# The Effectiveness of Strategies to Contain SARS-CoV-2: Testing, Vaccinations, and NPIs [*]

Janoś Gabler[a, b]
Tobias Raabe[c]
Klara Röhrl[a]
Hans-Martin von Gaudecker[b,d]

[a] *Bonn Graduate School of Economics*
[b] *IZA Institute of Labor Economics*
[c] *Private sector*
[d] *Rheinische Friedrich-Wilhelms-Universität Bonn*

July 14, 2021

In order to slow the spread of the CoViD-19 pandemic, governments around the world have enacted a wide set of policies limiting the transmission of the disease. Initially, these focused on non-pharmaceutical interventions; more recently, vaccinations and large-scale rapid testing have started to play a major role. The objective of this study is to explain the quantitative effects of these policies on determining the course of the pandemic, allowing for factors like seasonality or virus strains with different transmission profiles. To do so, the study develops an agent-based simulation model, which is estimated using data for the second and the third wave of the CoViD-19 pandemic in Germany. The paper finds that during a period where vaccination rates rose from 5% to 40%, seasonality and rapid testing had the largest effect on reducing infection numbers. Frequent large-scale rapid testing should remain part of strategies to contain CoViD-19; it can substitute for many non-pharmaceutical interventions that come at a much larger cost to individuals, society, and the economy.

**[Tobias 1]**

**Commenting is on!**
To switch it off, activate
`\PassOptionsToPackage`
`{final}{changes}`
in preamble.tex.

1

Since early 2020, the CoViD-19 pandemic has presented an enormous challenge to humanity on many dimensions. The development of highly effective vaccines holds the promise of containment in the medium term. However, most countries find themselves many months—and often years—away from reaching vaccination levels that would end the pandemic or even protect the most vulnerable (*17*). In the meantime, it is of utmost importance to employ an effective mix of strategies for containing the virus. The most frequent initial response was a set 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. Containment measures have become increasingly diverse and now include rapid testing, more nuanced NPIs, and contact tracing. Neither these policies' effect nor the influence of seasonal patterns or of more infectious virus strains are well understood in quantitative terms.

This paper develops 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. We apply the model to Germany, where new infections fell by almost 80% during the May 2021. Our analysis shows that, aside from seasonality, frequent and large-scale rapid testing caused the bulk of this decrease, which is in line with prior predictions (*20*). We conclude that it should have a large role for at least as long as vaccinations have not been offered to an entire population.

At the core of our agent-based model (*1*, *9*, we review more literature in Supplementary Material B.1) are physical contacts between heterogeneous agents (Figure 1a). Each contact between an infectious individual and somebody susceptible to the disease bears the risk of transmitting the virus. Contacts occur in up to four networks: Within the household, at work, at school, or in other settings (leisure activities, grocery shopping, medical appointments, etc.). Some contacts recur regularly, others 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. *21*, *10*). Within each network, meeting frequencies depend on age and geographical location (see Supplementary Material A.4).

The four contact networks are chosen so that the most common NPIs can be modeled in great 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 masking requirements

**[Klara 1]**
infection probability per contact for contact type $c$: $\beta_c$

**[Klara 2]**
probability to have $n$ contacts of contact type $c$: $\eta_{c,n}$

**[Klara 3]**
Share of contacts of age group and county of i to meet someone of j's age group and county: $\alpha_{work, age_i, age_j, county_i, county_j}$

**[Klara 4]**
$\rho_{attend,t}$

---

1. See Contreras *et al.* (*4*) for a theoretical equilibrium at low case numbers which is sustained with test-trace-and-isolate policies.

**(a)** Model for contacts and infections



**(b)** Disease progression



**(c)** Model for rapid tests



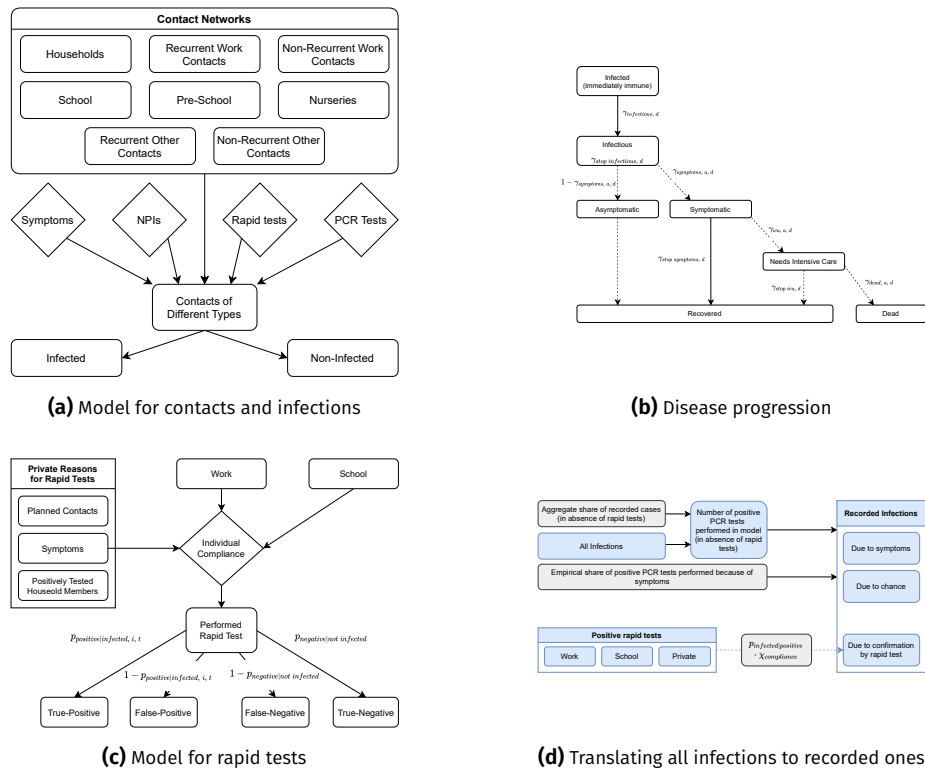**(d)** Translating all infections to recorded ones

**Figure 1.** Model description

*Note:* A description of the model can be found in Supplementary Material B. Figure 1a shows the influence on an agent's contacts to other agents. Demographic characteristics set the baseline number of contacts in different networks ($\eta$). She may reduce the number of contacts due to NPIs, showing symptoms, or testing positively for SARS-CoV-2 ($\tau$). Infections may occur when a susceptible agent meets an infectious agent; the probability depends on the type of contact ($\beta_c$), on seasonality ($\kappa_c$), and on NPIs ($\rho_{c,t}$). If infected, the infection progresses as depicted in Figure 1b. If rapid tests are available, agents' demand is modeled as in Figure 1c. 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 $\tau_{c,t}$). Figure 1d shows the model of translating all infections in the simulated data to age-specific recorded infections. The model uses data on the aggregate share of recorded cases ($\psi$), on the share of positive PCR tests triggered by symptoms ($\chi_{symptom}$), and on 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.

or air filters (*14*). Curfews may reduce the number of contacts in settings outside of work and school. In any setting, measures like masking requirements would reduce the probability of infection associated with a contact (*3*).

In the model, susceptibility to contracting the SARS-CoV-2 virus is dependent on age. A possible infection progresses as shown in Figure 1b. 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 *7*). The probabilities of transitioning between these states depend on age; their duration is random and calibrated

**[Klara 5]**
would enter into $\rho_{other,t}$

**[Klara 6]**
$\rho_{hygiene}$

**[Klara 7]**
$\zeta_a$

**[HM 1]**
New version of figure has lost the timing. Need to explain d-subscript, maybe even better to add stuff like 1-3 days back in? Needs to be comprehensible without recurring to appendix

**[HM 2]**
Would the model in Aleta *et al.* ( ) work?

2

to medical literature (for a detailed description see Supplementary Material A.1). Conditional on the type of contact, infectiousness is independent of age (*12*).

The model includes several other 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 vaccine roll-out, priority may depend on age and occupation; vaccine hesitancy is modelled by some individuals refusing vaccination offers. With some probability, vaccinated agents become immune and do not transmit the virus (*11*, *15*, *22*, *23*).

We include two types of tests. Polymerase chain reaction (PCR) tests 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 (*2*, *26*); sensitivity depends on the timing of the test relative to the onset of infectiousness. After a phase-in period, all tests that are demanded will be performed. Figure 1c shows our model for rapid test demand. Schools may require staff and students to be tested regularly. Rapid tests may be offered by employers to on-site workers. Individuals may 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 whether she takes up rapid tests.[2]

Modelling a population of agents according to actual demographic characteristics means that we can use a wide array of data to identify and estimate the model's many parameters.[3] 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. Other studies' estimates of the seasonality of infections can be incorporated directly. The remaining parameters—most notably, these include infection probabilities by contact network and the effects of some NPIs, see Supplementary Material B.10—will be chosen numerically so that the model matches features of the data (see *18*, 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 B.1.1.7 strain.

---

[2]. Positive test results or symptoms leads most individuals to reduce their contacts; this is why tests impact the actual contacts in Figure 1.

[3]. See Supplementary Material A for a complete description.

3

**[Klara 8]**
would be $\beta_c$

**[Klara 9]**
$\sigma_v$

**[Klara 10]**
$\gamma_{PCR, d}$ where $d$ = number of days

**[Klara 11]**
$p_{negative|not\ infected}$

**[Klara 12]**
"$p_{positive|infected, i, t}$", where $i$ is the individual and $t$ is the date to determine time until/since infectiousness

**[Klara 13]**
"$\pi_{teacher, t}$" etc.

The main issue with official case numbers is that they will contain only a fraction of all infections. In the German case, this specifically amounts to positive PCR tests. We thus model recorded cases as depicted in Figure 1d. 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 shares (see Figure C.9 for the share of detected cases by age group over time in our model). 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. Once rapid test become available, confirmation of a positive result is another reason leading to positive PCR tests.

The model is applied to the second and third waves of the CoViD-19 pandemic in Germany, covering the period mid-September 2020 to the end of May 2021. Figure 2 describes the evolution of the pandemic and of its drivers. The black line in Figure 2a shows officially recorded cases; the black line in Figure 2b the Oxford Response Stringency Index (8), which tracks the tightness of non-pharmaceutical interventions. The index is shown for illustration of the NPIs, we never use it directly. 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, masking requirements, but open schools and workplaces. 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. New infections remained constant throughout November before rising again in December, prompting 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. The third wave in the spring of 2021 is associated with the B.1.1.7 (Alpha) strain, which became dominant in March (Figure 2c).[4] In early March, some NPIs were relaxed; e.g., hairdressers and home improvement stores were allowed to open again to the public. There were many changes in details of regulations afterwards, but they did not change the overall stringency index.

By March 2021, the set of policy instruments had become much more diverse. Around the turn of the year, the first people were vaccinated with a focus on older age groups and medical staff (Figure 2d). Until the end of May, 43% had received at least one dose of a vaccine. In late 2020, rapid tests started to replace regular PCR tests for staff in many medical and nursing facilities. These had to be administered by medical doctors or in pharmacies. At-home tests approved by authorities became

---

4. B.1.617.2 (Delta) reached Germany in April but still accounted for less than 5% of cases at the end of our simulation period.
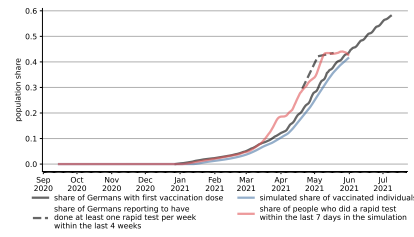
[HM 3]
added "mortality-based" – correct?

[Klara 14]
$\psi$

[Klara 15]
$\chi_{symptom}$

[Klara 16]
prob of demanding confirmation is $\chi_{confirmation}$

**(a)** Recorded cases: Empirical and simulated
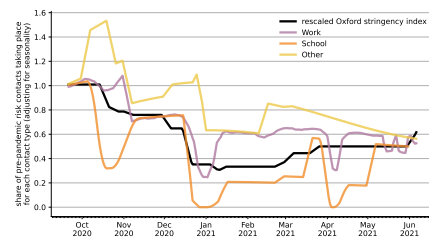
**(b)** Stringency of NPIs and infectious contacts

**(c)** Fraction of B.1.1.7 strain

**(d)** Tests and vaccinations

**Figure 2.** Evolution of the pandemic, its drivers, and model fit, September 2020 to May 2021

*Note:* Data sources are described in Supplementary Material A. Age- and region-specific analogues to Figure 2a can be found in Supplementary Material C.1. For legibility reasons, all lines in Figure 2b are rolling 7-day averages. The Oxford Response Stringency Index 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 in Figure 2b show the product of the effect of contact reductions, increased hygiene regulations, and seasonality. See Appendix A.5 for separate plots of the three factors by contact type.

available in mid-March. Rapid test centers were opened, and one test per person and week was made available free of charge. In several states, customers were only allowed to enter certain stores with a recent negative rapid test result. 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.

We draw simulated samples of agents from the population structure in September 2020 and use the model to predict recorded infection rates until the end of May 2021. See Supplementary Materials A.2 and B.9 for details. The blue line in Figure 2a shows that our model's predictions are very close to officially recorded cases in the aggregate. This is also true for infections by age and geographical region (see Supplementary Material C.1).

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 in Figure 2b 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%. Two aspects are particularly interesting. First, 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 B.1.1.7 variant. During the steep drop in recorded cases during May 2021, for 42% of the population took at least one rapid tests per week, the first-dose vaccination rate rose from 28% to 43%, and seasonality lowered the relative infectiousness of contacts.

In order to better understand the contributions of rapid tests, vaccinations, and seasonality on the evolution of infections in 2021, Figure 3 considers various scenarios. NPIs are always held constant at their values in the baseline scenario. Figure 3a shows the model fit (the blue line, same as in Figure 2a), a scenario without any of the three factors (red line), and three scenarios turning each of these factors on individually. Figure 3b does the same for total infections in the model. Figure 3c employs Shapley values (25) to decompose the difference in total infections between the scenario without any of the three factors and our main specification.

