.8. Rate of False Positive Rapid Tests by Channel

Note: The false positive rate is the share of positive tests that are given to people who are not infected. This share is large as can be expected with a very low baseline rate of positive individuals. As the incidence in the population drops, the false positive rate increases. An exception are the private rapid tests because they are – especially when the incidence is high – often triggered by events that make it likely that the test taker is infected and therefore their false positive rate is much lower. All results are averaged over 30 simulation runs. For legibility reasons, all lines are rolling 7-day averages.

These numbers show that the large effect of rapid tests on the infection dynamic is not driven by unrealistic assumptions about their sensitivity and specificity but rather by the fact that there was a very large number of infected individuals who did not (yet) know they were infected when the rapid test informed them. Detecting and isolating some of them is enough to slow down the overall infection dynamic and very cost effective when incidences are high.

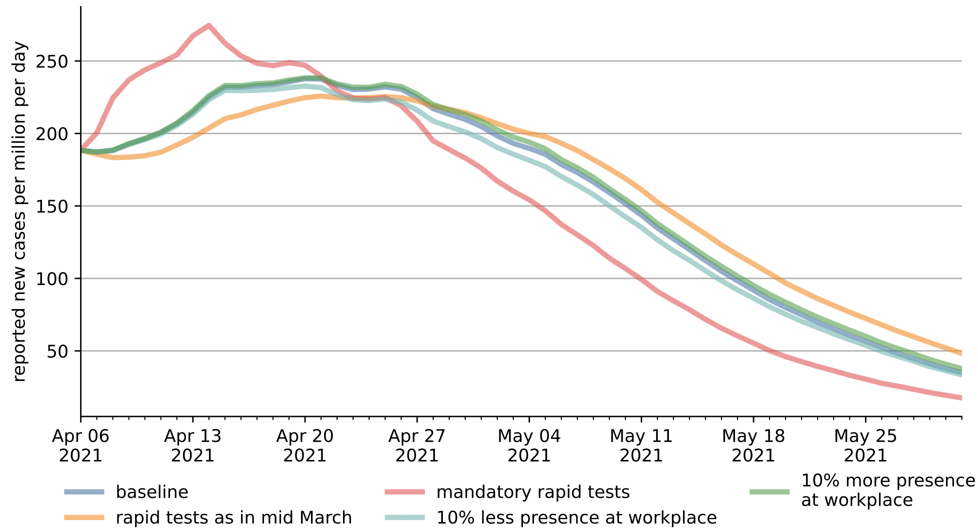
3.4.3 Can Rapid Tests Substitute for NPIs?

Lastly, we complement our analysis on the effectiveness of rapid tests by showing the effects of rapid test policies vis-à-vis two of the most contentious NPIs: work from home mandates and school closures. All scenarios start after Easter (April 6). Our analyses show that socially costly NPIs can be avoided through strong rapid testing policies.

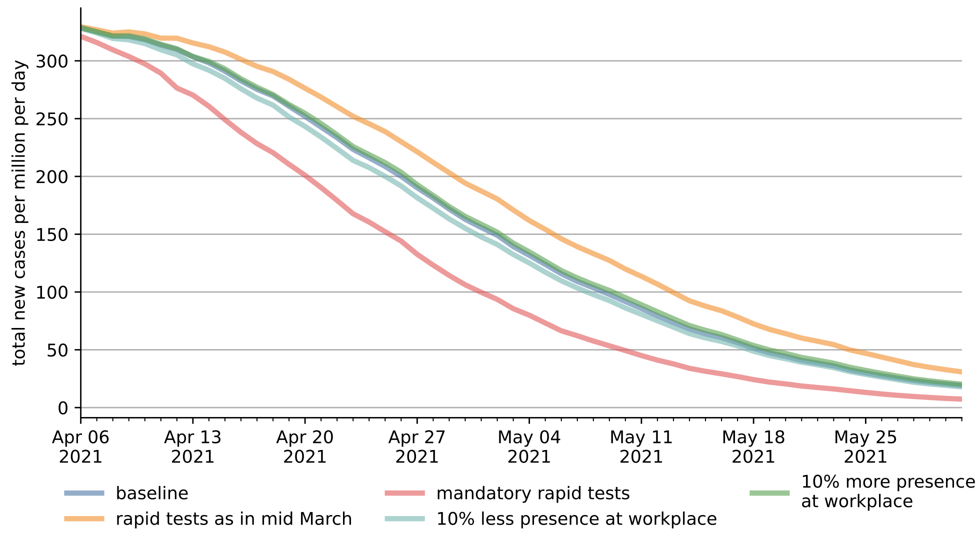
Figure 3.4.9 shows the effects of different work policies on the infections in the general population. We compare four scenarios with our baseline scenario: Keeping the share of workers having physical work contacts the same as in our baseline scenario, the orange line shows what would have happened with rapid testing in firms at the level of mid March (orange line) where only 14% of workers regularly did rapid tests. We also include a scenario what would have happened if rapid tests had become truly mandatory after Easter,[16] assuming a 95% compliance rate on both the employer and the employee side. On the work from home dimension we compare our baseline scenario with 10% more or less work from home compared to the baseline scenario.

Figure 3.4.9 shows that with a large fraction of workers receiving tests, testing at the workplace has larger effects than mandating employees to work from home. Whether the share of workers working at the usual workplace is reduced or increased by ten percent changes infection rates by 2.5% or less in either direction. Making testing mandatory twice a week – assuming independent compliance by employers and workers of 95% each – on the other hand would have reduced infections by 23%. Reducing rapid tests offers by employers to the level of March would have increased infections by 13%.

16. Starting on April 19th employers were required by law to provide two weekly tests to their employees (Bundesanzeiger, 2021b). However, voluntarily only 60% of workers regularly test themselves when offered tests (Betsch, Korn, Felgendreff, Eitze, Schmid, Sprengholz, Wieler, Schmich, Stollorz, Ramharter, Bosnjak, Omer, Thaiss, De Bock, and Von Rüden, 2021).



(a) Reported Cases



(b) Total Cases

Figure 3.4.9. The Effect of Different Work Scenarios on Reported and Total Cases

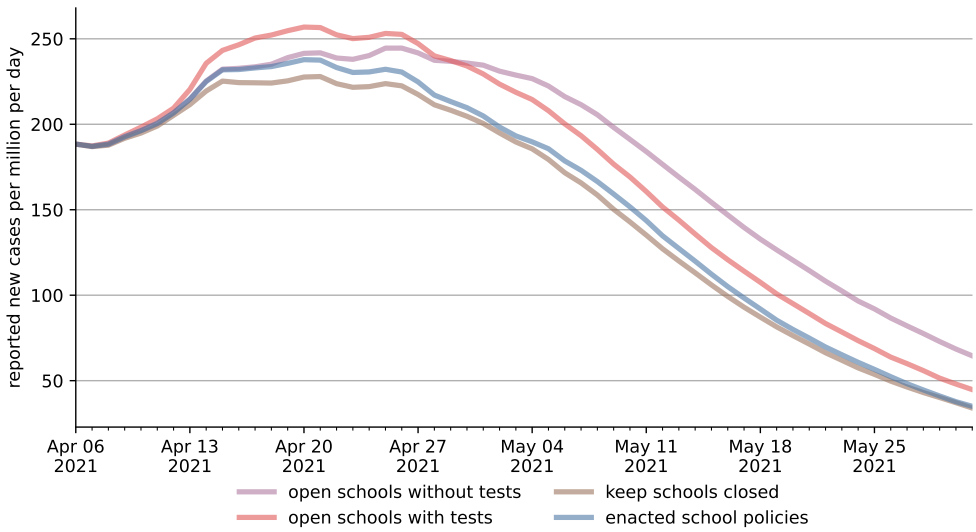
Note: The figure shows the development of cases after different hypothetical work policy changes take place at Easter until the end of our simulation period. We vary the share of workers that have physical work contacts (10% more or less compared to the share in the baseline scenario, 85% or 70% of workers, respectively) and how many tests are performed at work relative to our baseline scenario. As an ambitious scenario we implement mandatory tests for all employees that do not work from home, assuming 95% compliance on both the employer and the employee side. On the other hand, we show what would have happened if the test offers had fallen back to the level of mid March (only 14% of workers are tested regularly). The observed cases can be misleading because more testing leads to more detected cases.

The second commonly employed and also very contentious NPI we study are school closures. In Germany, as in many countries, schools switched to remote instruction to reduce infections. In April 2021 German schooling mostly consisted of generous emergency care with rotating on-site schooling for graduating classes due to the high incidence. In May where cases fell and schools gradually opened, we model the policy as rotating on-site schooling for most students (except for children eligible for emergency care and graduating classes who attend in full). We compare this baseline scenario to simply keeping schools completely closed (the brown line) and opening schools normally with and without tests (but maintaining our hygiene multiplier to account for mask wearing, ventilation etc.), the red and violet line, respectively.

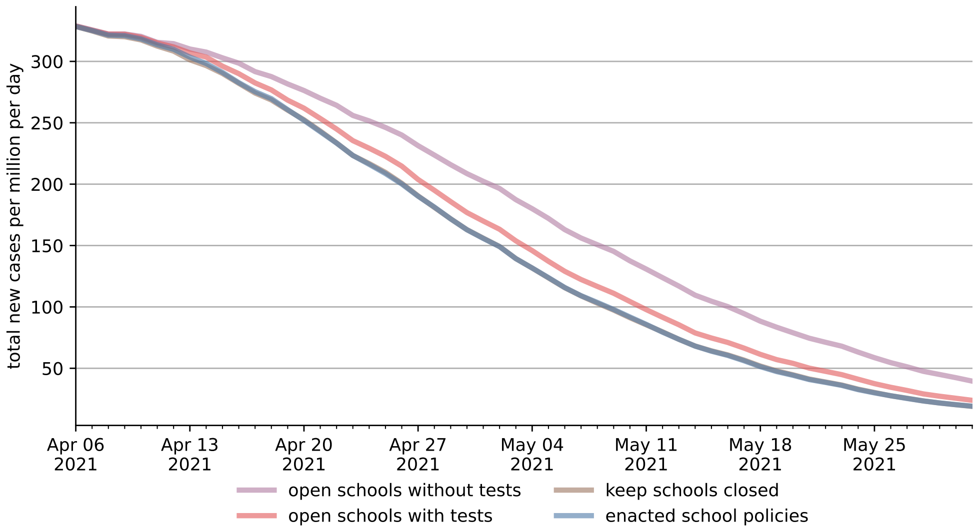
As can be seen in Figure 3.4.10 we estimate the realized scenario to have essentially the same effect as a situation with closed schools. Under fully opened schools with mandatory tests, total infections would have been 6% higher; this number rises to 20% without tests. These effect sizes are broadly in line with empirical studies (Berger, Fritz, and Kauermann, 2021; Vlachos, Hertegård, and B. Svaleryd, 2021). In light of the large negative effects school closures have on children and parents (Luijten et al., 2021; Melegari et al., 2021) – and in particular on those with low socio-economic status – these results in conjunction with hindsight bias suggest that opening schools combined with a testing strategy would have been beneficial. Tests, however, are crucial here. Had schools opened completely without any testing of students and staff this would have added up to 50 incidence points.

Our analysis has shown that during the transition to high levels of vaccination and possibly thereafter, large-scale rapid testing can substitute for some NPIs. This comes at a fraction of the cost. A week of the fairly strict lockdown in early 2021 is estimated to have cost around 20-30 Euros per capita (Wollmershäuser, 2021); retail prices for rapid tests were below one Euro in early June 2021 and below five Euros for firms. While we do not distinguish between self-administered rapid tests and point of care rapid tests, the former are likely to play a larger role for indication-driven testing. Widespread availability at low prices seems important. However, they rely on purely voluntary participation in a non-public setting. The benefit of point-of-care rapid tests as a precondition to participate in leisure activities as well as mandatory tests at the workplace or at school come from screening the entire population.