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 chan-

**(a)** Recorded cases: 2021 scenarios



**(b)** Total cases: 2021 scenarios
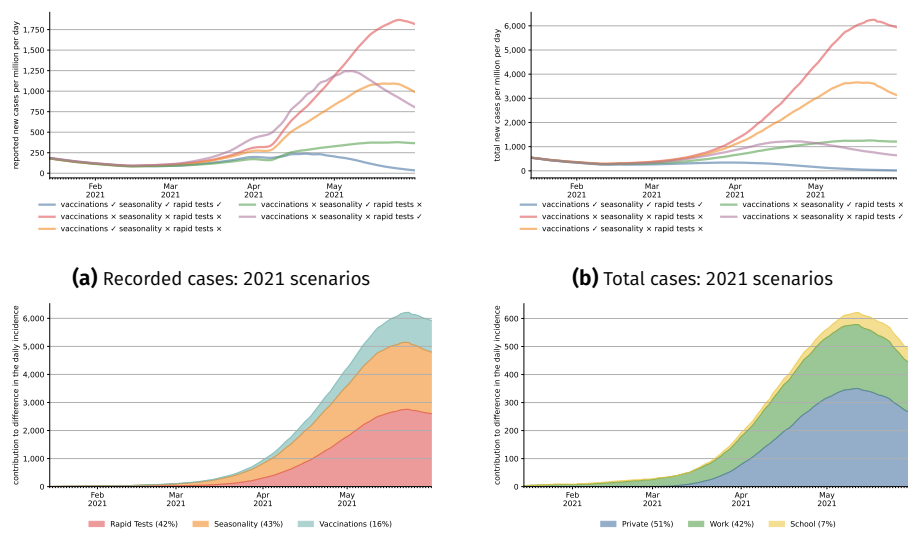


**(c)** Decomposition of the difference between the scenario without any of the three factors and the main scenario in Figure 3b.



**(d)** Decomposition of the difference between the scenario without rapid tests and the main scenario in Figure 3b.

**Figure 3.** The effect of different interventions on recorded and actual infections

*Note:* The blue line in Figure 3a is the same as in Figure 2a and refers to our baseline scenario, so does the blue line in Figure 3b. The red lines refer to a situation where NPIs evolve as in the baseline scenario and the B.1.1.7 variant is introduced in the same way; vaccinations, rapid tests, and seasonality remain at their January levels. The other scenarios turn each of these three factors on individually. The decompositions in Figures 3c and 3d are based on Shapley values, which are explained more thoroughly in Appendix B.11

nel active, the final level is always the highest in case of the vaccination campaign (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 numbers 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 was rate 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).

Seasonality has a large effect in slowing the spread of SARS-CoV-2. By May 31, both observed and recorded cases would be reduced by a factor of four if only seasonality mattered. However, in this period, cases would have kept on rising throughout, just at a much lower pace (this is in line with results in 6, 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. Here, it is crucial to differentiate between recorded cases and actual cases. Additional testing means that additional infections will be recorded which would otherwise remain undetected. Figure 3a 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 3b. Despite the fact that only 10% of the population performed weekly rapid tests in March on average, new infections on April 1 would have been reduced by 53% 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 3d decomposes the difference in the scenario without rapid tests only and the main specification into the three channels for rapid tests. Tests at schools have the smallest effect, which is largely explained by schools not operating at full capacity during our period of study and the relatively small number of students.[5] 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

[Klara 17]
This needs a source.

[HM 4]
Didn't we calculate that ourselves at some point?

---

5. 18% of our population are in the education sector (pupils, teachers, etc.); 46% are workers outside the education sector.

**(a)** Effects of different schooling scenarios


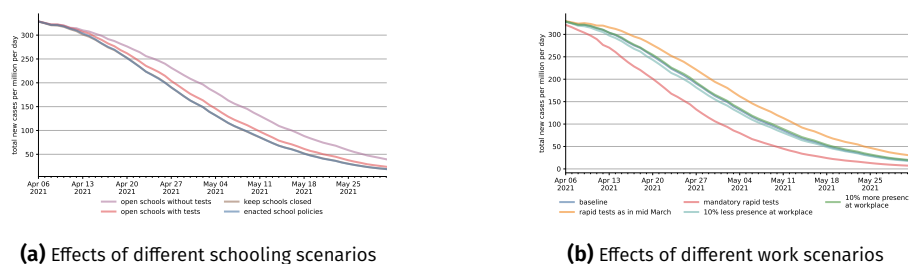
**(b)** Effects of different work scenarios

**Figure 4.** Effects of different scenarios for policies regarding schools and workplaces.

*Note:* Blue lines in both figures refer to our baseline scenario; they are the same as in Figure 3b. Interventions start at Easter because there were no capacity constraints for rapid tests afterwards.

form of contact tracing, which has been shown to be very effective (*4*, *5*, *13*). Indeed, a deeper analysis in Supplementary Material C.5 shows that the same amount of rapid tests administered randomly in the population would not nearly have been as effective.

Two of the most contentious NPIs concern schools and mandates to work from home. In many countries, schools switched to remote instruction during the first wave, so did Germany. After the summer break, they were operating at full capacity with increased hygiene measures, before being closed again from mid-December onward. Some states started opening them gradually in late February, but operation at normal capacity did not resume until the beginning of June. Figure 4a shows the effects of different policies regarding schools starting at Easter, at which point rapid tests had become widely available. 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 (e.g., *28*). To use another metric, the effective weekly reproduction number differs by 0.018 and 0.052, respectively. In light of the large negative effects school closures have on children and parents (*16*, *19*)—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. In other situations, and particular when rapid test are not available at scale, trade-offs may well be different.

Figure 4b 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—would have reduced infections by 23%. Reducing rapid

**[Klara 18]**
We should decide if we want to keep this section or prefer to show some rapid test statistics.

**[HM 5]**
Need a back-of the envelope calculation

9

tests offers by employers to the level of March would have increased infections by 13%.

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 Euros per capita (*29*); retail prices for rapid tests were below one Euro in early June 2021. 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. *27*); at the same time, these groups have a higher risk to contract CoViD-19 (*24*). 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 programmes allow a much quicker roll-out, making it arguably the most effective tool to contain the pandemic in the short run.

## References

1. A. Aleta, D. Martín-Corral, A. P. y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. L. Jr, S. Merler, A. Pentland, A. Vespignani, E. Moro, Y. Moreno, Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* **4**, 964–971 (2020).

2. L. E. Brümmer, S. Katzenschlager, M. Gaeddert, C. Erdmann, S. Schmitz, M. Bota, M. Grilli, J. Larmann, M. A. Weigand, N. R. Pollock, S. Carmona, S. Ongarello, J. Sacks, C. M. Denkinger, The accuracy of novel antigen rapid diagnostics for SARS-CoV-2: a living systematic review and meta-analysis. (2021).

3. Y. Cheng, N. Ma, C. Witt, S. Rapp, P. S. Wild, M. O. Andreae, U. Pöschl, H. Su, Face masks effectively limit the probability of SARS-CoV-2 transmission. *Science*, ISSN: 0036-8075, eprint: https://science.sciencemag.org/content/early/2021/05/19/science.abg6296.full.pdf, (https://science.sciencemag.org/content/early/2021/05/19/science.abg6296) (2021).

4. S. Contreras, J. Dehning, S. B. Mohr, S. Bauer, F. P. Spitzner, V. Priesemann, Low case numbers enable long-term stable pandemic control without lockdowns. *arXiv*, arXiv: 2011.11413 (q-bio.PE) (2021).

5. T. Fetzer, T. Graeber, Measuring the Scientific Effectiveness of Contact Tracing: Evidence from a Natural Experiment. *Forthcoming, Proceedings of the National Academy of Sciences* (2021).

6. T. Gavenčiak, J. T. Monrad, G. Leech, M. Sharma, S. Mindermann, J. M. Brauner, S. Bhatt, J. Kulveit, 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, (https://www.medrxiv.org/content/early/2021/06/13/2021.06.10.21258647) (2021).

7. V. Grimm, F. Mengel, M. Schmidt, Extensions of the SEIR model for the analysis of tailored social distancing and tracing approaches to cope with COVID-19. *Scientific Reports* **11** (2021).

8. T. Hale, T. Atav, L. Hallas, B. Kira, T. Phillips, A. Petherick, A. Pott, Variation in US states responses to COVID-19. *Blavatnik School of Government* (2020).

9. R. Hinch, W. J. M. Probert, A. Nurtay, M. Kendall, C. Wymatt, M. Hall, K. Lythgoe, A. B. Cruz, L. Zhao, A. Stewart, L. Ferritti, D. Montero, J. Warren, N. Mather, M. Abueg, N. Wu, A. Finkelstein, D. G. Bonsall, L. Abeler-Dorner, C. Fraser, OpenABM-Covid19 - an agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. (2020).

10. T. Hoang, P. Coletti, A. Melegaro, J. Wallinga, C. G. Grijalva, J. W. Edmunds, P. Beutels, N. Hens, A systematic review of social contact surveys to inform transmission models of close-contact infections. *Epidemiology* **30**, 723, (http://www.socialcontactdata.org/) (2019).

11. P. R. Hunter, J. Brainard, 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. (2021).

12. T. C. Jones, G. Biele, B. Mühlemann, T. Veith, J. Schneider, J. Beheim-Schwarzbach, T. Bleicker, J. Tesch, M. L. Schmidt, L. E. Sander, F. Kurth, P. Menzel, R. Schwarzer, M. Zuchowski, J. Hofmann, A. Krumbholz, A. Stein, A. Edelmann, V. M. Corman, C. Drosten, Estimating infectiousness throughout SARS-CoV-2 infection course. *Science*, ISSN: 0036-8075, eprint: https://science.sciencemag.org/content/early/2021/05/24/science.abi5273.full.pdf, (https://science.sciencemag.org/content/early/2021/05/24/science.abi5273) (2021).

13. M. E. Kretzschmar, G. Rozhnova, M. C. J. Bootsma, M. van Boven, J. H. H. M. van de Wijgert, M. J. M. Bonten, Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study. *The Lancet Public Health* **5**, e452–e459 (2020).

14. J. Lessler, M. K. Grabowski, K. H. Grantz, E. Badillo-Goicoechea, C. J. E. Metcalf, C. Lupton-Smith, A. S. Azman, E. A. Stuart, Household COVID-19 risk and in-person schooling. *Science* **372**, 1092–1097 (2021).

15. M. Levine-Tiefenbrun, I. Yelin, R. Katz, E. Herzel, Z. Golan, L. Schreiber, T. Wolf, V. Nadler, A. Ben-Tov, J. Kuint, S. Gazit, T. Patalon, G. Chodick, R. Kishony, Decreased SARS-CoV-2 viral load following vaccination. (2021).

16. M. A. J. Luijten, M. M. van Muilekom, L. Teela, T. J. C. Polderman, C. B. Terwee, J. Zijlmans, L. Klaufus, A. Popma, K. J. Oostrom, H. A. van Oers, L. Haverman, The impact of lockdown during the COVID-19 pandemic on mental and social health of children and adolescents. *Quality of Life Research* (2021).

17. E. Mathieu, H. Ritchie, E. Ortiz-Ospina, M. Roser, J. Hasell, C. Appel, C. Giattino, L. Rodés-Guirao, A global database of COVID-19 vaccinations. *Nature Human Behaviour* (2021).

18. D. McFadden, A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, 995–1026 (1989).

19. M. Melegari, M. Giallonardo, R. Sacco, L. Marcucci, S. Orecchio, O. Bruni, 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**, 113692 (2021).

20. M. J. Mina, K. G. Andersen, COVID-19 testing: One size does not fit all. *Science* **371**, 126–127, ISSN: 0036-8075, eprint: https://science.sciencemag.org/content/371/6525/126.full.pdf, (https://science.sciencemag.org/content/371/6525/126) (2021).

21. J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, *et al.*, Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine* **5** (2008).

22. E. Petter, O. Mor, N. Zuckerman, D. Oz-Levi, A. Younger, D. Aran, Y. Erlich, Initial real world evidence for lower viral load of individuals who have been vaccinated by BNT162b2. (2021).

23. E. Pritchard, P. C. Matthews, N. Stoesser, D. W. Eyre, O. Gethings, K.-D. Vihta, J. Jones, T. House, H. VanSteenHouse, I. Bell, J. I. Bell, J. N. Newton, J. Farrar, I. Diamond, E. Rourke, R. Studley, D. Crook, T. Peto, A. S. Walker, K. B. Pouwels, Impact of vaccination on SARS-CoV-2 cases in the community: a population-based study using the UK's COVID-19 Infection Survey. *medRxiv*, eprint: https://www.medrxiv.org/content/early/2021/04/23/2021.04.22.21255913.full.pdf, (https://www.medrxiv.org/content/early/2021/04/23/2021.04.22.21255913) (2021).

24. Robert Koch-Institut, "Corona-Monitoring bundesweit (RKI-SOEP-Studie)," de, tech. rep.

25. L. S. Shapley, *17. A value for n-person games* (Princeton University Press, 2016).

26. R. L. Smith, L. L. Gibson, P. P. Martinez, R. Ke, A. Mirza, M. Conte, N. Gallagher, A. Conte, L. Wang, R. Fredrickson, D. C. Edmonson, M. E. Baughman, K. K. Chiu, H. Choi, T. W. Jensen, K. R. Scardina, S. Bradley, S. L. Gloss, C. Reinhart, J. Yedetore, A. N. Owens, J. Broach, B. Barton, P. Lazar, D. Henness, T. Young, A. Dunnett, M. L. Robinson, H. H. Mostafa, A. Pekosz, Y. C. Manabe, W. J. Heetderks, D. D. McManus, C. B. Brooke, Longitudinal assessment of diagnostic test performance over the course of acute SARS-CoV-2 infection. *The Journal of Infectious Diseases* (2021).

27. S. Stillman, M. Tonin, "Communities and Testing for COVID-19," IZA Discussion Paper No. 14012, 2021.

28. J. Vlachos, E. Hertegård, H. B. Svaleryd, The effects of school closures on SARS-CoV-2 among parents and teachers. *Proceedings of the National Academy of Sciences* **118**, ISSN: 0027-8424, eprint: https://www.pnas.org/content/118/9/e2020834118.full.pdf, (https://www.pnas.org/content/118/9/e2020834118) (2021).

29. T. Wollmershäuser, Konjunkturelle Auswirkungen der zweiten Coronawelle auf ausgewählte Wirtschaftsbereiche. *ifo Schnelldienst Digital* **2**, (https://www.ifo.de/publikationen/2021/aufsatz-zeitschrift/konjunkturelle-auswirkungen-der-zweiten-corona-welle) (2021).

Supplementary Material for

# The Effectiveness of Strategies to Contain
# SARS-CoV-2: Testing, Vaccinations, and NPIs

Janoś Gabler,
Tobias Raabe,
Klara Röhrl,
Hans-Martin von Gaudecker

## Contents

## Appendix A:   Data and Parameters

The model is described by a large number of parameters that govern the number of contacts a person has, the reduction in contacts due to NPIs, the demand for rapid tests and PCR tests, the likelihood of becoming infected on each contact, the likelihood of developing light or strong symptoms or even dying from the disease as well as the duration each stage of the disease takes.

### A.1   Course of Disease

This section discusses the parameters governing the course of disease, their sources and how we arrived at the distributions used in the paper.[6] See Figure 1b for a summary of our disease progression model.

We denote the latent period—i.e., the time span between infection and the start of infectiousness—by $\gamma_{infectious}$. (1) 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 (2) 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.* (3) and in line with He *et al.* (4) we set the length of the presymptomatic stage of age group $a$, $\gamma_{symptoms,a}$ to be one or two days. The probability to become symptomatic for 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.* (5).

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 (6)) and that the viral load between symptomatic and asymptomatic individuals are similar (Zou *et al.* (7), Byrne *et al.* (8), Singanayagam *et al.* (9)). Our distribution of the duration of infectiousness is based on Byrne *et al.* (8). 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.[7] Excluding the most extreme combinations, we arrive at 3 to 11 days as the duration of infectiousness.

We use the duration to recovery of mild and moderate cases reported by (10, Figure S3, Panel 2) for the duration of symptoms for non-ICU requiring symptomatic

---

6. Additional information can be found in the online documentation.

7. 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.* (9).

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 we rely on data by ([11])) and ([12]) ($\gamma_{icu,a}$). For those who will require intensive care we follow Chen *et al.* ([13]) who estimate the time from symptom onset to ICU admission as $8.5 \pm 4$ days. This aligns well with numbers reported for the time from first symptoms to hospitalization: Gaythorpe *et al.* ([14]) report a mean of 5.76 with a standard deviation of four. 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.* ([12]). 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% day 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 ([15]) reports a mean duration of 18.8 days until recovery and a standard deviation of 12.21 days. We discretize this such that of those who recover in intensive care, 22% do so after one day, 30% after 15 days, 28% after 25 days and 18% after 45 days.

### A.2   The Synthetic Population

We build a synthetic population based on the German microcensus ([16]). 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 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 ([17]). For children between three and six years old, we assume all go to preschool (officially 92.5% according to ([17])). Children that attend a nursery meet in groups of four ([18]) plus one adult care taker every weekday when there are no school vacations. Preschool children meet in groups of nine ([18]) with two adult care takers. These groups are mixed with respect to age but all belong to the same state and mostly to the same county.

Every child that goes to school is part of three different classes that meet every day on weekdays when there are no vacations and no school policies are in place.[8]

---

8. We implement vacations on the federal state level.

Each class consists of approximately 23 students (*19*) and two teachers. 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 daily meeting work group. The group sizes vary to match the number of daily repeating work contacts reported by working individuals in (*20*). 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 A.2e. To match the number of weekly work groups we match each worker with up to 14 other workers into pairs to match the number of reported weekly work contacts shown in Figure A.2f. Each pair is assigned a weekday on which they always meet in the absence of work policies. 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's 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 that practice regularly together. Both daily and weekly groups are created analogously to the work groups but matching the numbers in Figure A.2b and Figure A.2c. 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.

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's 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 ($\xi$) which we calibrate to 15% (*21*). The vaccination groups are created to match the recommendations by the Ständige Impfkommission (*22*).[9] To cover that the Pfizer-

---

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

BioNTech vaccine was later approved for younger age groups we put adolescents and children into two groups that follow after the general population. These groups do not become eligible within our simulation frame until June. The way vaccinations are rolled out in our model is shown in Figure A.1.
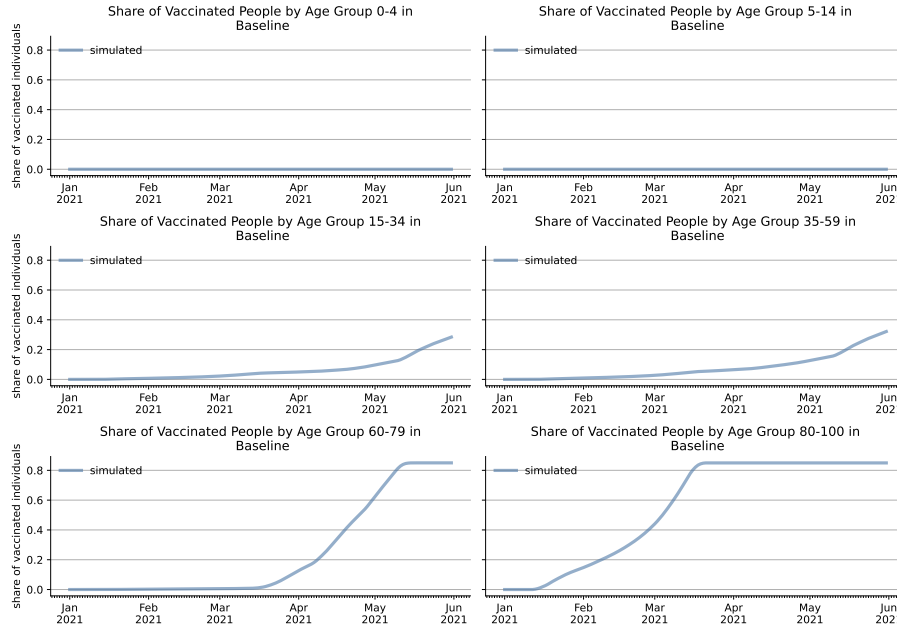


**Figure A.1.** Vaccination Rates by Age Group

*Note:* An individual's vaccination priority depends on her work contact priority, her age group and a random component to capture preconditions like diabetes. 15% of the population refuse to be vaccinated ($\xi$). Adolescents would be vaccinated after the general population and children last. The figure clearly shows that the first vaccinations go to some workers with very high work contact priority and to the 80 to 100 age group 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.

## A.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 (*20*). 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 ($\eta_c$) as shown in Figure A.2.

**(a)** Number of Non Recurrent Other Contacts

**(b)** Number of Daily Recurrent Other Contacts

**(c)** Number of Weekly Recurrent Other Contacts

**(d)** Number of Non Recurrent Work Contacts

**(e)** Number of Daily Recurrent Work Contacts

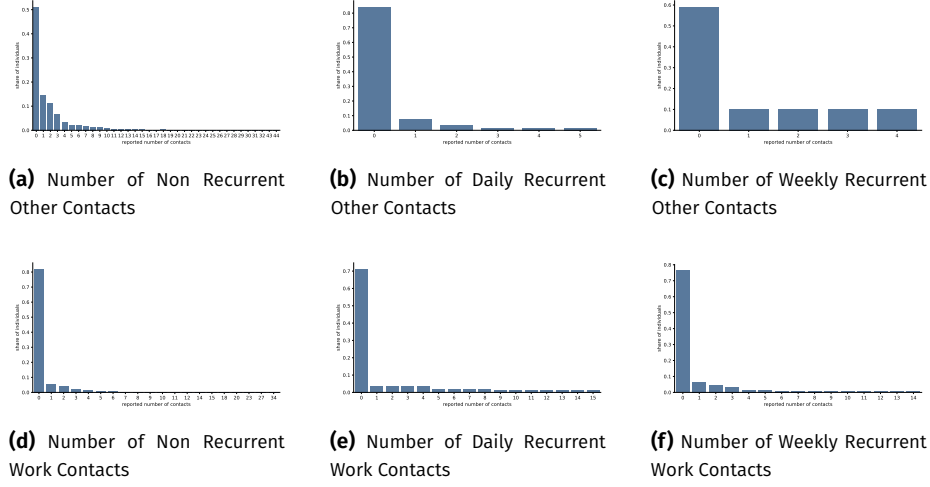**(f)** Number of Weekly Recurrent Work Contacts

**Figure A.2.** Number of Contacts of the Different Contact Types

*Note:* This figure shows the pre-pandemic number of contacts individuals report of different contact types ($\eta_c$). 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. For work contacts, meetings can only take place on work days. The pre-pandemic number of contacts with transmission potential is reduced by policies ($\rho$), seasonality ($\kappa$) and individual responses ($\tau$) to events such as receiving a positive rapid test to the number of actual contacts with transmission potential. The upper row shows the distribution of the number of other contacts individuals report ($\eta_{other}$). 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 lower row shows the distribution of the different types of work contacts ($\eta_{work}$). Work contacts only take place between working individuals.

An exception where we do not rely on the data by (*20*) are the household contacts. Since households are included in the the German microcensus (*16*) on which we build our synthetic population we simply assume for the household contacts that individuals meet all other household members every day. The number of household contacts that happen every day is shown in Figure A.3.
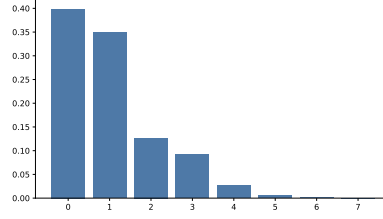
**Figure A.3.** Number of Household Contacts

*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.

## A.4 Assortativity

As explained in section B.5, 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 ($\alpha$). 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 (*20*).
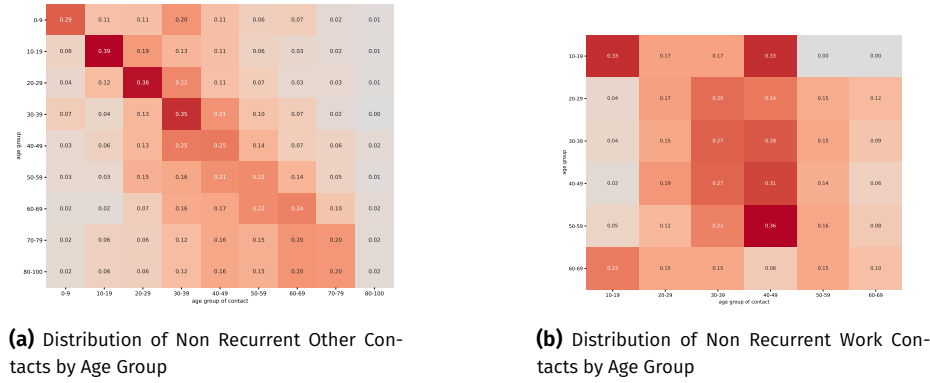


**(a)** Distribution of Non Recurrent Other Contacts by Age Group



**(b)** Distribution of Non Recurrent Work Contacts by Age Group

**Figure A.4.** Assortativity by Age Group for Non Recurrent Other and Work Contacts

*Note:* The figure shows the distribution of non recurrent contacts by age group for other contacts on the left and work contacts on the right. 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. For work contacts, we only show age groups that have a significant fraction of working individuals.

Figure A.4a shows that assortativity of the other 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 A.4b shows that assortativity by age is also important among work contacts.

For recurrent contacts, we constructed 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. 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.

## A.5  Policies

Our policies (denoted by $\rho$) usually affect one of three contact types: education, work and other contacts. Germany had no policies limiting contacts within households so there are no policies on them in our model.[10] For nurseries, preschools and schools we implement vacations as announced by the German federal states as well as school closures, emergency care and rotating schooling 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 A.5a. For work contacts we use the reductions in work mobility reported by the Google Mobility Data (23) to calibrate the reduction in physical work contacts ($\rho_{w, attend, t}$). Reductions in work contacts are not random but governed through a work contact priority where the policy changes the threshold below which workers stay home. Figure A.5b shows the share of workers that go to work over time at the federal German level. We use the data on the state level 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, leading to a 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 not calibrate the policies from data but estimated the policy effects. The estimation and values are detailed in Section B.10 and Figure B.3.

---

10. 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.
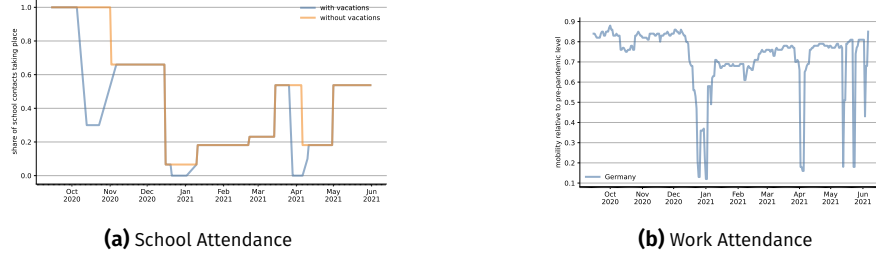
**(a)** School Attendance



**(b)** Work Attendance

**Figure A.5.** The Contact Reduction Effects of School and Work Attendance Policies

*Note:* The left figure shows the approximate share of school contacts taking place with and without vacations factored in. The dates on which schools have vacation are decided at the state level. Vacations are directly implemented in our model with no school contacts taking place on weekends and during vacations (by federal state) just like the schooling mode (full operation, emergency care, rotating schemes with half class sizes etc.). The figure is, thus, only an illustration that shows the approximate share of contacts taking place compared to the pre-pandemic level with and without vacations. The difference between the lines show when vacations take place and to what degree. For example all states have fall vacations but the timing varies strongly between states while all states close schools during the Christmas holidays. The right figure shows the work mobility as reported by (*23*). 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 we use. The figure shows the share for Germany as a whole. To capture the effect that local policies, school vacations, 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.