This is important because disadvantaged groups are less likely to be reached by testing campaigns relying on voluntary participation (e.g. Stillman and Tonin, 2021); at the same time, these groups have a higher risk to contract CoViD-19 (Robert Koch-Institut, 2021a). Mandatory tests at school and at the workplace will extend more into these groups. The same goes for individuals who exhibit a low level of compliance with CoViD-19-related regulations. Compared to vaccinations, rapid testing



(a) Reported Cases



(b) Total Cases

Figure 3.4.10. The Effect of Different School Scenarios on Reported and Total Cases

Note: The figure shows the development of cases after different hypothetical school policy changes take place at Easter until the end of our simulation period. Apart from the enacted school policies as our baseline we simulate how cases would have developed if schools had been closed completely as the strictest possible counterfactual scenario and two opening models: One where schools open normally (with hygiene measures) without any testing in the education sector and one where schools open normally but testing develops as in the baseline scenario.

programmes allow a much quicker roll-out, making it arguably the most effective tool to contain a pandemic in the short run.

3.5 Conclusion

This chapter has used the agent-based model from Chapter 2 to identify the role various policies have played in Germany's infection dynamics in spring 2021.

We show that most of the model parameters can be calibrated in a real world setting. Furthermore, the small set of remaining parameters can be estimated from empirically observed cases due to our unique model of case detection. The fitted model is able to reproduce two waves, the proliferation of a more infectious variant and match different infection patterns in different age groups.

Given our mostly policy-invariant parameters we have run counterfactual simulations to evaluate the role different policies played during the third wave in Germany. We decompose the influence of seasonality, vaccinations and rapid tests during that time. Contrary to popular opinion, vaccinations played a minor role in the decrease in cases, with over 40% of the joint effect being due to seasonality and rapid tests each.

Our detailed model of rapid test demand allows us to show that both screening rapid tests and contact tracing rapid tests are important factors behind this large effect. About half of the effect of rapid tests is due to the screening done by rapid tests at work and in schools. The other half is attributed to private rapid tests which are mostly triggered by symptomatic individuals and individuals with sick or positive household members. We further show that – even given their partially very imperfect testing performance – rapid tests are highly cost effective. This is especially true when incidences are high.

Going beyond the evaluation of the implemented policies, we show that stringent rapid testing can substitute for much more socially costly NPIs, such as work from home mandates and virtual schooling. According to our simulations Germany could have saved up to 50 cases per million inhabitants per day if twice weekly rapid testing had been mandatory for all employees who do not work from home. Even with the very imperfect rapid test coverage at German workplaces, whether 10% more or less individuals attended work would have had a negligible effect on the incidence. Similarly, given the thorough testing done in schools, had schools opened directly after Easter this would have only increased the national incidence by 9 cases per million per day.

Our results have important implications for countries that are still rolling out vaccines as well as for future epidemics whose pathogens share characteristics with SARS-CoV-2, such as aerosol transmission, presymptomatic infectiousness and a significant share of asymptomatic but infectious carriers – or the emergence of an immune escape variant of SARS-CoV-2. Firstly, vaccination campaigns aimed at protecting the elderly only have a limited effect on infections – even though they are able to prevent deaths – as other age groups have much more contacts and are therefore stronger drivers of the epidemic. Secondly, rapid tests – even when they are only sensitive once an individual is infectious – are a very cost-effective instrument to stop infection chains. This is especially true when infection rates are high. Therefore it is important that they are used to both screen high contact groups and to quickly reveal the infection state of individuals with an elevated probability to carry the virus. Thus, availability and incentives are key for such a policy to work. Thirdly, when rapid tests are done with a high enough frequency, they can substitute for more costly NPIs such as work from home mandates or school closures.

References

Abu-Raddad, Laith J, Hiam Chemaitelly, Joel A Malek, Ayeda A Ahmed, Yasmin A Mohamoud, Shameem Younuskunju, Houssein H Ayoub, Zaina Al Kanaani, Abdullatif Al Khal, Einas Al Kuwari, Adeel A Butt, Peter Coyle, Andrew Jeremijenko, Anvar Hassan Kaleeckal, Ali Nizar Latif, Riyazuddin Mohammad Shaik, Hanan F Abdul Rahim, Hadi M Yassine, Mohamed G Al Kuwari, Hamad Eid Al Romaihi, Mohamed H Al-Thani, and Roberto Bertollini. 2020. "Assessment of the Risk of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Reinfection in an Intense Reexposure Setting." *Clinical Infectious Diseases*, [89]

Berger, Ursula, Cornelius Fritz, and Göran Kauermann. 30, 2021. "CODAG Bericht Nr. 14: 2. Schulschließungen oder Schulöffnung mit Testpflicht? Epidemiologisch-statistische Aspekte sprechen für Schulöffnungen mit verpflichtenden Tests." Working paper. Ludwig-Maximilians-Universität München. [85, 133]

Bernal, Jamie Lopez, Nick Andrews, Charlotte Gower, Eileen Gallagher, Ruth Simmons, Simon Thelwall, Julia Stowe, Elise Tessier, Natalie Groves, Gavin Dabrera, Richard Myers, Colin N.J. Campbell, Gayatri Amirthalingam, Matt Edmunds, Maria Zambon, Kevin E. Brown, Susan Hopkins, Meera Chand, and Mary Ramsay. 2021. "Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant." *New England Journal of Medicine* 385 (7): 585–94. [90]

Bertelsmann Stiftung. 2019. "Ländermonitor Frühkindliche Bildungssysteme 2019." Working paper. Bertelsmann Stiftung. [91]

Betsch, Cornelia, Lars Korn, Lisa Felgendreff, Sarah Eitze, Philipp Schmid, Philipp Sprengholz, Lothar Wieler, Patrick Schmich, Volker Stollorz, Michael Ramharter, Michael Bosnjak, Saad B. Omer, Heidrun Thaiss, Freia De Bock, and Ursula von Rüden. 26, 2021. *Zusammenfassung und Empfehlungen Welle 37*. URL: https://projekte.uni-erfurt.de/cosmo2020/web/summary/37/ (visited on 06/27/2021). [101]

Betsch, Cornelia, Lars Korn, Lisa Felgendreff, Sarah Eitze, Philipp Schmid, Philipp Sprengholz, Lothar Wieler, Patrick Schmich, Volker Stollorz, Michael Ramharter, Michael Bosnjak, Saad B. Omer, Heidrun Thaiss, Freia De Bock, and Ursula Von Rüden. 2021. "COVID-19 Snapshot Monitoring (COSMO Germany) - Wave 41." [100, 101, 104, 105, 131]

Betsch, Cornelia, Lothar Wieler, Michael Bosnjak, Michael Ramharter, Volker Stollorz, Saad Omer, Lars Korn, Philipp Sprengholz, Lisa Felgendreff, Sarah Eitze, et al. 2021. "Germany COVID-19 Snapshot Monitoring (COSMO Germany): Rapid Tests." [121]

Bi, Qifang, Yongsheng Wu, Shujiang Mei, Chenfei Ye, Xuan Zou, Zhen Zhang, Xiaojian Liu, Lan Wei, Shaun A Truelove, Tong Zhang, Wei Gao, Cong Cheng, Xiujuan Tang, Xiaoliang Wu, Yu Wu, Binbin Sun, Suli Huang, Yu Sun, Juncen Zhang, Ting Ma, Justin Lessler, and Tiejian Feng. 2020. "Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study." *The Lancet Infectious Diseases* 20 (8): 911–19. [89]

Bundesanzeiger. 8, 2021a. "Verordnung zum Anspruch auf Testung in Bezug auf einen direkten Erregernachweis des Coronavirus SARS-CoV-2 (Coronavirus-Testverordnung - TestV)." [101]

Bundesanzeiger. 2021b. "Zweite Verordnung zur Änderung der SARS-CoV-2-Arbeitsschutzverordnung." [131]

Byrne, Andrew William, David McEvoy, Aine B Collins, Kevin Hunt, Miriam Casey, Ann Barber, Francis Butler, John Griffin, Elizabeth A Lane, Conor McAloon, Kirsty O'Brien, Patrick Wall, Kieran A Walsh, and Simon J More. 2020. "Inferred duration of infectious period of SARS-CoV-2: rapid scoping review and analysis of available evidence for asymptomatic and symptomatic COVID-19 cases." *BMJ Open* 10 (8): e039856. [88]

Chen, Jun, Tangkai Qi, Li Liu, Yun Ling, Zhiping Qian, Tao Li, Feng Li, Qingnian Xu, Yuyi Zhang, Shuibao Xu, Zhigang Song, Yigang Zeng, Yinzhong Shen, Yuxin Shi, Tongyu Zhu, and Hongzhou Lu. 2020. "Clinical progression of patients with COVID-19 in Shanghai, China." *Journal of Infection* 80 (5): e1–e6. [89]

Contreras, Sebastian, Jonas Dehning, Sebastian B. Mohr, Simon Bauer, F. Paul Spitzner, and Viola Priesemann. 2021. "Low case numbers enable long-term stable pandemic control without lockdowns." *Science Advances* 7 (41): eabg2243. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.abg2243. [125]

Davies, Nicholas G., Sam Abbott, Rosanna C. Barnard, Christopher I. Jarvis, Adam J. Kucharski, James D. Munday, Carl A. B. Pearson, Timothy W. Russell, Damien C. Tully, Alex D. Washburne, Tom Wenseleers, Amy Gimma, William Waites, Kerry L. M. Wong, Kevin van Zandvoort, Justin D. Silverman, Karla Diaz-Ordaz, Ruth Keogh, Rosalind M. Eggo, Sebastian Funk, Mark Jit, Katherine E. Atkins, and W. John Edmunds. 2021. "Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England." *Science*, (3): [87]

Davies, Nicholas G., Petra Klepac, Yang Liu, Kiesha Prem, Mark Jit, and Rosalind M. Eggo. 2020. "Age-dependent effects in the transmission and control of COVID-19 epidemics." *Nature Medicine* 26 (8): 1205–11. [87, 88]

Destatis. 2020. "Betreuungsquote von Kindern unter 6 Jahren nach Bundesländern." Working paper. Statistisches Bundesamt. [91]

Fetzer, Thiemo, and Thomas Graeber. 2021. "Measuring the Scientific Effectiveness of Contact Tracing: Evidence from a Natural Experiment." *Forthcoming, Proceedings of the National Academy of Sciences*, [125]

Forschungsdatenzentren Der Statistischen Ämter Des Bundes Und Der Länder. 2018. "Mikrozensus 2010, CF, Version 0." de. [90, 95]

Gavenčiak, Tomáš, Joshua Teperowski Monrad, Gavin Leech, Mrinank Sharma, Sören Mindermann, Jan Marcus Brauner, Samir Bhatt, and Jan Kulveit. 2021. "Seasonal variation in SARS-CoV-2 transmission in temperate climates." *medRxiv*, eprint: https://www.medrxiv.org/content/early/2021/06/13/2021.06.10.21258647.full.pdf. [87, 124]

**Gaythorpe, K, N Imai, G Cuomo-Dannenburg, M Baguelin, S Bhatia, A Boonyasiri, A Cori, Z Cucunuba Perez, A Dighe, I Dorigatti, R Fitzjohn, H Fu, W Green, J Griffin, A Hamlet, W Hinsley, N Hong, M Kwun, D Laydon, G Nedjati Gilani, L Okell, S Riley, H Thompson, S Van Elsland, R

Verity, E Volz, P Walker, H Wang, Y Wang, C Walters, C Whittaker, P Winskill, X Xi, C Donnelly, A Ghani, and N Ferguson. 2020. "Report 8: Symptom progression of COVID-19." [89]

Google, LLC. 2021. "Google COVID-19 Community Mobility Reports." Working paper. [98, 99, 116]

Harris, Ross J., Jennifer A. Hall, Asad Zaidi, Nick J. Andrews, J. Kevin Dunbar, and Gavin Dabrera. 2021. "Effect of Vaccination on Household Transmission of SARS-CoV-2 in England." *New England Journal of Medicine* 385 (8): 759–60. [89, 90]

He, Xi, Eric HY Lau, Peng Wu, Xilong Deng, Jian Wang, Xinxin Hao, Yiu Chung Lau, Jessica Y Wong, Yujuan Guan, Xinghua Tan, et al. 2020. "Temporal dynamics in viral shedding and transmissibility of COVID-19." *Nature medicine* 26 (5): 672–75. [88]

Hinch, Robert, William J M Probert, Anel Nurtay, Michelle Kendall, Chris Wymatt, Matthew Hall, Katrina Lythgoe, Ana Bulas Cruz, Lele Zhao, Andrea Stewart, Luca Ferritti, Daniel Montero, James Warren, Nicole Mather, Matthew Abueg, Neo Wu, Anthony Finkelstein, David G Bonsall, Lucie Abeler-Dorner, and Christophe Fraser. 21, 2021. *OpenABM-Covid19 - Parameter Dictionary*. URL: https://github.com/BDI-pathogens/OpenABM-Covid19/blob/14e24ce298ad523eba5e6f33ee9c70a90c527b3e/documentation/parameters/parameter_dictionary.md. [89]

Hunter, Paul R., and Julii Brainard. 2021. "Estimating the effectiveness of the Pfizer COVID-19 BNT162b2 vaccine after a single dose. A reanalysis of a study of real-world vaccination outcomes from Israel." *medRxiv*, (2): [90]

impfdashboard.de. 15, 2021. *Download der Datensätze*. URL: https://impfdashboard.de/daten. [89]

Isphording, Ingo E., Marc Lipfert, and Nico Pestel. 2021. "Does re-opening schools contribute to the spread of SARS-CoV-2? Evidence from staggered summer breaks in Germany." *Journal of Public Economics* 198 (6): 104426. [85]

Kretzschmar, Mirjam E, Ganna Rozhnova, Martin C J Bootsma, Michiel van Boven, Janneke H H M van de Wijgert, and Marc J M Bonten. 2020. "Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study." *Lancet Public Health* 5 (8): e452–e459. [125]

Kühn, Martin J., Daniel Abele, Tanmay Mitra, Wadim Koslow, Majid Abedi, Kathrin Rack, Martin Siggel, Sahamoddin Khailaie, Margrit Klitz, Sebastian Binder, Luca Spataro, Jonas Gilg, Jan Kleinert, Matthias Häberle, Lena Plötzke, Christoph D. Spinner, Melanie Stecher, Xiao Xiang Zhu, Achim Basermann, and Michael Meyer-Hermann. 2021. "Assessment of effective mitigation and prediction of the spread of SARS-CoV-2 in Germany using demographic information and spatial resolution." *Mathematical Biosciences* 339: 108648. [87]

Land Nordrhein-Westfalen. 16, 2021. *Corona-Teststrategie in Nordrhein-Westfalen: Bereits rund 4,4 Millionen Bürgertestungen*. URL: https://www.land.nrw/de/pressemitteilung/corona-teststrategie-nordrhein-westfalen-bereits-rund-44-millionen-buergertestungen. [126]

Levine-Tiefenbrun, Matan, Idan Yelin, Rachel Katz, Esma Herzel, Ziv Golan, Licita Schreiber, Tamar Wolf, Varda Nadler, Amir Ben-Tov, Jacob Kuint, Sivan Gazit, Tal Patalon, Gabriel Chodick, and Roy Kishony. 2021. "Initial report of decreased SARS-CoV-2 viral load after inoculation with the BNT162b2 vaccine." *Nature medicine* 27 (5): 790–92. [90]

Lipsitch, Marc, and Rebecca Kahn. 2021. "Interpreting vaccine efficacy trial results for infection and transmission." *Vaccine* 39 (30): 4082–88. [90]

Luijten, Michiel A. J., Maud M. van Muilekom, Lorynn Teela, Tinca J. C. Polderman, Caroline B. Terwee, Josjan Zijlmans, Leonie Klaufus, Arne Popma, Kim J. Oostrom, Hedy A. van Oers, and Lotte Haverman. 2021. "The impact of lockdown during the COVID-19 pandemic on mental and social health of children and adolescents." *Quality of Life Research*, (5): [133]

McAloon, Conor, Áine Collins, Kevin Hunt, Ann Barber, Andrew W Byrne, Francis Butler, Miriam Casey, John Griffin, Elizabeth Lane, David McEvoy, Patrick Wall, Martin Green, Luke O'Grady, and Simon J More. 2020. "Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research." *BMJ Open* 10 (8): e039652. [88]

McFadden, Daniel. 1989. "A method of simulated moments for estimation of discrete response models without numerical integration." *Econometrica: Journal of the Econometric Society*, 995–1026. [86, 106]

Melegari, M.G., M. Giallonardo, R. Sacco, L. Marcucci, S. Orecchio, and O. Bruni. 2021. "Identifying the impact of the confinement of Covid-19 on emotional-mood and behavioural dimensions in children and adolescents with attention deficit hyperactivity disorder (ADHD)." *Psychiatry Research* 296 (2): 113692. [133]

Mossong, Joël, Niel Hens, Mark Jit, Philippe Beutels, Kari Auranen, Rafael Mikolajczyk, Marco Massari, Stefania Salmaso, Gianpaolo Scalia Tomba, Jacco Wallinga, et al. 2008. "Social contacts and mixing patterns relevant to the spread of infectious diseases." *PLoS medicine* 5 (3): [91, 93, 95]

OECD. 2013. *Bildung auf einen Blick 2013: OECD-Indikatoren*, 530. [91]

Paul, Christian, Anna Steinberg, Christoph Schaller, Kathy Haemmerl, Leonie Thum, Marcel Parciak, Michael Knaf, Sarina Meyer, Tina Parciak, and Urs Grupp. 2020. *Dunkelzifferradar*. Edited by Prototype Fund. URL: https://prototypefund.de/project/dunkelzifferradar/. [103, 104]

Pavelka, Martin, Kevin Van-Zandvoort, Sam Abbott, Katharine Sherratt, Marek Majdan, Pavol Jarčuška, Marek Krajčí, Stefan Flasche, and Sebastian Funk. 2021. "The impact of population-wide rapid antigen testing on SARS-CoV-2 prevalence in Slovakia." *Science* 372 (6542): 635–41. [85]

Peak, Corey M, Rebecca Kahn, Yonatan H Grad, Lauren M Childs, Ruoran Li, Marc Lipsitch, and Caroline O Buckee. 2020. "Individual quarantine versus active monitoring of contacts for the mitigation of COVID-19: a modelling study." *Lancet Infectious Diseases* 20 (9): 1025–33. [88]

Petter, Ella, Orna Mor, Neta Zuckerman, Danit Oz-Levi, Asaf Younger, Dvir Aran, and Yaniv Erlich. 2021. "Initial real world evidence for lower viral load of individuals who have been vaccinated by BNT162b2." *medRxiv*, (2): [90]

Presse- und Informationsamt der Bundesregierung. 24, 2021. *Erste Zulassungen für Selbsttests*. URL: https://www.bundesregierung.de/breg-de/themen/coronavirus/zulassung-schnelltest-1861354 (visited on 06/24/2021). [101]

Pritchard, Emma, Philippa C. Matthews, Nicole Stoesser, David W. Eyre, Owen Gethings, Karina-Doris Vihta, Joel Jones, Thomas House, Harper VanSteenHouse, Iain Bell, John I. Bell, John N. Newton, Jeremy Farrar, Ian Diamond, Emma Rourke, Ruth Studley, Derrick Crook, Tim E. A. Peto, A. Sarah Walker, and Koen B. Pouwels. 2021. "Impact of vaccination on new SARS-CoV-2 infections in the United Kingdom." *Nature Medicine*, [89]

Rehrmann, Marc-Oliver. 13, 2021. *Drosten: "Schnelltests sind wohl weniger zuverlässig als gedacht"*. URL: https://www.ndr.de/nachrichten/info/Drosten-Schnelltests-sind-wohl-weniger-zuverlaessig-als-gedacht,coronavirusupdate178.html. [127]

Robert Koch Instititue. 7, 2021. *COVID-19-Fälle nach Meldewoche und Geschlecht sowie Anteile mit für COVID-19 relevanten Symptomen, Anteile Hospitalisierter/Verstorbener und Altersmittelwert/-median*. URL: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Daten/Klinische_Aspekte.xlsx. [104]

Robert Koch Institute. 6, 2020. *Testkriterien für die SARS-CoV-2 Diagnostik: Anpassungen für die Herbst- und Wintersaison 2020/2021.* URL: https://web.archive.org/web/20201103122248/https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Teststrategie/Testkiterien_Herbst_Winter.html. [103]

Robert Koch-Institut. 2020. *Laborbasierte Surveillance SARS-CoV-2.* URL: https://ars.rki.de/Content/COVID19/Main.aspx. [104]

Robert Koch-Institut. 2021a. "Corona-Monitoring bundesweit (RKI-SOEP-Studie)." de. Working paper. Robert Koch-Institut. [133]

Robert Koch-Institut. 29, 2021b. "COVID-19 Impfquoten-Monitoring in Deutschland (COVIMO): Report 5 - Fokuserhebung Impfquoten." Working paper. [92]

Robert Koch-Institut. 30, 2021c. *Todesfälle nach Sterbedatum.* URL: https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/COVID-19_Todesfaelle.html. [124]

Shapley, Lloyd S. 2016. *17. A value for n-person games.* Princeton University Press. [122]

Singanayagam, Anika, Monika Patel, Andre Charlett, Jamie Lopez Bernal, Vanessa Saliba, Joanna Ellis, Shamez Ladhani, Maria Zambon, and Robin Gopal. 2020. "Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020." *Eurosurveillance* 25 (32): [88, 89]

Stillman, Steven, and Mirco Tonin. 2021. "Communities and Testing for COVID-19." IZA Discussion Paper No. 14012. [133]

Stokes, Erin K., Laura D. Zambrano, Kayla N. Anderson, Ellyn P. Marder, Kala M. Raz, Suad El Burai Felix, Yunfeng Tie, and Kathleen E. Fullerton. 2020. "Coronavirus Disease 2019 Case Surveillance — United States, January 22–May 30, 2020." *MMWR. Morbidity and Mortality Weekly Report* 69 (24): 759–65. [89]

Süddeutsche Zeitung. 4, 2021. *80 Prozent der positiven Corona-Schnelltests falsch-positiv.* URL: https://www.sueddeutsche.de/gesundheit/gesundheit-hamburg-80-prozent-der-positiven-corona-schnelltests-falsch-positiv-dpa.urn-newsml-dpa-com-20090101-210704-99-252628. [130]

The Brussels Times. 22, 2020. *Rapid tests are not the solution to restart normal life, expert says.* URL: https://www.brusselstimes.com/news/belgium-all-news/146693/rapid-tests-are-not-the-solution-to-restart-normal-life-expert-says-herman-goossens-antigen-pcr/. [127]