## A.6 Rapid Test Demand

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

(1) someone plans to have work contacts

(2) someone is an employee of an educational facility or a school pupil

(3) a household member has tested positive or developed symptoms

(4) someone has developed symptoms but has not received a PCR test

(5) someone plans to participate in a weekly non-work meeting

For work contacts, we know from the COSMO study ((*24*), 20th/21st of April) 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 ($\pi_{w, s, t}$). Mid march, 20% of employers offered tests to their employees (*25*). In the second half of March, 23% of employees reported being offered weekly rapid tests by their employer (*26*). This share increased to 61% until the first days of April (*27*, *28*). Until mid April 72% of workers were expected to receive a weekly test offer (*27*, *28*). However, according to surveys conducted in mid April (*24*), less than two thirds of individuals with work contacts receive a test offer. Starting on April 19th employers were required by law to provide two weekly tests to their employees (*29*). We assume that compliance is

incomplete and only 80% of employers actually offer tests. We interpolate between these points linearly, arriving at the blue line in Figure A.6. In addition, we increase the frequency of testing ($\theta_{t,work}$) from weekly to twice weekly during April.

We assume that employees in educational facilities start getting tested in 2021 and that by March 1st 30% of them ($\pi_{teacher,t}$) are tested weekly ($\theta_{before\,Easter,educ} = 7$). The share increases to 90% for the week before Easter. At that time both Bavaria (30) and Baden-Württemberg (31) were offering tests to teachers and North-Rhine Westphalia (32) and Lower Saxony (33) were already testing students and tests for students and teachers were already mandatory in Saxony (34). After Easter we assume that 95% of teachers get tested twice per week ($\theta_{after\,Easter,educ} = 3$).

Tests for students started later (31, 32) so we assume that they only start in February and only 10% of 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% (32). After Easter the share of students receiving twice weekly tests is set to 75%. This is based on tests becoming mandatory in Bavaria (35) and North Rhine-Westphalia (36) after their Easter breaks and on the 19th in Baden-Württemberg (37), after which we assume twice weekly rapid tests to be mandatory for all students in Germany. Again, we interpolate linearly between these points and arrive at the purple line for teachers and the red line for school students in Figure A.6.

To limit our degrees of freedom, we only have one parameter that governs how many individuals do a rapid test because of any of the private demand reasons ($\pi_{private,t}$).[11] We assume that there is no private rapid test demand until March when both the citizens' tests and rapid tests for lay people started to become available (38, 39) and other access to rapid tests was very limited.

According to the COSMO study (40) 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 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 which we estimate in our model. The resulting share of private rapid test demand is shown as the green line in Figure A.6 (also see Section B.10 for details on the estimation).

---

11. The reasons that can lead to an individual doing a rapid test for private reasons are own symptoms but no PCR test, planned weekly leisure meeting or a symptomatic or positively tested household member
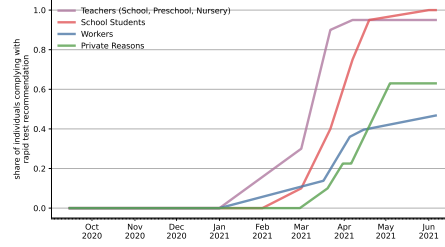
**Figure A.6.** Share of Individuals Doing a Rapid Test.

*Note:* Rapid test demand can be triggered by individuals planning to have education contacts ($\pi_{teacher, t}$ 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 ($\pi_{private, t}$). 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 a worker in May. In that time workers are encouraged to test themselves twice weekly (or every three days, i.e. $\theta_{May, work} = 3$) but there is no general requirement to test themselves. If the worker has not done a test within the last four days in our model she will demand a test if her (time-constant) compliance parameter belongs to the upper 60% in the population.

## A.7 Share of Detected Cases

One important feature of our model is that we distinguish between undetected and detected cases and that we model which cases are detected and which are not (see Section B.7 for a detailed description for how we model both rapid and PCR tests). For our model it is important to have an estimate for the share of cases that is detected in the absence of rapid tests ($\psi_t$). For this we rely on the (Dunkelzifferradar Project *41*) 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 the 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[12]

Since vaccinations started after Christmas 2020 and these were predominantly given to nursing homes 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 also achieve in our model an increase in the share of detected cases but this is driven from inside our model through increased rapid testing which lead follow-up PCR tests when they are positive (see Section C.3 and B.7).

---

12. The testing policy changed in November (*42*). However, this only moved the rare PCR tests more towards vulnerable groups.

Lastly, we model reductions in the share known 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 known cases. The resulting share of detected cases in the absence of rapid tests is shown in Figure A.7.
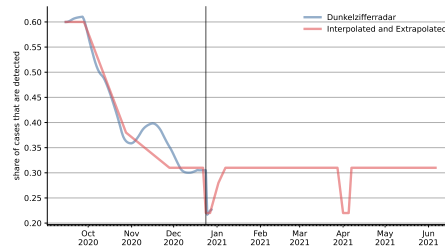


.

**Figure A.7.** 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 known cases that was estimated through the case fatality ratio by the (Dunkelzifferradar Project *41*) 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 also 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 C.9 for how the share of detected cases develops in our model for each age group

# Appendix B:   Detailed Model Description

## B.1   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.* (*43*) and Cuevas (*44*)). 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.* (*15*) is the closest in spirit to ours. They model detailed contact networks for workplaces, schools and households

**[HM 6]**
I'd rather be close to Aleta et al. ;-)

26

and can evaluate the effect of several NPIs. The main focus of their application are contact tracing policies.

Aleta *et al.* (*45*) 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.* (*46*) simulate the entire Austrian population. The use data from the first wave (February 21 to April 9, 2020) to calibrate ther model and predict the effect of different NPIs and contact tracing policies until November 2020. They the same data provided by Mossong *et al.* (*20*) as we use to calibrate contact networks for households, workplaces and schools. The model focuses on analyzing the effect of different contact tracing strategies and not on modelling enacted Austrian policies over a long period of time.

Moreover, there are several forthcoming applications of agent-based simulation models with contact networks. Examples are Basurto *et al.* (*47*), Delli Gatti and Reissl (*48*) and Mellacher (*49*).

**[HM 7]**
What is that supposed to mean?

**[Janos 1]**
Does that mean we should not cite the working papers here?

Our model combines elements from the above models and adds other ones. 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 (*50*). Despite having few free parameters our model does en excellent job in explaining observed case numbers and the spread of the B.1.1.7 mutation over more than nine months of data.

2. We have an extremely fine grained of schools and preschools. We can thus easily model all NPIs that have been implemented in Germany in the past 9 months. 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 over since the start of the second wave. Since the vast majority of cases has occurred in that time period and we also model unobserved infections our simulations take into account that many people are already immune because they are recovered from an infection and that this 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. We fit estimates of the share of known cases from other studies for time periods where those are available but also estimate a plausible share of known cases after vaccinations become widespread and methodologies that rely on a case fatality ratio to identify the share of known cases are not applicable anymore.

## B.2 Summary

To predict and quantify the effects of a wide variety of fine-grained social distancing policies, vaccinations and rapid testing, we propose a different model structure. Our model inherits many features from prototypical agent-based simulation models but replaces the contacts between moving particles by contacts between individuals who work, go to school, live in a household and enjoy leisure activities.

The structure of the model is depicted in Figure 1a.

We distinguish between eight types of contact models which are all listed in Figure 1a: households, recurrent and random work contacts, recurrent and random leisure contacts, and nursery, preschool, and school contacts.

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.

The infection probabilities of contacts vary with contact type, age of the susceptible person, and the virus strain of the infected person. Moreover, they follow a seasonal pattern. The strength of the seasonality effect is higher for contacts that are easy to be 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).

Once a person is infected, the disease progresses in a fairly standard way which is depicted in Figure 1b. Asymptomatic cases and cases with mild symptoms are infectious for some time and recover eventually. Cases with severe symptoms additionally require hospitalization and lead to either recovery or death.

After rapid tests become available, people who work or go to school can receive rapid tests there. Moreover, people can decide to make a rapid test if they develop symptoms, have many planned contacts or observe cases in their contact network. People who have a positive rapid test demand a confirmatory PCR test with a certain probability. Moreover, PCR tests can be demanded because of symptoms or randomly.

**[Klara 20]**

at the moment it's implemented as planing to participate in a weekly leisure meeting. Should we be more explicit here?

This rich model of PCR and rapid tests leads to a share of detected cases that varies over time and across age groups. It also allows to quantify the effect of changes in testing policies on the dynamic of infections.

People who have symptoms, received a positive test, or had a risk contact can reduce their number of contacts across all contact types endogenously. The extent to which this is done is calibrated from survey data.

The model makes it very simple to translate policies into model quantities. For example, school closures imply the complete suspension of school contacts. A strict lockdown implies shutting down work contacts of all people who are not employed in a systemically relevant sector. It is also possible to have more sophisticated policies

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

An important feature of the model is that the number of contacts an individual has of each contact type can be calibrated from publicly available data ([20]). This in turn allows us to estimate policy-invariant infection probabilities from time series of infection and death rates using the method of simulated moments ([50]). Since the infection probabilities are time-invariant, data collected since the beginning of the pandemic can be used for estimation. Moreover, since we model the testing strategies that were in place at each point in time, we can correct the estimates for the fact that not all infections are observed.

The model has a very 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 the additional model features.

## B.3   Modeling 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 only those 50 people had contact with an infectious person and the disease has an infection probability per contact ($\beta$) of one or whether everyone met one infectious person but the disease has an infection probability of only 5 percent per contact. SEIR models do not distinguish between the number of contacts ($\eta$) and the infectiousness of each contact ($\beta$). 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 and type 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 1a and can be further grouped in 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: Each student meets all of his 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 states. 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. 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.

The number of random and recurrent contacts at the workplace, during leisure activities and at home is calibrated with data provided by Mossong *et al.* (*20*). For details see Section A.3. In particular, we sample the number of contacts or group sizes from empirical distributions. It would also be possible to use economic or other behavioral models to predict the number of contacts.

## B.4    Reducing Numbers of Contacts via NPIs

Our model makes it very 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 use the full time series of infections and fatality rates to estimate the model parameters.[13]

---

13. See Avery *et al.* (*51*) for an explanation why it can be harmful to use too long time series to estimate simple SEIR type models.

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

On an abstract level, a policy is a functions that modifies the number of contacts of one contact type. 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 (*52*).

The most complex policies are typically found in the education sector. Since the beginning of 2021 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. On top of that there are different policies for allowing young students whose parents work full time to attend school even on days where they normally would not. For details on how we calibrate these policies see Section A.5.

Importantly, policies can depend on the health states of participating individuals. This allows to quarantine entire school classes if one student tested positive or to implement official or private contact tracing.

For some policies the exact effect on each contact type is not easy to determine. If this refers to a policy has been active during the estimation period, it is possible to estimate such parameters by fitting the model to time series data of infection rates. This is only possible if the policy was not active during the whole estimation period and thus the infection probabilities can be identified separately. We do this to account for hygiene measures at school and in the workplace that have been in effect since November 2020.

Not all things that reduce contacts compared to the pre-pandemic level are driven by NPIs. Therefore, we also model endogenous contact reductions that can depend on the health state of individuals, known risk contacts or the local incidence of infections. Examples are strong contact reductions for symptomatic individuals or those who have a positive PCR or rapid tests or contact reductions when a houshold member tested positive. The extent to which contacts are reduced can be calibrated from surveys. For an application of our model showcasing private contact tracing in the context of the Christmas holidays see (*53*).

### B.5 Matching Individuals

The empirical data described above only allows to estimate the number of contacts each person has. In order to simulate transmissions of Covid-19, the numbers of contacts has to be translated into actual meetings between people. This is achieved by a matching algorithm:

As described in section B.3, 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 value 0 (non-participating) and 1 (participating) and can depend on the weekday, date and health state. This can be used to model vacations, weekends or symptomatic people who stay home (see section B.4 for details).

Given those 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

$$P(infection) \ = \ \beta_c \cdot s_{c,t} \cdot \sigma_v \cdot \zeta_a \tag{B.1}$$

where $beta_c$ 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 B.4), $\sigma_v$ is the infectiousness factor of virus variant $v$ and $\zeta_a$ 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 suspected role of aerosol transmission for Covid-19 (*54*, *55*). 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 randomly 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 introduce 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:

$$f_{zj} = \underbrace{\alpha_{c,a_i,a_j,county_i,county_j}}_{\text{Group Probability}} \cdot \underbrace{\frac{K_{z,j,c}}{\sum_{l=1}^{N} K_{z,l,c} \cdot \mathbb{I}_{county_l=county_j \wedge a_l=a_j}}}_{\text{Individual probability}} \tag{B.2}$$

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 B.1. Finally we update the remaining numbers of unmatched contacts by setting:

$$K_{z+1,i,c} = K_{z,i,c} - 1 K_{z+1,j,c} = K_{z,j,c} - 1 \tag{B.3}$$

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 evaluate Equation B.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 B.2. Next we sample an individual from this group with the Individual probabilities defined in Equation B.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.[14]

## B.6   Course of Disease

The disease progression in the model is fairly standard. It is depicted in Figure 1b and the values and source of the relevant parameters are describes in Section A.1.