Thompson, Mark G, Jefferey L Burgess, Allison L Naleway, Harmony L Tyner, Sarang K Yoon, Jennifer Meece, Lauren EW Olsho, Alberto J Caban-Martinez, Ashley Fowlkes, Karen Lutrick, et al. 2021. "Interim estimates of vaccine effectiveness of BNT162b2 and mRNA-1273 COVID-19 vaccines in preventing SARS-CoV-2 infection among health care personnel, first responders, and other essential and frontline workers—eight US locations, December 2020–March 2021." *Morbidity and Mortality Weekly Report* 70 (13): 495. [90]

Vlachos, Jonas, Edvin Hertegård, and Helena B. Svaleryd. 2021. "The effects of school closures on SARS-CoV-2 among parents and teachers." *Proceedings of the National Academy of Sciences* 118 (9): eprint: https://www.pnas.org/content/118/9/e2020834118.full.pdf. [85, 133]

Vygen-Bonnet, Sabine, Judith Koch, Christian Bogdan, Thomas Harder, Ulrich Heininger, Kerstin Kling, Martina Littmann, Joerg Meerpohl, Heidi Meyer, Thomas Mertens, Nora Schmid-Küpke, Stefan Scholz, Martin Terhardt, Marina Treskova-Schwarzbach, Klaus Überla, Marianne van der Sande, Ole Wichmann, Sabine Wicker, Ursula Wiedermann, Verina Wild, and Rüdiger von Kries. 2020. "Beschluss und Wissenschaftliche Begründung der Ständigen Impfkommission (STIKO) für die COVID-19-Impfempfehlung." (2): 3–63. [92]

Wollmershäuser, Timo. 2021. “Konjunkturelle Auswirkungen der zweiten Coronawelle auf ausgewählte Wirtschaftsbereiche.” *ifo Schnelldienst Digital* 2 (5): [133]

Yin, Guosheng, and Huaqing Jin. 2020. “Comparison of Transmissibility of Coronavirus Between Symptomatic and Asymptomatic Patients: Reanalysis of the Ningbo COVID-19 Data.” *JMIR Public Health and Surveillance* 6 (2): e19464. [88]

Zhao, Shi, Biao Tang, Salihu S Musa, Shujuan Ma, Jiayue Zhang, Minyan Zeng, Qingping Yun, Wei Guo, Yixiang Zheng, Zuyao Yang, Zhihang Peng, Marc KC Chong, Mohammad Javanbakht, Daihai He, and Maggie H. Wang. 2021. “Estimating the generation interval and inferring the latent period of COVID-19 from the contact tracing data.” *Epidemics* 36: 100482. [88]

Zou, Lirong, Feng Ruan, Mingxing Huang, Lijun Liang, Huitao Huang, Zhongsi Hong, Jianxiang Yu, Min Kang, Yingchao Song, Jinyu Xia, Qianfang Guo, Tie Song, Jianfeng He, Hui-Ling Yen, Malik Peiris, and Jie Wu. 2020. “SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients.” *New England Journal of Medicine* 382 (12): 1177–79. [88]

Appendix 3.A Germany's Policy Response to CoViD-19

Germany responded quite quickly to the first wave of CoViD-19 cases in March 2020 with a first lockdown that started around March 20th. It very effectively reduced case numbers and first measures were lifted at the end of April. The summer was characterized by very low numbers and these only started to increase again with the end of the summer vacations in August and September.

This did not lead to stricter rules which resulted in the second wave. In the seven weeks between mid September and early November, cases increased by a factor of ten. Restrictions were somewhat tightened in mid-October and again in early November (a so called "lockdown light"). New infections remained constant throughout November before rising again in December and Germany remained at incidences at or above 200 from November until January. The stark increase to nearly 300 cases per million inhabitants per day in December prompted the most stringent lockdown to this date. Schools and daycare centers were closed, so were customer-facing businesses except for grocery and drug stores. From the peak of the second wave just before Christmas until the trough in mid-February, newly detected cases decreased by almost three quarters. Around the turn of the year, the first people were vaccinated with a focus on older age groups and medical staff (Figures 3.3.8 and 3.2.2).

Despite warnings of the new Alpha variant, some NPIs were relaxed in early March. For example hairdressers and home improvement stores were allowed to open again to the public. The third wave from March to June 2021 is associated with the Alpha variant, which became dominant in March (Figure 3.3.7). There were many changes to details of regulations afterwards but they did not change the overall stringency index, as can be seen in Figure 3.A.1. Instead, Germany's epidemic management focused on two new containment strategies: rapid tests and vaccinations.

Until the end of May, Germany achieved that 43% of adults had received at least one dose of a vaccine. For individuals, who were not yet eligible to be vaccinated, widespread rapid testing was pursued to avoid stricter lockdown measures. While rapid tests had already replaced regular PCR tests for staff in many medical and nursing facilities in late 2020, those tests had to be administered by medical doctors or in pharmacies and were not widely available. In the beginning of March one test per person and week was made available free of charge (Bundesanzeiger, 2021a) and rapid test centers opened in many places, esp. in urban areas. This resulted in a massive rollout of rapid tests (Land Nordrhein-Westfalen, 2021a). In several states, customers were only allowed to enter certain stores with a recent negative rapid test result. At-home tests approved by authorities became available in mid-March. In mid April, rapid tests became mandatory in schools and employers were required to provide two rapid tests per week to their employees (Bundesanzeiger, 2021b). The

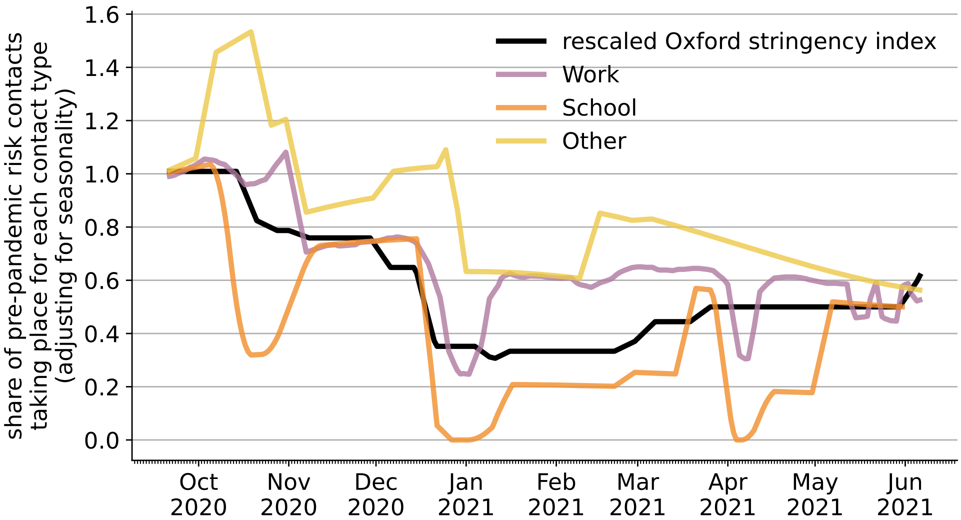


Figure 3.A.1. Stringency of NPIs and Infectious Contacts

Note: For legibility reasons, all lines are rolling 7-day averages. The Oxford Response Stringency Index (Hale et al., 2020) is scaled as $2 \cdot (1 - x/100)$, so that a value of one refers to the situation at the start of our sample period and zero means that all NPIs included in the index are turned on. The other lines show the product of the effect of contact reductions, increased hygiene regulations, and seasonality. The effects of various mechanisms can be disentangled due to the distinct temporal variation in the drivers of the pandemic. Next to the stringency index, the three lines summarize how contact reductions, increased hygiene regulations, and seasonality evolved since early September for each of the three broad contact networks. For example, a value of 0.75 for the work multiplier means that if the environment was the same as in September (levels of infection rates, no rapid tests or vaccinations, only the wildtype virus present), infections at the workplace would be reduced by 25%. See Section 3.2.5 for details.

combination of continued NPIs, massive rapid testing, increasing vaccination rates and favorable seasonality led cases to plummet starting at the end of April. Within a month the incidence fell from 250 to 50 cases per million inhabitants per day.

The Delta variant was first detected in Germany in April; at the end of our simulation period (end of May 2021) it accounted for less than 5% of cases still.

These developments are characteristic of many countries: The initial focus on NPIs to slow the spread of the disease has been accompanied by vaccines and a growing acceptance and use of rapid tests. At broadly similar points in time, novel strains of the virus have started to pose additional challenges.

Thus, Germany had very different combinations of NPIs and a lot of variation in its infection dynamic throughout our study period which is very helpful to identify both the infection probabilities of different contact types as well as the few policy parameters which we are unable to calibrate from surveys or other sources.

For a visualization of the policy stringency in our different contact networks, see Figure 3.A.1. The black line shows the Oxford Response Stringency Index (Hale et al., 2020), which tracks the tightness of non-pharmaceutical interventions. The index is shown for illustration of the NPIs, we never use it directly.[17] Next to the stringency index, the three lines summarize how contact reductions, increased hygiene regulations, and seasonality evolved since early September for each of the three broad contact networks.

Two aspects are particularly interesting: First, despite quite a bit of variation and even contrary movements, all lines broadly follow the stringency index and they would do so even more if we left out seasonality and school vacations (roughly the last two weeks of October, two weeks each around Christmas and Easter, and some days in late May). Second, the most stringent regulations coincide with the period of decreasing infection rates between late December 2020 and mid-February 2021. The subsequent reversal of the trend is associated with the spread of the Alpha variant. During the steep drop in recorded cases during May 2021, for 42% of the population took at least one rapid tests per week (Figure 3.3.9), the first-dose vaccination rate rose from 28% to 43% (Figure 3.3.8), and seasonality lowered the relative infectiousness of contacts (Figure 3.2.1).

[17]For legibility reasons, we transform the index so that lower values represent higher levels of restrictions. A value of zero means all measures incorporated in the index are turned on. The value one represents the situation in mid-September, with restrictions on gatherings and public events, face mask requirements, but open schools and workplaces.

Appendix 3.B Education Policies

3.B.1 School Policies

Until November schools were open normally. Starting in November, we assume that increased hygiene measures were taken. Schools stayed open until mid December. From mid December until January 10th schools closed and only offered so called "emergency care" for young children whose parents could credibly demonstrate that both had to work and had no other child care arrangement. Approximately 25% of primary school children and 5% of secondary students attended school as a result (Redaktionsnetzwerk Deutschland, 2021).

After January 10th when parents had returned to work the rules for emergency care were relaxed and approximately a third of primary school children and 10% of secondary students attended school as a result (Redaktionsnetzwerk Deutschland, 2021). In addition, graduating classes (most adolescents between 16 and 18) were allowed to return to school (Lehnardt, 2021; Stuttgarter Nachrichten, 2021) in a rotating scheme where each class was split in two groups. Relying on anecdotal evidence we assume that the groups rotate on a daily basis.

Starting on February 22nd primary school children were also allowed to return to school on a rotating basis until mid March. We summarize the school policy from mid March until Easter as all students being on a rotating school schedule. In addition, children that qualify for emergency care also attend on days where their group is scheduled to not attend school physically.

After the Easter break schools were mostly closed again. Part of this was a federal law, the so called "Bundesnotbremse" (Bundesgesetzblatt, 2021) that set rules for schools based on local incidences that were binding at the time. As a result, most states adjusted their schooling policies and during April most schools were closed with emergency care arrangements as in the time from January 10th to February 21st. As cases fell schools were allowed to gradually open. We summarize this as students being on the same rotating schedule as from mid March to Easter starting on May 1st (Bayerisches Staatsministerium für Unterricht und Kultus, 2021a,b; Landesregierung von Baden-Württemberg, 2021a,b,c; Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2021a,b,c,d).

3.B.2 Preschool and Nursery Policies

The policies for preschools and nurseries are similar to the school policies but simpler. Until November children attended completely normally, starting in November with increased hygiene measures. Nurseries and preschools stayed open until mid December. From mid December until January 10th, nurseries and preschools were nearly completely closed. If parents could credibly demonstrate that both parents work in systemically relevant professions and no other child care arrangement was possible, nurseries and preschools offered so called "emergency care". We assume 10% of children qualified and used emergency care during this time.

After January 10th when parents had returned to work the rules for emergency care were relaxed and we assume a third of children attended nursery and preschool. This policy stayed in place until February 20th. Afterwards, preschools and nurseries were open normally (maintaining increased hygiene measures) until mid March. Then during the third wave the restrictions of February were put back into place until end of April when nurseries and preschools opened again and stayed open for the rest of our simulation period – maintaining increased hygiene measures. (Bayerisches Staatsiministerium für Familie, Arbeit und Soziales, 2021a; Landesregierung von Baden-Württemberg, 2021a,d,e; Ministerium für Kinder, Familie, Flüchtlinge und Integration des Landes Nordrhein-Westfalen, 2021; Bayerisches Staatsiministerium für Familie, Arbeit und Soziales, 2021b).

Appendix 3.C Rapid Test Demand

3.C.1 Work Rapid Test Offers

Mid march, 20% of employers offered tests to their employees (Deutscher Industrie- und Handelskammertag, 2021). In the second half of March, 23% of employees reported being offered weekly rapid tests by their employer (Ahlers, Lübker, and Jung, 2021).[18] This share increased to 61% until the first days of April (Bonin, Krause-Pilatus, and Rinne, 2021a,d). Until mid April 72% of workers were expected to receive a weekly test offer (Bonin, Krause-Pilatus, and Rinne, 2021a,d). However, according to surveys conducted in mid April (Betsch et al., 2021), less than two thirds of individuals with work contacts did.[19] Starting on April 19th employers were required by law to provide two weekly tests to their employees (Bundesanzeiger, 2021b). However, compliance was very incomplete and only reached 80% (Ahlers and Lübker, 2021; Bellmann et al., 2021).

[18] Bonin, Krause-Pilatus, and Rinne (2021c) reports a somewhat higher share of 35% around March 20.

[19] Bonin, Krause-Pilatus, and Rinne (2021b) report 80% in mid April which is about 12% points less than what we arrive at.

3.C.2 Education Rapid Tests

We assume that employees in educational facilities start getting tested in 2021 and that by March 1st 30% of them ($\pi_{educator,\,t}$) are tested weekly ($\theta_{before\,Easter,\,educ} = 7$) (Land Baden-Württemberg, 2021; Land Bayern, 2021; Land Nordrhein-Westfalen, 2021b; Süddeutsche Zeitung, 2021; Zeit, 2021). The share increases to 90% for the week before Easter. At that time both Bavaria (Bayerisches Staatsministerium für Gesundheit und Pflege, 2021) and Baden-Württemberg (Ministerium für Kultus, Jugend und Sport Baden Württemberg, 2021) were offering tests to educators and North-Rhine Westphalia (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2021f) and Lower Saxony (Niedersächsisches Kultusministerium, 2021) were already testing school students and tests for school students and educators were already mandatory in Saxony (Sächsisches Staatsministerium für Kultus, 2021). After Easter we assume that 95% of educators get tested twice per week ($\theta_{after\,Easter,\,educ} = 3$).

Tests for school students started later (Ministerium für Kultus, Jugend und Sport Baden Württemberg, 2021; Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2021f) so we assume that they only start in February and only 10% of school students get tested by March 1st ($\pi_{students,\,t}$). Relying on the same sources as above we approximate that by the week before Easter this share had increased to 40% (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2021f). After Easter the share of school students receiving twice weekly tests is set to 75%. This is based on tests becoming mandatory in Bavaria (Bayerische Staatskanzlei, 2021) and North Rhine-Westphalia (Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, 2021e) after their Easter breaks and on April 19 in Baden-Württemberg (Ministerium für Kultus, Jugend und Sport Baden-Württemberg, 2021), after which we assume twice weekly rapid tests to be mandatory for all school students in Germany. Again, we interpolate linearly between these points and arrive at the purple line for educators and the red line for school students in Figure 3.2.10.

Appendix 3.D Shapley Values

We decompose the effects of different NPIs and seasonality on the infection rates with Shapley values. Shapley values (Shapley, 2016) are a concept in game theory to divide payoffs between a coalition of players. It allows to assign a single value to the contribution of an NPI or seasonality which takes into account substitutional and complementary effects with other factors.

More formally, define a coalitional game with N players and a super-additive function ν which maps subsets of N to the real numbers. The function ν is also called the characteristic function and assigns a value to a coalition. Then, the Shapley value ϕ for player i is

$$\phi_i(\nu) = \frac{1}{|N|!} \sum_{S \subseteq N \setminus \{i\}} |S|!(|N| - |S| - 1)!(\nu(S \cup \{i\}) - \nu(S))$$

The last term $(\nu(S \cup \{i\}) - \nu(S))$ is the marginal contribution of player i minus the coalition without player i. Then, compute the sum of marginal contributions over all subsets S of N which do not include player i. Each marginal contribution has to be multiplied by all combinations of other players in S which precede i and all possible combinations of remaining players which follow player i in the coalition. To arrive at the Shapley value for player i, divide the sum by the total number of combinations.

The Shapley value has some properties.

Efficiency The sum of Shapley values is equal to the value of a coalition formed by all players.

Symmetry The Shapley does not depend on the label of a player but only on its position in the characteristic function.

Linearity The Shapley value depends linearly on the values from the characteristic function ν.

Dummy Axiom The Shapley value of a player who contributes nothing to any coalition is 0.

To produce Figure 3.4.3 and Figure 3.4.4, we calculate the Shapley values of each factor in the comparison on the cumulative number of saved infections between the main scenario and the scenario without any of the factors for every day. Then, we divide up the saved infections on a particular day according to the Shapley values for the same day which yields the daily saved infections for each factor.

Appendix 3.E Model Robustness

Achieving a good in-sample fit does not necessarily guarantee that our model will also be able to make out of sample predictions. For example, it could be that the results are very sensitive to the exact number of vaccinations, the work mobility multiplier ($\rho_{w, attend, t}$) or the number of performed rapid tests (governed by the π parameters) – all of which are things that cannot be known exactly ex-ante.

In this section we compare simulated infections that use all available data with out of sample predictions that only use data that was available at March 1 2021.

For the out of sample predictions we predict the number of vaccinations between March and June with a simple linear regression model that was fitted on vaccine data from February. This prediction model is pessimistic compared to the actual number of vaccinations. The work mobility multiplier ($\rho_{w, attend, t}$) is predicted to be constant at a value of 0.75, which is an approximate average of the second half of February. This turned out to be optimistic.

The area that is fraught with the most uncertainty is the introduction of rapid tests, because it comprises both supply and demand factors. Moreover, accurately predicting the number of rapid tests is expected to be important because rapid tests play a large role for the transmission dynamic.

We therefore make a scenario analysis with different assumptions on the availability of rapid tests. The number of rapid tests performed in each scenario can be seen in Figure 3.E.1. All scenarios are the same until March 1 and have the same level of rapid tests when all supply constraints are resolved. They differ in the date at which the full number of tests is reached. For students ($\pi_{students, t}$) and teachers ($\pi_{teacher, t}$) the full number of rapid tests is reached after the Easter holidays in all scenarios. For rapid tests in the workplace ($\pi_{w, s, t}$) and private rapid tests ($\pi_{private, t}$) it is reached between May 1 and June 10, depending on the scenario.

Moreover, the out of sample predictions assume that the share of detected cases (ψ_t) that would have been obtained without rapid tests is not affected by the Easter holidays because the extent to which this was the case was estimated from case numbers in April.

The results of the out of sample prediction are displayed in Figure 3.E.2. While all scenarios considerably deviate from the ex-post scenario, they all reproduce the steep increase of cases until the end of April, followed by a decline until June. We can therefore conclude that our main results are not sensitive to measurement errors in the number of rapid tests, vaccinations or mobility data.

Appendix 3.F Comparison to Other Studies on Rapid Tests

Another form of validating our model is to see how well our main results align with other studies that evaluate the effect of large scale rapid testing. Of course, this has to be taken with a grain of salt as the effect of any rapid testing policy depends on the incidence of the disease in the population, how well other testing policies such as PCR tests are working, the effect of seasonality and NPIs that are in place. Nevertheless, it is reassuring that other studies find effect sizes in the same order of magnitude.

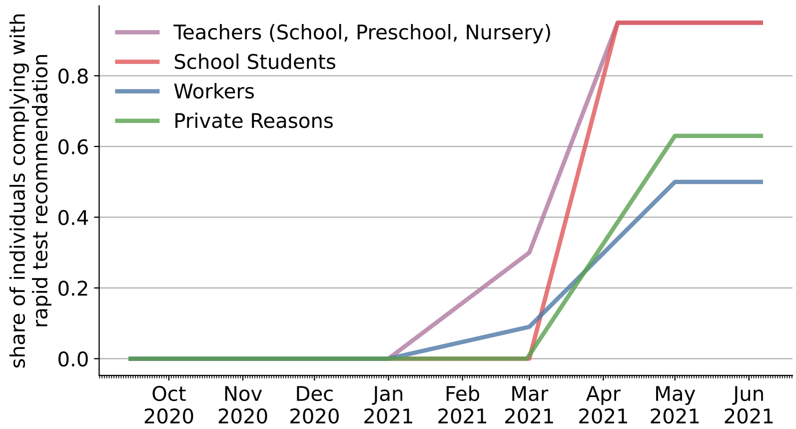
Pavelka et al. (2021) estimate that a mass testing campaign in Slovakia in October and November 2020 where approximately 65 % of the population took a rapid test within a two week period lead to a reduction in case numbers of 70 % three weeks after the start of the intervention. Moreover, they find that this strong reduction in cases cannot be explained by isolation of people who tested positive alone but only when they took into account that household members of people who tested positive reduced their contacts.

While we do not model the exact scenario of Pavelka et al. (2021), we can roughly compare their estimates with our predictions for the difference between the baseline scenario and and a scenario without rapid tests. In May about 45% of people do at least one rapid test in every week. Taking into account that there are many repeated testers the number of people who do a test within a two week period is probably slightly less than the 65% from the intervention in Slovakia. On the other hand, we have many people who do more than one rapid test in that time which also leads to the detection of cases. Our model predicts that the observed incidence with tests is approximately 65% lower than without tests after three weeks. Thus we have an effect size in the same order of magnitude but are slightly less optimistic regarding the efficiency of rapid tests.