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. Thus they can potentially infect many other people (56). The probability to develop symptoms with Covid-19 is highly age dependent with 75% of children not developing clinical symptoms (2).

A small share of symptomatic people will develop strong symptoms that require intensive care. The exact share and time span 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.

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 [15].

**[Janos 2]**
Replace this by a link to the relevant table with parameters in the data appendix, once we have that table.

## B.7   Testing

Having a realistic model of PCR and rapid tests is crucial for two reasons: Firstly, only via a testing model the simulated infections from the model can be made comparable

---

14. 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.

15. Detailed information on the calibration of the disease parameters is available as part of our online documentation.

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 aim for a simpler model of PCR and rapid tests that can be calibrated with available data on test demand and availability and – nevertheless – can produce a share of undetected cases that varies over time and across age groups and agrees with other estimates over the time periods where they are available.

PCR tests are modeled since the beginning of the simulation and determine whether a infections 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 tests demand a confirmatory PCR test. However, positive rapid tests can have a strong effect on the infection dynamics 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 (*41*). Next, we determine how many of these tests should go to symptomatic and asymptomatic individuals from data by the RKI. Then, we sample the individuals to which those tests are allocated from the pools of symptomatic and asymptomatic infected but not yet tested 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 in time periods where many tests are done because of symptoms which is in line with the estimates from the literature.

At the beginning of 2021, two challenges arise: Firstly, the externally estimated share of detected cases from Paul *et al.* (*41*) can no longer be used because it is based on case fatality rates which drastically change due to vaccinations. 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.

The second challenge is solved by mechanistic rapid test demand models for the workplace, schools and by private individuals. The calibration of these models is described in Section A.6. Figure 2d shows that the number of performed rapid tests in the model fits the empirical data well (where empirical data is available).

In contrast to PCR tests, rapid tests are not perfect and can be falsely positive or falsely negative. While the specificity of rapid tests is constant over time, their sensitivity strongly depends on whether the tested individual is already infectious and if so for how long he has been infectious. Before the onset of infectiousness the sensitivity is very 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, a recent study has found no basis for this concern.

While rapid tests do not directly enter official case numbers, many positively tested individuals confirm their rapid test result with a PCR test. Importantly, those PCR tests are made in addition to the tests that would have been done otherwise. Section C.3 discusses the effect of rapid tests on the share of detected cases.

**[Janos 5]**
cite charite study on self administered tests

**[Janos 6]**
Need to link to data appendix once the number is there and maybe mention the number

**[Klara 21]**
Talk here about reactions to tests and symptoms with contact reductions. (If not here, find other place where to talk about this in detail.)

**[Klara 22]**
Philipp suggests to write that we choose upper bounds

## B.8 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 (57).

We follow (58) and (59) in modeling seasonality in the transmission of SARS-CoV-2 as a multiplicative factor on infection probabilities. The factor follows a sine curve that reaches its maximum at January first 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:

$$s_{c,t} = 1 + 0.5\kappa_c sin\left(\pi\left(\frac{1}{2} + \frac{t}{182.5}\right)\right) - 0.5\kappa_c \qquad (B.4)$$

Where $\kappa_k$ is difference in the seasonality factor between peak infectiousness and lowest infectiousness.

The subscript $k$ 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
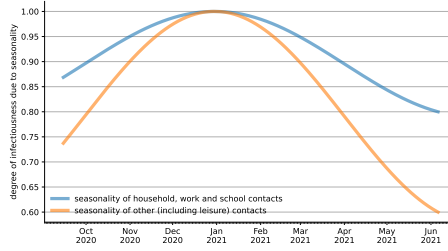
**Figure B.1.** Seasonality by Type of Contact

*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.

doing leisure activities) are mostly happening outside in the summer. Therefore, transmission via those contacts should have a stronger seasonal pattern.

We calibrate $\kappa_{strong}$ to 0.42 and $\kappa_{weak}$ to 0.21. This is in line with (*59*) and (*58*). The two seasonality curves are shown in Figure B.1

## B.9   Initial Conditions

Consider a situation where you want to start a simulation with the beginning 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 your 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 (school closures), etc.. The distribution of courses of diseases in the population at the begin 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 available characteristics like age and geographic information. The matching must be done for each day of a longer time frame like a month to have individuals with possible health states. Then, health statuses evolve until the begin of the simulation period without simulating infections by contacts. We also correct reported infections for a reporting lag and scale them up to arrive at the true number of infections.

## B.10    Estimated Parameters

We estimate parameters that cannot be calibrated outside of the model with the method of simulated moments (*50*) by minimizing the distance between simulated and observed infection rates and fatality rates (disaggregated by region and age groups). Since our model includes a lot of randomness, we average simulated infection rates over several model runs.

All estimated parameters are described in Table B.1.

**Table B.1.** Estimated Parameters

| notation | estimate | note |
|---|---|---|
| **Infection Probabilities** | | |
| $\beta_{household}$ | 0.1 | base probability of getting infected by an infectious household member |
| $\beta_{school}$ | 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 |
| $\beta_{work}$ | 0.1475 | base infection probability for work contacts |
| $\beta_{other}$ | 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 |
| **B.1.1.7 Introduction** | | |
| $\omega_{B.1.1.7,\ Jan\ 31}$ | 0.986 | number of B.1.1.7 cases per 100 000 individuals to import on January 31st. Imported B.1.1.7 cases gradually rise from January 1st where 0 cases are imported. No cases are imported in other months. |
| **Rapid Test Introduction** | | |
| $\pi_{private,t}$ | Figure A.6 | 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. In between those levels are linearly interpolated. The remaining rapid tests demands are calibrated from surveys. See Section A.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 extremely 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) 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 B.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 strand of the infected person. The base infection probability is only equal to the actual infection probability when all of those factors are 1. This would be the case for a contact between an 80+ year old susceptible person with a person who is infected with the B.1.1.7 strand of the virus on January first.

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 1 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 actually have physical contact. Thus the infection probabilities for those contact types are actually the product of the probability to actually 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 intuitive to look at how many infections were actually caused by each contact type. This is depicted in Figure B.2. We can see that work and other contacts are the main driver 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 C.14 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 stricter mask mandates and distancing rules were introduced. It is estimated to reduce infectiousness of contacts by one third.

**Figure B.2.** Daily share of infections by contact type

*Note:* Daily hare 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.
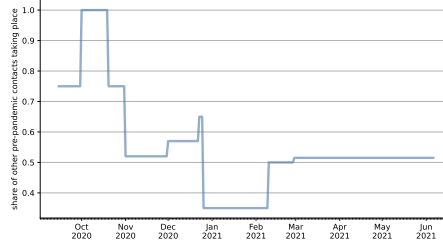
**Figure B.3.** Share of Pre-Pandemic Other Contacts Taking Place with Infection Potential

*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 B.1

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 coincides with policy changes and is not determined from the data. 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 also depicted in Figure B.3.

While we estimate 9 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 where 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 Table B.1 are thus a combination of contact reduction and hygiene measures.

Finally we estimate one parameter that governs the introduction of the B.1.1.7 virus 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 B.1.1.7 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, below we give a rough intuition which features of the data help us to estimate each parameter.

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) and they affect different subgroups of the population differently (e.g. $\beta_{school}$ most strongly affects kids whereas $\beta_{work}$ has the strongest effect on adults in working age and $\beta_{other}$ 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 B.1.1.7 mutation can be identified from the share of that virus strand 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.

## B.11 Shapley Values

We decompose the effects of different NPIs and seasonality on the infection rates with Shapley values. Shapley values (*60*) 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 $v$ which maps subsets of $N$ to the real numbers. The function $v$ is also called the characteristic function and assigns a value to a coalition. Then, the Shapley value $\phi$ for player $i$ is

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

The last term $(v(S \cup \{i\}) - v(S))$ 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 $v$.

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

To produce Figure 3c and Figure 3d, 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.

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

- The 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 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.

## B.12 Overview Model Parameters

**Table B.2.** Contacts, Matching and Policies

| name | notation | time active | source | description |
|---|---|---|---|---|
| n_contacts | $\eta_{c,n}$ | always | Mossong *et al.* (*20*) | probability to have $n$ contacts in contact type $c$. For non recurrent contacts the number of contacts is drawn for each individual every day. For recurrent contacts individuals are assigned group identifiers that determine the groups in which they meet either daily or weekly. The number of groups individuals belong to are calibrated to arrive at the number of recurrent contacts reported in Mossong *et al.* (*20*). |
| degree_of_assortativity | $\alpha_{c,a_i,a_j,county_i,county_j}$ | always | Mossong *et al.* (*20*) | 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. The degree of assortativity depends on the contact type 'c'. |
| work_attend_multiplier | $\rho_{w,attend,t}$ | always | Google (*23*) | share of workers that continue to have work contacts, i.e. do not work from home on date $t$. We use the reduction in work mobility reported by Google (*23*) as a proxy for the share of workers that work from home. |
| hygiene_multiplier | $\rho_{hygiene}$ | since November 2020 | estimated | reduction in transmission during work and educational contacts due to stricter hygiene measures such as wearing face masks. |
| other_multiplier | $\rho_{other,t}$ | always | estimated | reduction in the transmission during other contacts, such as leisure. This incorporates both hygiene measures as well as the reduction of physical meetings. There are nine breakpoints for the whole estimation period. |

**[HM 8]**

Copied from main text. Do we describe the parameter for fraction of vaccination-induced immunity/sterility somewhere? I could not find it. Nothing special about the 75%....
Also, start out more positively. Something like "We do not distinguish between first and second doses for computational reasons. We thus use a weighted average ..."
75% is lower than what is usually reported for after the second dose of the Biontech/Pfizer vaccine, which is most commonly used in Germany. We choose it because our model neither includes booster shots, nor does it allow vaccinated individuals who became immune to transmit the disease(21–23). If anything, these assumptions would overstate the effect of vaccines for our study period. This would be different if a large fraction of vaccinated individuals had received a second dose already.

**Table B.3.** Infection Probabilities and Virus Variants

| name | notation | time active | source | description |
|------|----------|-------------|--------|-------------|
| infection_prob | $\beta_c$ | 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. |
| susceptibility | $\zeta_a$ | always | Davies *et al.* ([2]) | 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 | $\kappa_c$ | always | Gavenčiak *et al.* ([59]) | 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 B.8 for an explanation. |
| variant_infectiousness | $\sigma_v$ | always | Davies *et al.* ([64]) | 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 $t$ |

**Table B.4.** Disease and Vaccination Model

| name | notation | time active | source | description |
|---|---|---|---|---|
| p_duration_immune | $\gamma_{immune,\,d}$ | always | see Section A.1 | probability to stay immune for $d$ days after having contracted CoViD-19 |
| p_duration_until_infectious | $\gamma_{infectious,\,d}$ | always | see Section A.1 | probability to become infectious $d$ days after infection |
| p_duration_of_infectiousness | $\gamma_{stop\,infectious,\,d}$ | always | see Section A.1 | probability that infectiousness lasts $d$ days |
| p_duration_until_symptoms | $\gamma_{symptoms,\,a,\,d}$ | always | see Section A.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 A.1 | probability for individuals of age group $a$ that symptoms last $d$ days |
| p_duration_until_icu | $\gamma_{icu,\,a,\,d}$ | always | see Section A.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 A.1 | probability to recover after $d$ days from requiring intensive care |
| p_duration_until_death | $\gamma_{dead,\,a,\,d}$ | always | see Section A.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 | Pritchard *et al.* (*63*) | probability to develop immunity $d$ days (possibly infinite) after being vaccinated |
| share_vaccine_refusers | $\xi$ | always | Frisch (*65*) | share of individuals refusing to be vaccinated |

**Table B.5.** Rapid Testing

| name | notation | time active | source | description |
|------|----------|-------------|--------|-------------|
| rapid_test_specificity | $p_{negative\|not\ infected}$ | 2021 | Brümmer *et al.* (*66*) | the probability of an uninfected person to receive a negative rapid test result. |
| rapid_test_sensitivity | $p_{positive\|infected,\ i,\ t}$ | 2021 | Smith *et al.* (*67*) | 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. |
| share_accepting_work_rapid_test_offer | $\pi_{w,\ d}$ | 2021 | Betsch *et al.* (*24*) | 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. |
| share_workers_receiving_rapid_test_offer | $\pi_{w,\ s,\ t}$ | 2021 | see A.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 A.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 A.6 | share of school pupils that do rapid tests at school on date $t$. |
| share_private_rapid_test_demand | $\pi_{private,\ t}$ | 2021 | see A.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 | | Frequency with which individuals test themselves in educational settings. Weekly before Easter 2021, twice weekly after Easter 2021. |
| rapid_test_work_freq | $\theta_{t,\ work}$ | 2021 | | Frequency with which complying workers test themselves. Weekly before April 2021. Then the frequency increases to twice weekly during April 2021. |

**Table B.6.** 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 (68) | probability that it takes $d$ days between the performance of a PCR test and receiving the result. |
| share_of_tests_for_symptomatics | $\chi_{symptom}$ | always | calibrated from RKI | share of positive PCR tests that are performed on individuals because of Covid-19 symptoms |
| share_w_positive_rapid_test_ requesting_pcr | $\chi_{confirmation}$ | 2021 | Betsch et al. (24) | share of individuals with positive rapid test that seek a PCR test |
| share_known_cases_without_ rapid_tests | $\psi_t$ | always | Paul et al. (41) | share of cases that would be detected in the absence of rapid tests (see Section B.7) |
| symptomatic_multiplier | $\tau_{symptoms, c}$ | always | | share of symptomatic individuals that still have contacts of type $c$ |
| positive_pcr_multiplier | $\tau_{positive\ PCR, c}$ | always | | 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. (24) | share of individuals with a recent positive rapid test that still have contacts of type $c$ |

## Appendix C:   Additional Results

### C.1   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 the B.1.1.7 mutation, vaccinations.

Overall, our model achieves an excellent fit of the two waves of infections with few free parameters (Figure C.1a). As a result the effective replication number $R_t$ also closely follows that reported by the RKI (see Figure C.1b). 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 test.