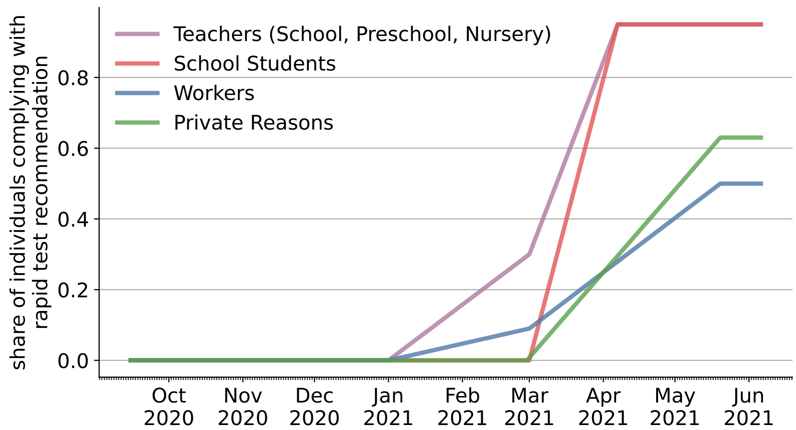
Berger, Fritz, and Kauermann (2021) analyse the effect of twice weekly rapid testing in schools. They have two main findings: Firstly, rapid tests reduced the share of undetected cases among students by a factor between two and four. Secondly, open schools with mandatory testing might lead to the same or even lower numbers of infections than closed schools. The estimates are based on infection numbers after the Easter holiday.

Again, we do not directly simulate their scenarios but can roughly compare our results to theirs. We estimate a share of undetected cases of approximately 75% among school age children (five to 14 years) at the beginning of April, see Figure 3.G.1. This drops to slightly less than 40% at the end of our simulation period. Thus in the long run, mandatory tests at schools led to a reduction of the share of undetected cases by a factor of more than 1.8 which is just slightly below the factor of two to four predicted by Berger, Fritz, and Kauermann (2021).

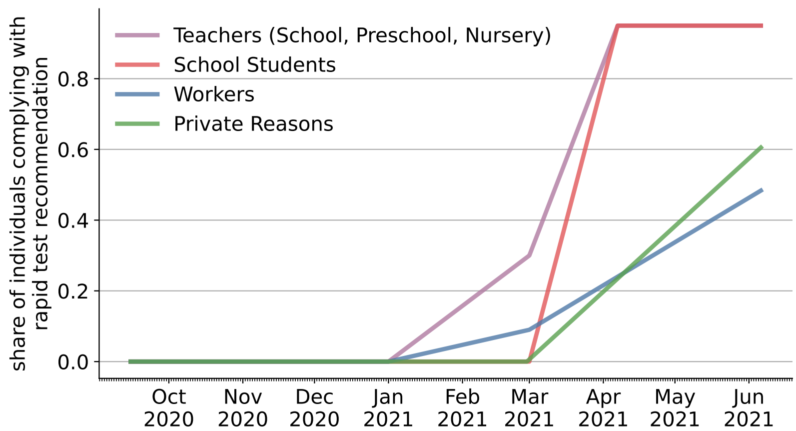
Similarly we are slightly less optimistic for the effect of opening schools with testing compared to closing schools. While they predict that opening schools could even be beneficial we estimate that it would lead to a slight increase in case numbers see Figure 3.4.10).



(a) Rapid Test Parameters: Early Scenario



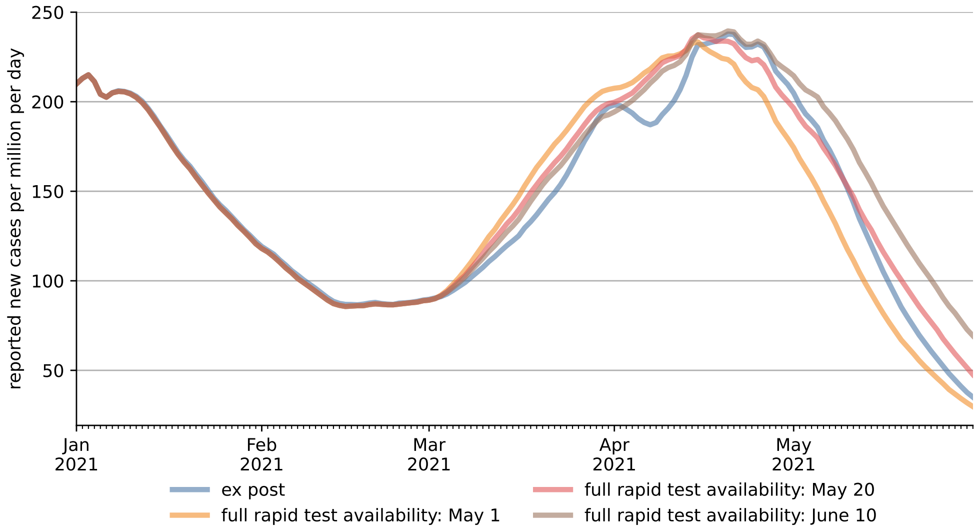
(b) Rapid Test Parameters: Medium Scenario



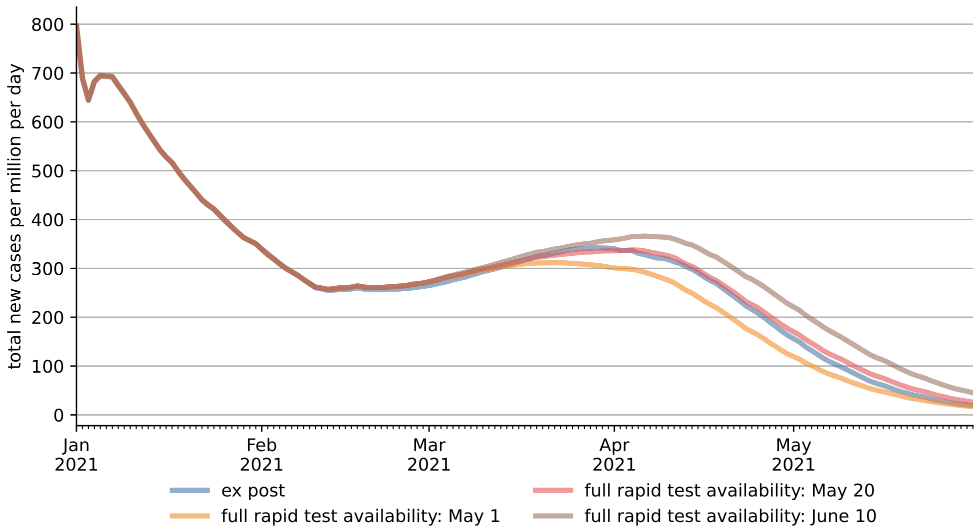
(c) Rapid Test Parameters: Late Scenario

Figure 3.E.1. Rapid Test Introduction in the Three Scenarios

Note: Number of rapid tests performed in the different prediction scenarios. All scenarios are the same until March 1 and have the same level of rapid tests when all supply constraints are resolved. They differ in the date at which the full number of tests is reached. For students ($\pi_{students,t}$) and teachers ($\pi_{teacher,t}$) the full number of rapid tests is reached after the Easter holidays in all scenarios. For rapid tests in the workplace ($\pi_{w,s,t}$) and private rapid tests ($\pi_{private,t}$) it is reached between May 1 and June 10, depending on the scenario.



(a) Reported Cases



(b) Total Cases

Figure 3.E.2. Out of Sample Prediction for Reported and Total Cases from March to June 2021

Note: The ex-post scenario is an in-sample prediction that uses all available information and is very close to actual case numbers. For the other scenarios data on vaccinations, work mobility and rapid tests that became available after March 1 have been replaced by prediction models that are calibrated with data from February. Moreover, they do not model a lower number of detected cases over the Easter holidays. The different scenarios make different assumptions on the date at which full availability of rapid tests is reached. While the out of sample predictions differ substantially for the exact case numbers at the beginning of June (between 20 and 70 cases per million), they can all reproduce the decline in case numbers that is jointly driven by seasonality, large scale rapid tests and vaccinations. For legibility reasons, all lines are rolling 7-day averages. Each line is the average over 30 simulation runs.

Appendix 3.G Share of Detected Cases

This section shows the share of detected cases for different age groups that endogenously result in our simulations. See Section 2.9 for an explanation of how we model the detection of cases and Section 3.2.7 for the calibration of the relevant parameters.

The share of detected cases falls by 15 percentage points for most age groups from October until Christmas when the number of CoViD-19 infections increased tenfold and PCR tests were a limited resource (Robert Koch Institute, 2020). As rapid tests become available in 2021 and more and more individuals receive positive rapid tests and seek PCR tests, the share of detected cases starts to increase. While first rapid tests become available in January of 2021 the effect only becomes substantial after March when access to rapid tests was greatly expanded.

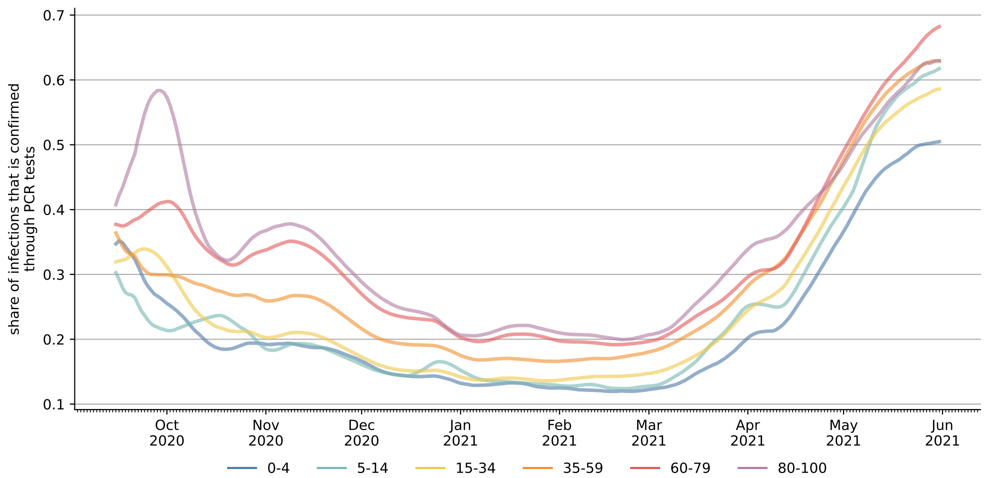


Figure 3.G.1. Share of Detected Cases by Age Group

Note: The figure shows the share of cases that is reported as an official case for each age group in our baseline simulations. For legibility reasons, all lines are rolling 7-day averages of the average of 30 simulation runs.

Overall, the share of detected cases is higher in older age groups (Figure 3.G.1). Our estimates suggest that – in the absence of rapid testing – the detection rate is 80% higher on average for individuals above age 80 compared to school age children. This is because the likelihood to develop symptoms increases with age and symptomatic cases are more likely to be detected. However, the role of rapid tests is here also very visible. The difference between age groups declines as rapid tests become more widely available – and are mostly taken up by age groups that still go to school or work. An especially strong case of this are school age children (5-14, green line), which overtake the next age group in May 2021. This comes from a particularly strong increase in their share of detected cases after Easter, when weekly rapid tests become mandatory in schools.