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 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 C.1b) 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 ($\psi_t$) falls. This is the case in the fall of 2020 (see Figure A.7) 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.
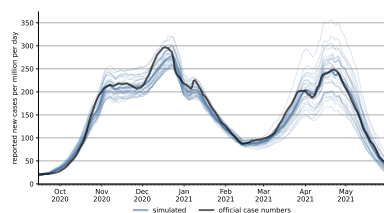
**(a)** Observed Incidence in the Model and as Reported by the RKI
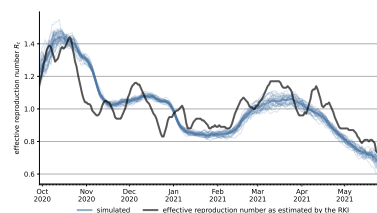


**(b)** Effective Replication Number $R_t$ in the Model and as Reported by the RKI

**Figure C.1.** 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 left figure shows the daily incidence rate per million for the simulated reported infection rates. The mean infection rate is the thick blue line. Single simulation runs are plotted in lighter and thinner lines. The official case numbers as reported by the RKI are plotted in black. The fit is overall very good. The higher the mean incidence and the stronger the growth the more variance there is between simulation runs. The right 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 ($\psi_t$) 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.

Zooming into the different age groups in Figure C.2, 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 years old 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 too optimistic assumptions on school reopenings.

**[Klara 23]**
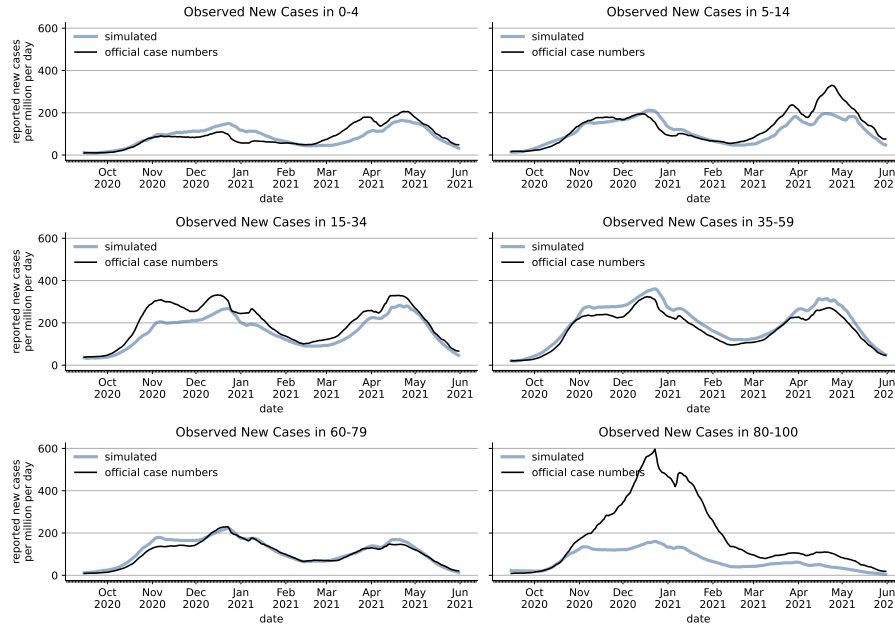cautious or slow? I think optimistic might be ambiguous here.

**Figure C.2.** 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.

Our model fit is also very good for the different German federal states. 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 (*23*) at the state level combined with county and age group specific initial conditions (see Section B.9) 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.
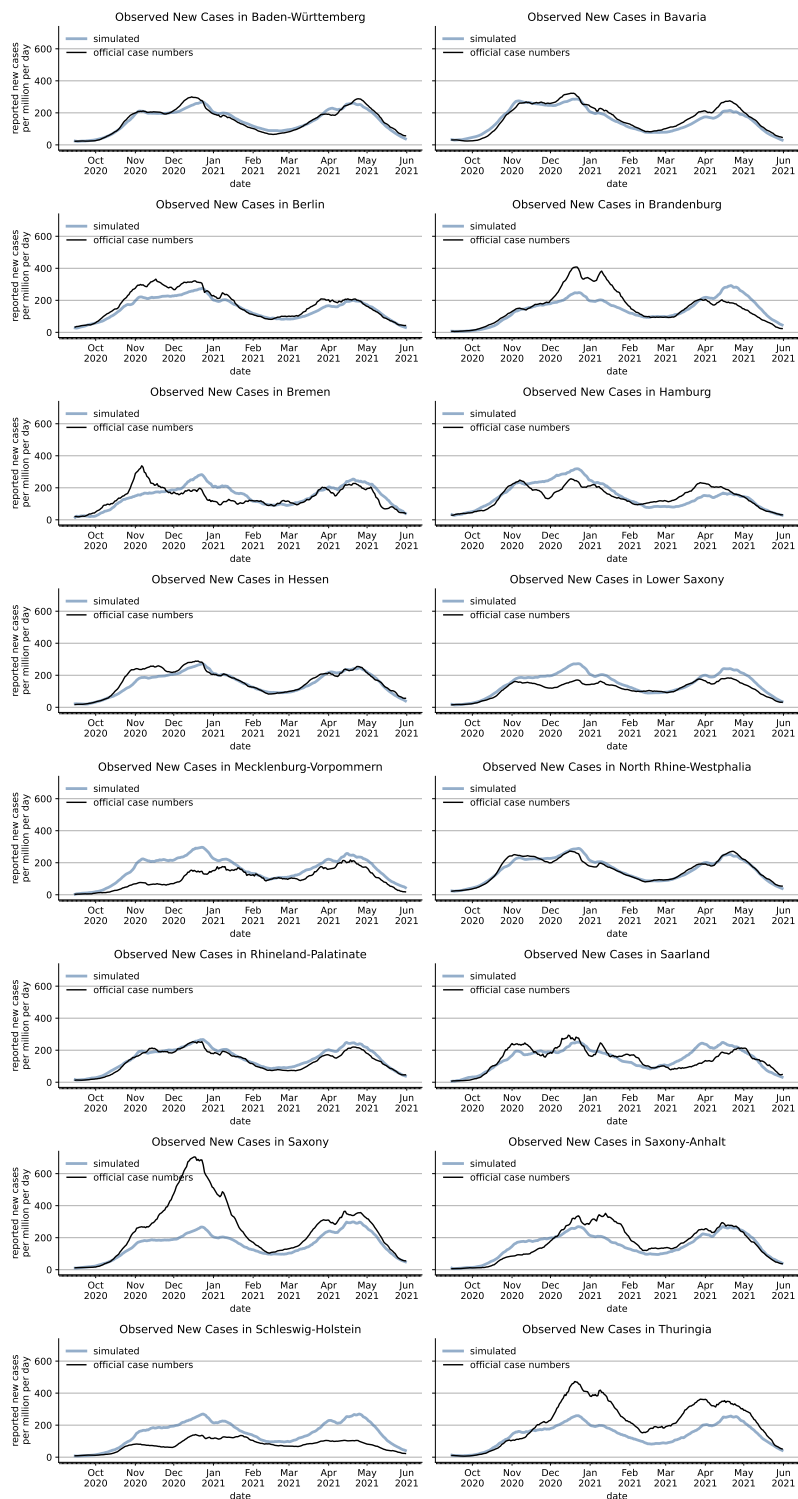
**Figure C.3.** Simulated and Empirical Infections by Federal State

*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.

We fit the proliferation of the B.1.1.7 variant quite exactly despite only introducing a few cases in January ($\omega_{B.1.1.7,t}$) as can be seen in Figure C.4a. Since we only model B.1.1.7 and do not include other variants, B.1.1.7 reaches a share of nearly 100% by May while the true rate plateaued at 90%. By the end of May B.1.617.2 gained traction in Germany. However, given that B.1.617.2 made up less than 5% even at the end of our simulation period, we did not include it in our model.
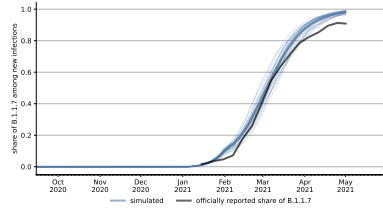


**(a)** Share of B.1.1.7 in the Model and as Reported by the RKI
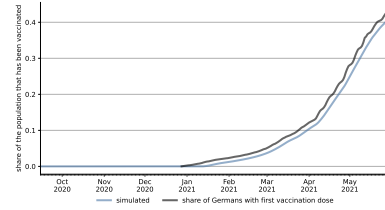


**(b)** Share of Vaccinated Individuals in the Model and the German Population

**Figure C.4.** Model Fit of the Share of B.1.1.7 and Vaccinations

*Note:* The left figure shows the share of B.1.1.7 as reported by the RKI and as calculated in our model. We only introduce a few cases over the course of January. From then B.1.1.7 takes over endogenously through its increased infectiousness ($\sigma_{B.1.1.7}$). We model no other features of B.1.1.7. We introduce at most 0.986 cases per 100,000 inhabitants. The right figure shows the rate of individuals that are vaccinated in our synthetic population versus in the general German population. Note that we excluded the vaccinations that were given to nursing homes (approx. one percent of the German population). Overall, our model covers a time frame that goes from zero vaccinated individuals to a state where over 40% of the population are vaccinated. Our vaccinations work imperfectly but we do not model different different vaccines nor do we distinguish between first and second shot.

The fit of the share of vaccinated individuals can be seen in Figure C.4b. 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 A.2 and Figure A.1 for the vaccination rates in our model by age group).

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. For the share of individuals that have ever done a rapid test our model is conservative. There are two reasons for this: Firstly, we do not model people who have made 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. Fortunately, Section C.2 shows that our main results are robust to changes in the exact shares of individuals demanding rapid tests.
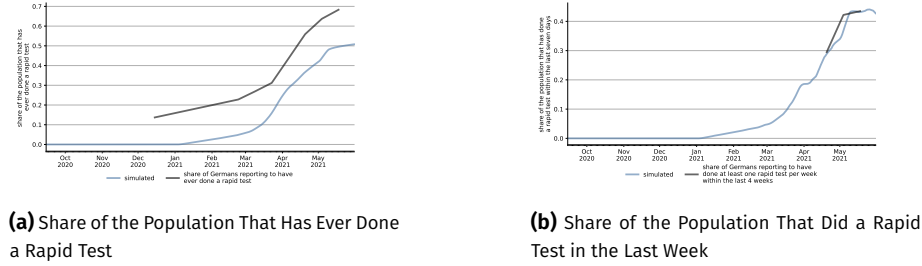


**(a)** Share of the Population That Has Ever Done a Rapid Test

**(b)** Share of the Population That Did a Rapid Test in the Last Week

**Figure C.5.** Share of Individuals With Rapid Tests

*Note:* The figure compares the share of individuals who have ever done a rapid test or done a rapid test within the last week in our simulations to the shares reported in Betsch *et al.* (69). The left panel compares the share of individuals who have ever done a rapid test. The right panel 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. Overall our calibration of rapid tests are slightly conservative. The overall share is below that in the study. We fit the share of weekly tests quite exactly.

## C.2 Model Validation

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 rapid tests (governed by the $\pi$ parameters) that are performed – all of which are things that cannot be exactly known 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 C.7. 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)** Rapid Test Parameters: Early Scenario

**(b)** Rapid Test Parameters: Medium Scenario

**(c)** Rapid Test Parameters: Late Scenario
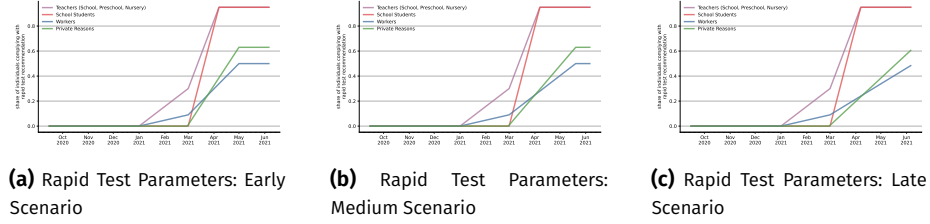
**Figure C.7.** Rapid Test Introduction in the Three Scenarios

*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.

Moreover, the out of sample predictions assume that the share of detected cases ($\psi_t$) 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 C.8. 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.
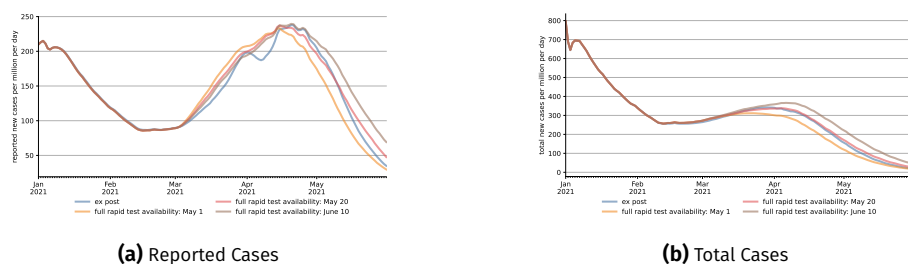
**(a)** Reported Cases



**(b)** Total Cases

**Figure C.8.** 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.

## C.3 Share of Detected Cases

We show the age group specific share of detected cases in Figure C.9.[16]

In 2020 the share of detected cases is much higher in older age groups as they are much more likely to develop symptoms and symptomatic individuals are much more likely to be tested. Starting in 2021 the share of detected cases not only depends on the overall share of detected cases ($\psi_t$) we feed into our model but also on the number of performed rapid tests (governed by the $\pi$ parameters). As more and more (often asymptomatic and presymptomatic) individuals receive positive rapid tests and seek and get confirming PCR tests, the share of detected cases starts to increase, especially after March when rapid tests became much more widely available. This effect is very visible for school pupils (5-14, green line). As twice weekly rapid tests become mandatory in schools and students start going to school after the Easter break, the share of detected cases for that group increases drastically from 25% to over 60% between April and June.