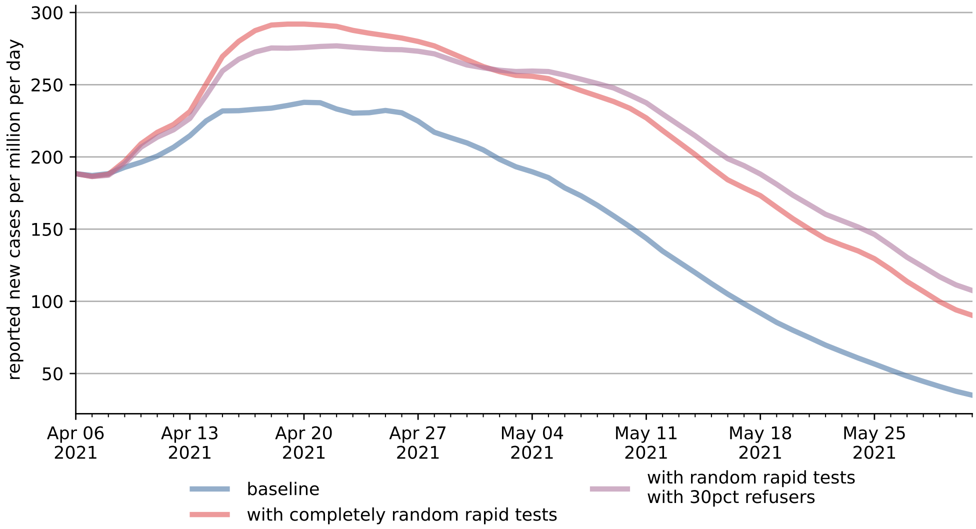
Appendix 3.H The Importance of the Rapid Test Policy

Here we shed some light on the role our rapid test demand channels play for the effect of rapid tests on case numbers. To do so we ran two scenarios where we allocated rapid tests either completely randomly in the entire population or among 70% of the population to account for the fact that a share of the population might refuse or be very hard to reach with rapid tests.[20]

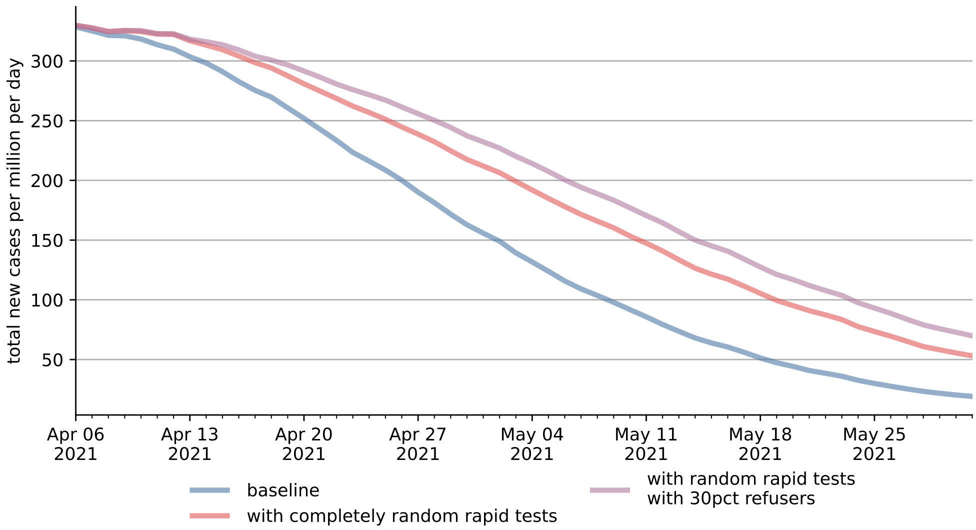
Figure 3.H.1 shows how the incidence of detected and total cases develops in the two random scenarios (red and purple line) relative to our baseline scenario (blue line). The differences between the two random scenarios are small. This is likely due to only a small fraction of the population being tested on any given day. Apart from this, two things stand out: Firstly, the total number of cases falls much faster in our baseline scenario compared to the two random scenarios. Secondly, this is not because the share of detected cases is higher in the baseline scenario; in fact, it is even slightly lower until end of April.

There are two mechanisms behind this: Firstly, tests at the workplace and schools predominantly target groups with many contacts. Thus, catching infections in these groups prevents more infections than in the general population. Secondly, rapid tests that are done because of private contact tracing are more effective at interrupting infection chains because they catch many infections in an early stage. Isolating infected individuals early on means that there are fewer days on which they can infect others.

[20] We calculate the number of rapid tests as in our baseline model. This leads to similar numbers of rapid tests. However given the higher incidence in our random scenarios these scenarios have a slightly higher number of rapid tests.



(a) Reported Cases



(b) Total Cases

Figure 3.H.1. The Role of Targeted and Compliance Driven Rapid Test Demand

Note: The figure shows the development of cases in two scenarios where rapid tests are distributed randomly in the population compared to our baseline scenario after Easter. In the baseline scenario rapid tests are targeted to workers, students, teachers and individuals at high risk of being infected including a weekly or twice weekly spacing between rapid tests. In the scenario with 30% refusers tests are randomly distributed among 70% of the population who are identified as compliers.

Appendix 3.I Reproducibility

The source code used for this paper is open source and available under the MIT License. It is split into two parts

- The source code for the model can be found at https://github.com/covid-19-impact-lab/sid/ and its documentation at https://sid-dev.readthedocs.io.
- The source code for the application to Germany can be found at https://github.com/covid-19-impact-lab/sid-germany/ with a shorter documentation at https://sid-germany.readthedocs.io.

We are grateful to the authors and contributors of the following software packages upon which our software is built: conda (Anaconda, 2016), conda-forge (conda-forge community, 2015) dask (Rocklin, 2015), estimagic (Gabler, 2020), holoviews (Stevens, Rudiger, and Bednar, 2015), matplotlib (Hunter, 2007), numba (Lam, Pitrou, and Seibert, 2015), numpy (Harris et al., 2020), pandas (McKinney, 2010; The pandas development team, 2020), pytask (Raabe, 2020), Python (Van Rossum and Drake Jr, 1995), scipy (Virtanen et al., 2020), and seaborn (Waskom, 2021).

References

Ahlers, Elke, and Malte Lübker. 2021. "Corona-Schnelltests am Arbeitsplatz: Einblicke aus der WSI-Lohnspiegel-Datenbank." WSI Policy Brief, No. 56. Type. Wirtschafts- und Sozialwissenschaftliches Institut, Hans-Böckler-Stiftung. [146]

Ahlers, Elke, Malte Lübker, and Rainer Jung. 2021. "Schleppender Start für wöchentliche Corona-Schnelltests am Arbeitsplatz: Nur 23 Prozent der Beschäftigten haben schon Zugang." Working paper. Wirtschafts- und Sozialwissenschaftliches Institut. [146]

Anaconda. 2016. "Anaconda Software Distribution." Version Vers. 2-2.4.0. [157]

Bayerische Staatskanzlei. 7, 2021. *Bericht aus der Kabinettssitzung vom 7. April 2021*. Pressemitteilung. URL: https://www.bayern.de/bericht-aus-der-kabinettssitzung-vom-7-april-2021/?seite=2453 (visited on 06/24/2021). [147]

Bayerisches Staatsiministerium für Familie, Arbeit und Soziales. 14, 2021a. *Kindertagesbetreuung: Aktuelle Informationen zum Coronavirus (SARS-CoV-2)*. URL: https://www.stmas.bayern.de/coronavirus-info/corona-kindertagesbetreuung.php (visited on 07/14/2021). [146]

Bayerisches Staatsiministerium für Familie, Arbeit und Soziales. 2021b. *404. Newsletter: Allgemeine Informationen zur Kindertagesbetreuung: Informationen zum Coronavirus (SARS-CoV-2)*. URL: https://www.stmas.bayern.de/imperia/md/content/stmas/stmas_inet/service-kinder/newsletter/404-newsletter.pdf (visited on 03/08/2021). [146]

Bayerisches Staatsministerium für Gesundheit und Pflege. 26, 2021. *Holetschek: Bayern hat frühzeitig Selbsttests gesichert - Schulen und Kitas bekommen schon nächste Woche 1,3 Millionen Selbsttests*. Pressemitteilung Nr. 44/GP. URL: https://www.stmgp.bayern.de/presse/holetschek-bayern-hat-fruehzeitig-selbsttests-gesichert-schulen-und-kitas-bekommen-schon/ (visited on 06/24/2021). [147]

Bayerisches Staatsministerium für Unterricht und Kultus. 30, 2021a. *FAQ zum Unterrichtsbetrieb an Bayerns Schulen*. URL: https://www.km.bayern.de/allgemein/meldung/7047/faq-zum-unterrichtsbetrieb-an-bayerns-schulen.html. Accessed on 2021-04-30. [145]

Bayerisches Staatsministerium für Unterricht und Kultus. 22, 2021b. *Informationen zum Unterrichtsbetrieb nach den Osterferien*. URL: https://www.km.bayern.de/lehrer/meldung/7244/informationen-zum-unterrichtsbetrieb-nach-den-osterferien.html. [145]

Bellmann, Lutz, Patrick Gleiser, Christian Kagerl, Eva Kleifgen, Corinna König, Ute Leber, Laura Pohlan, Duncan Roth, Malte Schierholz, Jens Stegmaier, and Armin Aminian. 18, 2021. *Das Angebot der Betriebe an Covid-19-Tests für die Beschäftigten hat deutlich zugenommen*. Insitut für Arbeitsmarkt- und Berufsforschung. URL: https://www.iab-forum.de/das-angebot-der-betriebe-an-covid-19-tests-fuer-die-beschaeftigten-hat-deutlich-zugenommen/ (visited on 07/15/2021). [146]

Berger, Ursula, Cornelius Fritz, and Göran Kauermann. 30, 2021. "CODAG Bericht Nr. 14: 2. Schulschließungen oder Schulöffnung mit Testpflicht? Epidemiologisch-statistische Aspekte sprechen für Schulöffnungen mit verpflichtenden Tests." Working paper. Ludwig-Maximilians-Universität München. [150]

Betsch, Cornelia, Lars Korn, Lisa Felgendreff, Sarah Eitze, Philipp Schmid, Philipp Sprengholz, Lothar Wieler, Patrick Schmich, Volker Stollorz, Michael Ramharter, Michael Bosnjak, Saad B. Omer, Heidrun Thaiss, Freia De Bock, and Ursula Von Rüden. 2021. "COVID-19 Snapshot Monitoring (COSMO Germany) - Wave 41." [146]

Bonin, Holger, Annabelle Krause-Pilatus, and Ulf Rinne. 2021a. "Arbeitssituation und Belastungsempfinden im Kontext der Corona-Pandemie im April 2021. Ergebnisse einer repräsentativen Befragung von abhängig Beschäftigten." Forschungsbericht 570/4. Institut zur Zukunft der Arbeit, Bundesministerium für Arbeit und Soziales. [146]

Bonin, Holger, Annabelle Krause-Pilatus, and Ulf Rinne. 2021b. "Arbeitssituation und Belastungsempfinden im Kontext der Corona-Pandemie im Mai 2021. Ergebnisse einer repräsentativen Befragung von abhängig Beschäftigten." Forschungsbericht 570/5. Institut zur Zukunft der Arbeit, Bundesministerium für Arbeit und Soziales. [146]

Bonin, Holger, Annabelle Krause-Pilatus, and Ulf Rinne. 2021c. "Arbeitssituation und Belastungsempfinden im Kontext der Corona-Pandemie im März 2021. Ergebnisse einer repräsentativen Befragung von abhängig Beschäftigten." Forschungsbericht 570/2. Institut zur Zukunft der Arbeit, Bundesministerium für Arbeit und Soziales. [146]

Bonin, Holger, Annabelle Krause-Pilatus, and Ulf Rinne. 15, 2021d. "Verbreitung von Homeoffice und Corona-Tests der Arbeitgeber Ende März-Anfang April 2021. Ergebnisse einer repräsentativen Befragung von abhängig Beschäftigten." Forschungsbericht 570/3. [146]

Bundesanzeiger. 8, 2021a. "Verordnung zum Anspruch auf Testung in Bezug auf einen direkten Erregernachweis des Coronavirus SARS-CoV-2 (Coronavirus-Testverordnung - TestV)." [142]

Bundesanzeiger. 2021b. "Zweite Verordnung zur Änderung der SARS-CoV-2-Arbeitsschutzverordnung." [142, 146]