---

16. We explain how we model the detection of cases in Section B.7 and show the overall share of detected cases that we use as input for our testing model in Section A.7.
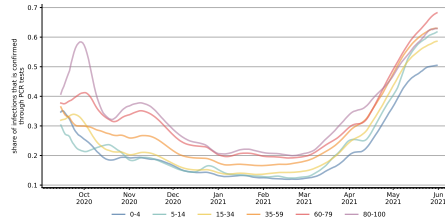
**Figure C.9.** Share of Detected Cases by Age Group

*Note:* The figure shows the share of cases that is reported as an official case via PCR confirmation for each age group in our population. We use the overall share of detected cases ($\psi_t$) that was estimated through the case fatality ratio by Paul *et al.* (41) 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. To get from an overall share of detected cases to the share of cases that is detected in each age group we use that asymptomatic cases are much less likely to be detected. As our model covers age specific asymptomatic rates this endogenously leads to group specific share of detected cases that suggest that infections in younger age groups are under-detected. Starting in 2021 in addition to the overall numbers of detected cases through $\psi_t$, cases are also detected through confirmation of positive rapid tests. This leads to an increase in the share of detected cases for all age groups but in particular for the younger age groups that are covered extensively with rapid tests through the rapid test requirement for participating in school. Age groups that have high shares of workers also have disproportionately increases in their detection rates.

## C.4    Simulated Rapid Tests

In order to make most use out of limited data sources on rapid test usage, we model the number of performed rapid tests as a result of time invariant willingness to do rapid tests and time varying supply side factors and events that trigger rapid tests. Thus, the $\pi$ parameters governing when individuals do rapid tests described in Section A.6 are only indirectly related to the number of rapid tests that are actually performed in the model. When it comes to positive and negative rapid tests, there is even an additional layer because rapid tests are imperfectly sensitive and specific.

In this section we look at how rapid tests expanded in our simulations over time and to what degree they are useful as a screening device despite their imperfections.

We start with the share of the population doing a rapid test, receiving a positive rapid test and receiving a negative rapid test over time by the channel through which the test was demanded in Figures C.10a, C.10b, respectively. 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 work rapid tests are a little ragged because of public holidays. For education rapid tests 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 image is very similar for the share of positive tests, except that the overall number of positive tests starts decreasing in May as rapid test expansion comes to a halt and cases fall, especially the positive share of private rapid tests falls as less and less individuals are triggered to seek a rapid test because of a risk contact in their household.
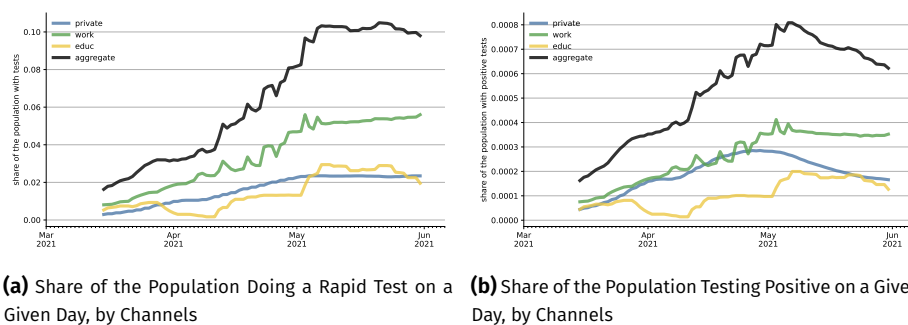
**(a)** Share of the Population Doing a Rapid Test on a Given Day, by Channels

**(b)** Share of the Population Testing Positive on a Given Day, by Channels

**Figure C.10.** Rapid Test Shares in the Population by Channel

*Note:* Rapid tests in the education setting 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. Work rapid tests 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 the time frame and the frequency of testing is also increased. Private tests are demanded by individuals for one of three reasons: having developed symptoms without access to a PCR test, having a household member that has tested positive or developed symptoms or having planned a weekly meeting with friends. Panel a shows the share of the population doing a rapid test on a given day. Panel b shows the share of the population testing positive on a given day (true and false positives).

Next, we show the tests split by whether they are true positive, false positive, true negative or false negative (see Figure C.11) in numbers per million individuals to make the metric comparable to incidences.

The number of true positives (Figure C.11a) rapidly increases and peaks at the end of April with over 200 cases per million detected through rapid tests per day. This means that our model suggests that 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 useless. It is rather the combination of large scale screening at work and in schools and very efficient follow up tests whenever those screening tests detected a case. Shapley values (Figure 3d) 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.

Such a large effect of rapid tests seems to be at odds with the general perception that they are not very reliable. However, this is not the case, which can be seen from the number of false negative tests. At its peak at the beginning of May, there are more than 80 false negative rapid tests per million inhabitants per day. While this number might sound low in isolation, it becomes quite large when set in context: We estimate that at at the the same time only 200 cases per million are detected by rapid tests. This means that even at the time where discovery of cases via rapid tests is at its peak, almost 30% of people who are infected and make a test are not discovered by the test.

> **[HM 9]**
> I do not understand. 80 fals neg for 200 true pos does not seem too reliable? I guess the important bit is the reference cat. – catching 200 pos vs. nothing
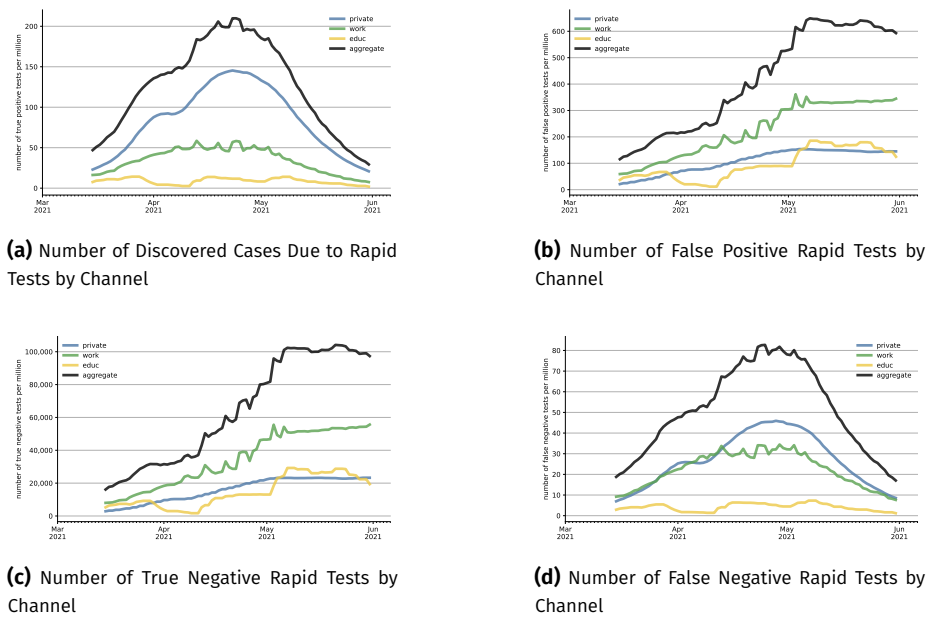
59

**(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 C.11.** Rapid Test Results

*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.

This shows clearly that the large effect of rapid tests on the infection dynamic is not driven by unrealistic assumptions about their sensitivity but rather by the fact that there was a very large number of infected individuals who did not know they are positive. Detecting and isolating some of them is enough to slow down the overall infection dynamic.

A similar picture arises, when looking at the false positive rate, i.e. the share of positive tests that go to people who are not infected. Figure C.12a 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. The false positive rate increases over time. This is due to the low prevalence of infections in the population, which falls over time. 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 low 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.
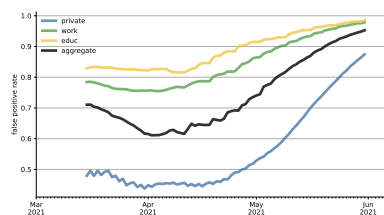
**(a)** Rate of False Positive Rapid Tests by Channel
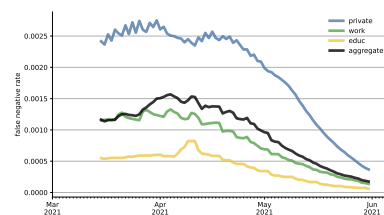


**(b)** Rate of False Negative Rapid Tests by Channel

**Figure C.12.** Rapid Test Rates by Channel

*Note:* The left panel shows 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. The right panel shows the false negative rate in the population, i.e. the share of negative test results that are mistakenly given to infected individuals. This is very low because there are many truly negative tests in times of low incidences and large scale screening tests.

## C.5   Scenarios
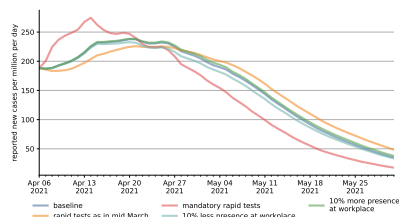
**(a)** Reported Cases
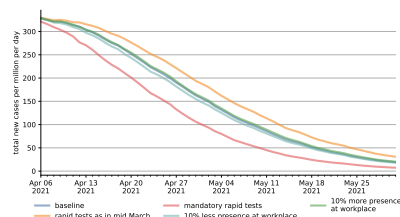


**(b)** Total Cases

**Figure C.13.** The Effect of Different Work Scenarios on Reported and Total Cases

*Note:* The figure shows the development of cases after the policy changes took place at Easter until the end of our simulation period (end of May). We vary the share of workers that work from home and how many tests are performed at work relative to our baseline scenario. Making it mandatory to test all employees that do not work from home markedly reduces cases – even when only assuming 95% compliance on both the employer and the employee side. As before, the observed cases can be misleading because more testing leads to more detected cases. It takes two to three weeks for the reduction in new infections effect to dominate the increased detection effect. Furthermore, the two opposing effects lead to a smaller effect size than is actually the case.
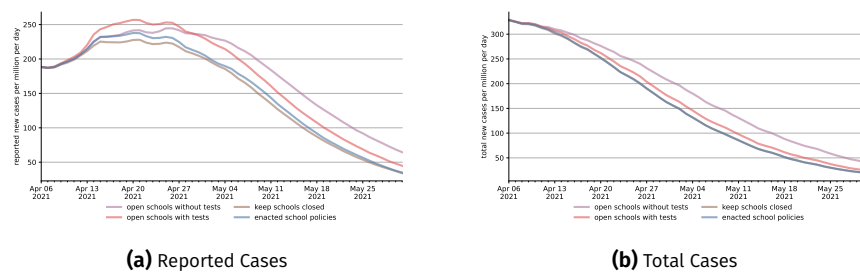
**(a)** Reported Cases



**(b)** Total Cases

**Figure C.14.** The Effect of Different School Scenarios on Reported and Total Cases

*Note:* The figure shows the development of cases after the policy changes took place at Easter until the end of our simulation period (end of May). 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 shares develop as in the baseline scenario. Our simulations suggest that the enacted policies were as effective as keeping schools closed. Opening schools with the testing schemes that were in place after Easter would have had a small effect on the overall incidence. However, this is mainly due to the stringent testing that was in place in schools by that time. Had schools opened without testing requirements the total incidence would have been up to 50 points higher, though this would have been less visible in the reported cases.
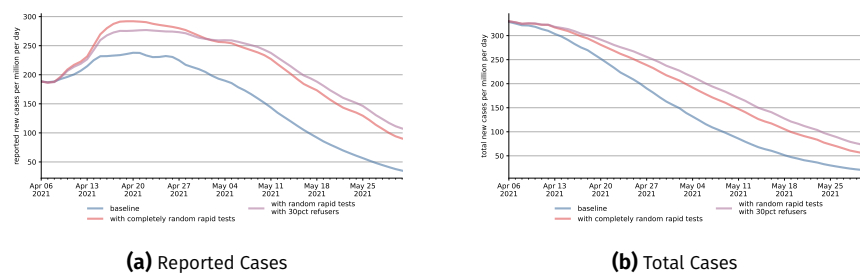


**(a)** Reported Cases



**(b)** Total Cases

**Figure C.15.** The Role of Targeted and Compliance Driven Rapid Test Demand

*Note:* To be written

## C.6   Comparison with Other Studies

- Liverpool Study
- CODAG report

**[HM 10]**

I would suggest just including this in model validation. "Another way to gauge the plausibility of our estimates is to compare our quantitative effects to evaluations of specific measures."

## References

*1.* S. Zhao, B. Tang, S. S. Musa, S. Ma, J. Zhang, M. Zeng, Q. Yun, W. Guo, Y. Zheng, Z. Yang, Z. Peng, M. K. Chong, M. Javanbakht, D. He, M. H. Wang, Estimating the generation interval and inferring the latent period of COVID-19 from the contact tracing data. *Epidemics* **36**, 100482 (2021).

2. N. G. Davies, P. Klepac, Y. Liu, K. Prem, M. Jit, R. M. Eggo, Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nature Medicine* **26**, 1205–1211 (2020).

3. C. M. Peak, R. Kahn, Y. H. Grad, L. M. Childs, R. Li, M. Lipsitch, C. O. Buckee, Individual quarantine versus active monitoring of contacts for the mitigation of COVID-19: a modelling study. *The Lancet Infectious Diseases* **20**, 1025–1033, (https://doi.org/10.1016/s1473-3099(20)30361-3) (2020).