Bundesgesetzblatt. 22, 2021. "Viertes Gesetz zum Schutz der Bevölkerung bei einer epidemischen Lage von nationaler Tragweite vom 22. April 2021." [145]

conda-forge community. 2015. "The conda-forge Project: Community-based Software Distribution Built on the conda Package Format and Ecosystem." [157]

Deutscher Industrie- und Handelskammertag. 2021. "Umfrage der IHK-Organisation zu Corona-Tests in den Unternehmen." Working paper. [146]

Gabler, Janos. 2020. "estimagic – A Python Tool for the Estimation of (Structural) Econometric Models." [157]

Hale, Thomas, Tilbe Atav, Laura Hallas, Beatriz Kira, Toby Phillips, Anna Petherick, and Annalena Pott. 2020. "Variation in US states responses to COVID-19." *Blavatnik School of Government*, [143, 144]

Harris, Charles R., K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. 2020. "Array programming with NumPy." *Nature* 585 (7825): 357–62. [157]

Hunter, John D. 2007. "Matplotlib: A 2D graphics environment." *Computing in science & engineering* 9 (03): 90–95. [157]

Lam, Siu Kwan, Antoine Pitrou, and Stanley Seibert. 2015. "Numba: A llvm-based python jit compiler." In *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, 1–6. [157]

Land Baden-Württemberg. 18, 2021. *Zwei anlasslose Schnelltests pro Woche für Schul- und Kitapersonal.* URL: https://www.baden-wuerttemberg.de/de/service/presse/pressemitteilung/pid/zwei-anlasslose-schnelltests-pro-woche-fuer-schul-und-kitapersonal/. [147]

Land Bayern. 26, 2021. *Bayerische Kitas und Schulen bekommen Corona-Selbsttests.* URL: https://www.br.de/nachrichten/bayern/kitas-und-schulen-in-bayern-bekommen-corona-selbsttests,SQ8F48k. [147]

Land Nordrhein-Westfalen. 16, 2021a. *Corona-Teststrategie in Nordrhein-Westfalen: Bereits rund 4,4 Millionen Bürgertestungen*. URL: https://www.land.nrw/de/pressemitteilung/corona-teststrategie-nordrhein-westfalen-bereits-rund-44-millionen-buergertestungen. [142]

Land Nordrhein-Westfalen. 15, 2021b. *Umfassende Informationen für die Schulen zu Corona-Selbsttests für Schülerinnen und Schüler*. URL: https://www.land.nrw/de/pressemitteilung/umfassende-informationen-fuer-die-schulen-zu-corona-selbsttests-fuer-schuelerinnen. [147]

Landesregierung von Baden-Württemberg. 23, 2021a. *Auswirkungen der Bundesnotbremse auf den Schul- und Kitabetrieb in Baden-Württemberg*. URL: https://km-bw.de/Lde/startseite/service/2021-04-23+Auswirkungen+der+Bundesnotbremse+auf+den+Schul-+und+Kitabetrieb+in+Baden-Wuerttemberg (visited on 04/23/2021). [145, 146]

Landesregierung von Baden-Württemberg. 25, 2021b. *Coronavirus: Aktuelle Informationen für Schulen und Kindertageseinrichtungen*. URL: https://km-bw.de/Coronavirus (visited on 03/25/2021). Accessed 2021-03-25. [145]

Landesregierung von Baden-Württemberg. 15, 2021c. *Präsenzbetrieb und inzidenzunabhängige Testpflicht ab dem 19. April*. URL: https://km-bw.de/,Lde/startseite/service/2021-04-15-schulbetrieb-ab-dem-19april (visited on 04/15/2021). [145]

Landesregierung von Baden-Württemberg. 14, 2021d. *Pressemitteilung: Landesweite Schließung von Schulen und Kindergärten*. URL: https://www.baden-wuerttemberg.de/de/service/presse/pressemitteilung/pid/landesweite-schliessung-von-schulen-und-kindergaerten/ (visited on 03/13/2021). [146]

Landesregierung von Baden-Württemberg. 14, 2021e. *Pressemitteilung: Schnelltests in Kitas und Kindergärten*. URL: https://www.baden-wuerttemberg.de/de/service/presse/pressemitteilung/pid/schnelltests-in-kitas-und-kindergaerten/ (visited on 04/08/2021). [146]

Lehnardt, Esther. 1, 2021. *Schulen öffnen für Abschlussklassen - Wegen Corona-Mutation nicht in ganz Bayern*. URL: https://www.swp.de/panorama/corona-schule-bayern-schulen-oeffnen-fuer-abschlussklassen-wegen-mutation-nicht-ueberall-kultusministerium-piazolo-54801404.html. [145]

McKinney, Wes. 2010. "Data Structures for Statistical Computing in Python." In *Proceedings of the 9th Python in Science Conference*. Edited by Stéfan van der Walt and Jarrod Millman, 56–61. [157]

Ministerium für Kinder, Familie, Flüchtlinge und Integration des Landes Nordrhein-Westfalen. 7, 2021. "Verordnung zum Schutz vor Neuinfizierungen mit dem Coronavirus SARS-CoV-2im Bereich der Betreuungsinfrastruktur (Coronabetreuungsverordnung – CoronaBetrVO)Vom 7. Januar 2021: In der ab dem 8. März 2021 gültigen Fassung." [146]

Ministerium für Kultus, Jugend und Sport Baden Württemberg. 2021. "Erweiterte Teststrategie - Selbsttests an Schulen." *Pressemitteilungen*, (4): [147]

Ministerium für Kultus, Jugend und Sport Baden-Württemberg. 2, 2021. *Informationen zum Schulbetrieb nach den Osterferien*. Pressemitteilung. URL: https://km-bw.de/,Lde/9112691 (visited on 06/24/2021). [147]

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. 23, 2021a. *„Notbremse" im Infektionsschutzgesetz des Bundes*. URL: https://schulministerium.nrw/presse/pressemitteilungen/notbremse-im-infektionsschutzgesetz-des-bundes-23-04-2021 (visited on 04/23/2021). [145]

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. 8, 2021b. *[08.04.2021] Informationen zum Schulbetrieb in NRW*. URL: https://schulministerium.nrw/08042021-informationen-zum-schulbetrieb-nrw (visited on 04/08/2021). [145]

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. 22, 2021c. *[22.04.2021] Informationen zum Schulbetrieb ab 26. April 2021*. URL: https://www.schulministerium.nrw/22042021-informationen-zum-schulbetrieb-ab-26-april-2021 (visited on 04/22/2021). [145]

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. 14, 2021d. *Pressemitteilung: Ministerin Gebauer: Bleiben auf dem Weg der Vorsicht: Wechselunterricht mit klaren Testvorgaben ab Montag*. URL: https://schulministerium.nrw/presse/pressemitteilungen/ministerin-gebauer-bleiben-auf-dem-weg-der-vorsicht-14-04-2021 (visited on 04/14/2021). [145]

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. 14, 2021e. *Schulbetrieb im Wechselunterricht ab Montag, 19.April 2021 // Coronaselbsttests an Schulen - Testpflicht*. URL: https://www.schulministerium.nrw/14042021-schulbetrieb-im-wechselunterricht-ab-montag-19-april-2021-coronaselbsttests-schulen (visited on 06/27/2021). [147]

Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen. 15, 2021f. *Umfassende Informationen für die Schulen zu Corona-Selbsttests für Schülerinnen und Schüler*. Pressemitteilung. URL: https://www.land.nrw/de/pressemitteilung/umfassende-informationen-fuer-die-schulen-zu-corona-selbsttests-fuer-schuelerinnen (visited on 06/27/2021). [147]

Niedersächsisches Kultusministerium. 18, 2021. *Selbsttests an Schulen laufen an - Auslieferung der ersten 400.000 Tests beginnt heute, Testwoche startet Montag*. Pressemitteilung. URL: https://www.mk.niedersachsen.de/startseite/aktuelles/presseinformationen/selbsttests-an-schulen-laufen-an-erste-400-000-tests-ab-heute-in-der-auslieferung-testwoche-startet-montag-198555.html (visited on 06/24/2021). [147]

Pavelka, Martin, Kevin Van-Zandvoort, Sam Abbott, Katharine Sherratt, Marek Majdan, Pavol Jarčuška, Marek Krajčí, Stefan Flasche, and Sebastian Funk. 2021. "The impact of population-wide rapid antigen testing on SARS-CoV-2 prevalence in Slovakia." *Science* 372 (6542): 635–41. [150]

Raabe, Tobias. 2020. "pytask – A Python tool for managing scientific workflows." [157]

Redaktionsnetzwerk Deutschland. 27, 2021. *Steigende Nutzung der Notbetreuung: Etwa jedes dritte Kita-Kind wird abgegeben*. URL: https://www.rnd.de/politik/notbetreuung-im-lockdown-etwa-jedes-dritte-kita-kind-betroffen-die-bundeslander-ubersicht-3ETNUHXENVFTLN4UGENFV7IPU4.html. [145]

Robert Koch Institute. 6, 2020. *Testkriterien für die SARS-CoV-2 Diagnostik: Anpassungen für die Herbst- und Wintersaison 2020/2021*. URL: https://web.archive.org/web/20201103122248/https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Teststrategie/Testkiterien_Herbst_Winter.html. [154]

Rocklin, Matthew. 2015. "Dask: Parallel computation with blocked algorithms and task scheduling." In *Proceedings of the 14th python in science conference*. Vol. 130, Citeseer, 136. [157]

Sächsisches Staatsministerium für Kultus. 16, 2021. *Selbsttests für Schüler und Lehrer ab Mittwoch*. URL: https://www.medienservice.sachsen.de/medien/news/248791 (visited on 06/24/2021). [147]

Shapley, Lloyd S. 2016. *17. A value for n-person games*. Princeton University Press. [147]

Stevens, Jean-Luc R, Philipp Rudiger, and J Bednar. 2015. "HoloViews: Building complex visualizations easily for reproducible science." In *Proceedings of the 14th Python in Science Conference*. Citeseer, 61–69. [157]

Stuttgarter Nachrichten. 6, 2021. *Präsenzunterricht für Abschlussklassen bereits ab 18. Januar geplant*. URL: https://www.stuttgarter-nachrichten.de/inhalt.coronavirus-in-baden-wuerttemberg-praesenzunterricht-fuer-abschlussklassen-bereits-ab-18-januar-geplant.d47dd03c-a6b6-458b-bb41-77bf6641325f.html. [145]

Süddeutsche Zeitung. 19, 2021. *Wie soll das ein Kind schaffen, dem man noch die Schuhe binden muss?* URL: https://www.sueddeutsche.de/muenchen/muenchen-schule-selbsttest-lehrer-protest-1.5240397. [147]

The pandas development team. 2020. *pandas-dev/pandas: Pandas*. Version latest. [157]

Van Rossum, Guido, and Fred L Drake Jr. 1995. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam. [157]

Virtanen, Pauli, Ralf Gommers, Travis E Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, et al. 2020. “SciPy 1.0: fundamental algorithms for scientific computing in Python.” *Nature methods* 17 (3): 261–72. [157]

Waskom, Michael L. 2021. “Seaborn: statistical data visualization.” *Journal of Open Source Software* 6 (60): 3021. [157]

Zeit. 4, 2021. *Teste sich, wer kann*. URL: https://www.zeit.de/politik/deutschland/2021-03/corona-schnelltest-selbsttest-anwendung-kosten-discounter-verkaufsstart-faq?utm_referrer=https%3A%2F%2Fwww.google.com%2F#sind-genug-tests-da. [147]