4. X. He, E. H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, *et al.*, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature medicine* **26**, 672–675 (2020).

5. C. McAloon, Á. Collins, K. Hunt, A. Barber, A. W. Byrne, F. Butler, M. Casey, J. Griffin, E. Lane, D. McEvoy, P. Wall, M. Green, L. O'Grady, S. J. More, Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. *BMJ Open* **10**, e039652, (https://doi.org/10.1136/bmjopen-2020-039652) (2020).

6. G. Yin, H. Jin, Comparison of Transmissibility of Coronavirus Between Symptomatic and Asymptomatic Patients: Reanalysis of the Ningbo COVID-19 Data. *JMIR Public Health and Surveillance* **6**, e19464 (2020).

7. L. Zou, F. Ruan, M. Huang, L. Liang, H. Huang, Z. Hong, J. Yu, M. Kang, Y. Song, J. Xia, Q. Guo, T. Song, J. He, H.-L. Yen, M. Peiris, J. Wu, SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *New England Journal of Medicine* **382**, 1177–1179 (2020).

8. A. W. Byrne, D. McEvoy, A. B. Collins, K. Hunt, M. Casey, A. Barber, F. Butler, J. Griffin, E. A. Lane, C. McAloon, K. O'Brien, P. Wall, K. A. Walsh, S. J. More, 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**, e039856 (2020).

9. A. Singanayagam, M. Patel, A. Charlett, J. L. Bernal, V. Saliba, J. Ellis, S. Ladhani, M. Zambon, R. Gopal, Duration of infectiousness and correlation with RT-PCR cycle threshold values in cases of COVID-19, England, January to May 2020. *Eurosurveillance* **25** (2020).

10. Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, X. Liu, L. Wei, S. A. Truelove, T. Zhang, W. Gao, C. Cheng, X. Tang, X. Wu, Y. Wu, B. Sun, S. Huang, Y. Sun, J. Zhang, T. Ma, J. Lessler, T. Feng, Epidemiology and Transmission of COVID-19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. (2020).

11. E. K. Stokes, L. D. Zambrano, K. N. Anderson, E. P. Marder, K. M. Raz, S. E. B. Felix, Y. Tie, K. E. Fullerton, Coronavirus Disease 2019 Case Surveillance — United States, January 22–May 30, 2020. *MMWR. Morbidity and Mortality Weekly Report* **69**, 759–765 (2020).

12. R. Hinch, W. J. M. Probert, A. Nurtay, M. Kendall, C. Wymatt, M. Hall, K. Lythgoe, A. B. Cruz, L. Zhao, A. Stewart, L. Ferritti, D. Montero, J. Warren, N. Mather, M. Abueg, N. Wu, A. Finkelstein, D. G. Bonsall, L. Abeler-Dorner, C. Fraser, *OpenABM-Covid19 - Parameter Dictionary* (https : / / github . com / BDI - pathogens / OpenABM - Covid19 / blob / 14e24ce298ad523eba5e6f33ee9c70a90c527b3e / documentation / parameters/parameter_dictionary.md).

13. J. Chen, T. Qi, L. Liu, Y. Ling, Z. Qian, T. Li, F. Li, Q. Xu, Y. Zhang, S. Xu, Z. Song, Y. Zeng, Y. Shen, Y. Shi, T. Zhu, H. Lu, Clinical progression of patients with COVID-19 in Shanghai, China. *Journal of Infection* **80**, e1–e6 (2020).

14. K. Gaythorpe, 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, N. Ferguson, Report 8: Symptom progression of COVID-19. (2020).

15. R. Hinch, W. J. M. Probert, A. Nurtay, M. Kendall, C. Wymatt, M. Hall, K. Lythgoe, A. B. Cruz, L. Zhao, A. Stewart, L. Ferritti, D. Montero, J. Warren, N. Mather, M. Abueg, N. Wu, A. Finkelstein, D. G. Bonsall, L. Abeler-Dorner, C. Fraser, OpenABM-Covid19 - an agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. (2020).

16. Forschungsdatenzentren Der Statistischen Ämter Des Bundes Und Der Länder, *Mikrozensus 2010, CF, Version 0*, de, 2018.

17. Destatis, "Betreuungsquote von Kindern unter 6 Jahren nach Bundesländern," tech. rep. (Statistisches Bundesamt, 2020), (https://www.destatis.de/DE/ Themen / Gesellschaft - Umwelt / Soziales / Kindertagesbetreuung / Tabellen / betreuungsquote-2018.html).

18. B. Stiftung, "Ländermonitor Frühkindliche Bildungssysteme 2019," tech. rep. (Bertelsmann Stiftung, 2019), (https://www.laendermonitor.de/fileadmin/ files / Projekte / Laendermonitoring _ Fruehkindliche _ Bildungssysteme / Laendermonitor_2019/IN_Keyfacts_alle_Bundeslaender_2019_01.pdf).

19. OECD, *Bildung auf einen Blick 2013: OECD-Indikatoren*, p. 530, (https://www. oecd-ilibrary.org/content/publication/eag-2013-de).

20. J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G. S. Tomba, J. Wallinga, *et al.*, Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS medicine* **5** (2008).

21. Robert Koch-Institut, "COVID-19 Impfquoten-Monitoring in Deutschland (COVIMO): Report 5 - Fokuserhebung Impfquoten," tech. rep. (Robert Koch-Institut, 2021), (https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_ Coronavirus/Projekte_RKI/COVIMO_Reports/covimo_studie_bericht_5.pdf? __blob=publicationFile).

22. S. Vygen-Bonnet, J. Koch, C. Bogdan, T. Harder, U. Heininger, K. Kling, M. Littmann, J. Meerpohl, H. Meyer, T. Mertens, N. Schmid-Küpke, S. Scholz, M. Terhardt, M. Treskova-Schwarzbach, K. Überla, M. van der Sande, O. Wichmann, S. Wicker, U. Wiedermann, V. Wild, R. von Kries, Beschluss und Wissenschaftliche Begründung der Ständigen Impfkommission (STIKO) für die COVID-19-Impfempfehlung. 3–63 (2020).

23. L. Google, "Google COVID-19 Community Mobility Reports," tech. rep., (https://www.google.com/covid19/mobility/).

24. C. Betsch, L. Korn, L. Felgendreff, S. Eitze, P. Schmid, P. Sprengholz, L. Wieler, P. Schmich, V. Stollorz, M. Ramharter, M. Bosnjak, S. B. Omer, H. Thaiss, F. De Bock, U. Von Rüden, *COVID-19 Snapshot Monitoring (COSMO Germany) - Wave 41*, de, 2021.

25. DIHK, "Umfrage der IHK-Organisation zu Corona-Tests in den Unternehmen," tech. rep. (Deutscher Industrie- und Handelskammertag, 2021), (https://www.dihk.de/resource/blob/48068/e4fd68d36b074e63530cbe34f98e318a/startumfrage-coronatests-data.pdf).

26. E. Ahlers, M. Lübker, R. Jung, "Schleppender Start für wöchentliche Corona-Schnelltests am Arbeitsplatz: Nur 23 Prozent der Beschäftigten haben schon Zugang," tech. rep. (Wirtschafts- und Sozialwissenschaftlichen Institut, 2021), (https://www.boeckler.de/de/pressemitteilungen-2675-schleppender-start-fur-wochentliche-corona-schnelltests-am-arbeitsplatz-31982.htm).

27. Bundesministerium für Wirtschaft und Energie, *Corona-Tests in der Wirtschaft: Ergebnisse der Umfragen bei Beschäftigten und Unternehmen*, Pressemitteilung (2021; https://www.bmwi.de/Redaktion/DE/Pressemitteilungen/2021/04/20210408-corona-tests-in-der-wirtschaft-ergebnisse-umfragen.html).

28. H. Bonin, A. Krause-Pilatus, U. Rinne, "Verbreitung von Homeoffice und Corona-Tests der Arbeitgeber Ende März-Anfang April 2021, Ergebnisse einer repräsentativen Befragung von abhängig Beschäftigten," Forschungsbericht 570/3, (2021; https://www.bmas.de/SharedDocs/Downloads/DE/Publikationen/Forschungsberichte/fb-570-3-verbreitung-homeoffice-corona-tests-ende-maerz-april-2021.pdf;jsessionid=C652579EFEAB79C3DA916404FA8AAC75.delivery1-replication?__blob=publicationFile&v=3).

29. Bundesanzeiger, *Zweite Verordnung zur Änderung der SARS-CoV-2-Arbeitsschutzverordnung*, 2021, (https://www.bmas.de/SharedDocs/Downloads/DE/Gesetze/zweite-aenderungsverordnung-sars-cov-2-arbeitsschutzverordnung.pdf).

30. Bayerisches Staatsministerium für Gesundheit und Pflege, *Holetschek: Bayern hat frühzeitig Selbsttests gesichert - Schulen und Kitas bekommen schon nächste Woche 1,3 Millionen Selbsttests*, Pressemitteilung Nr. 44/GP (2021; https://www.stmgp.bayern.de/presse/holetschek-bayern-hat-fruehzeitig-selbsttests-gesichert-schulen-und-kitas-bekommen-schon/).

31. Ministerium für Kultus, Jugend und Sport Baden Württemberg, Erweiterte Teststrategie - Selbsttests an Schulen. *Pressemitteilungen*, (https://km-bw.de/,Lde/startseite/service/2021-04-08-selbsttests-an-schulen) (2021).

32. Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, *Umfassende Informationen für die Schulen zu Corona-Selbsttests für Schülerinnen und Schüler*, Pressemitteilung (2021; https://www.land.nrw/de/pressemitteilung/umfassende-informationen-fuer-die-schulen-zu-corona-selbsttests-fuer-schuelerinnen).

33. Niedersächsisches Kultusministerium, *Selbsttests an Schulen laufen an - Auslieferung der ersten 400.000 Tests beginnt heute, Testwoche startet Montag*, Pressemitteilung (2021; 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).

34. Sächsisches Staatsministerium für Kultus, *Selbsttests für Schüler und Lehrer ab Mittwoch* (2021; https://www.medienservice.sachsen.de/medien/news/248791).

35. Bayerische Staatskanzlei, *Bericht aus der Kabinettssitzung vom 7. April 2021*, Pressemitteilung (2021; https://www.bayern.de/bericht-aus-der-kabinettssitzung-vom-7-april-2021/?seite=2453).

36. Ministerium für Schule und Bildung des Landes Nordrhein-Westfalen, *Schulbetrieb im Wechselunterricht ab Montag, 19.April 2021 // Coronaselbsttests an Schulen - Testpflicht* (2021; https://www.schulministerium.nrw/14042021-schulbetrieb-im-wechselunterricht-ab-montag-19-april-2021-coronaselbsttests-schulen).

37. Ministerium für Kultus, Jugend und Sport Baden-Württemberg, *Informationen zum Schulbetrieb nach den Osterferien*, Pressemitteilung (2021; https://km-bw.de/,Lde/9112691).

38. Bundesanzeiger, *Verordnung zum Anspruch auf Testung in Bezug auf einen direkten Erregernachweis des Coronavirus SARS-CoV-2 (Coronavirus-Testverordnung - TestV)*, 2021, (https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/3_Downloads/C/Coronavirus/Verordnungen/Corona-TestV_BAnz_AT_09.03.2021_V1.pdf).

39. Presse- und Informationsamt der Bundesregierung, *Erste Zulassungen für Selbsttests* (2021; https://www.bundesregierung.de/breg-de/themen/coronavirus/zulassung-schnell-test-1861354).

40. C. Betsch, L. Korn, L. Felgendreff, S. Eitze, P. Schmid, P. Sprengholz, L. Wieler, P. Schmich, V. Stollorz, M. Ramharter, M. Bosnjak, S. B. Omer, H. Thaiss, F. De Bock, U. von Rüden, *Zusammenfassung und Empfehlungen Welle 37* (2021; https://projekte.uni-erfurt.de/cosmo2020/web/summary/37/).

41. C. Paul, A. Steinberg, C. Schaller, K. Haemmerl, L. Thum, M. Parciak, M. Knaf, S. Meyer, T. Parciak, U. Grupp, *Dunkelzifferradar*, ed. by P. Fund (https://prototypefund.de/project/dunkelzifferradar/).

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

43. P. C. Silva, P. V. Batista, H. S. Lima, M. A. Alves, F. G. Guimarães, R. C. Silva, COVID-ABS: An agent-based model of COVID-19 epidemic to simulate health and economic effects of social distancing interventions. *Chaos, Solitons & Fractals* **139**, 110088, (https://doi.org/10.1016/j.chaos.2020.110088) (2020).

44. E. Cuevas, An agent-based model to evaluate the COVID-19 transmission risks in facilities. *Computers in Biology and Medicine* **121**, 103827, (https://doi.org/10.1016/j.compbiomed.2020.103827) (2020).

45. A. Aleta, D. Martín-Corral, A. P. y Piontti, M. Ajelli, M. Litvinova, M. Chinazzi, N. E. Dean, M. E. Halloran, I. M. L. Jr, S. Merler, A. Pentland, A. Vespignani, E. Moro, Y. Moreno, Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behaviour* **4**, 964–971 (2020).

46. M. Bicher, C. Rippinger, C. Urach, D. Brunmeir, U. Siebert, N. Popper, Evaluation of Contact-Tracing Policies against the Spread of SARS-CoV-2 in Austria: An Agent-Based Simulation. *Medical Decision Making,* 0272989X2110133 (2021).

47. A. Basurto, H. Dawid, P. Harting, J. Hepp, D. Kohlweyer, Economic and Epidemic Implications of Virus Containment Policies: Insights from Agent-Based Simulations. *SSRN Electronic Journal* (2020).

48. D. Delli Gatti, S. Reissl, ABC: An Agent Based Exploration of the Macroeconomic Effects of Covid-19. *CESifo Working Paper Series*, (https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3748964#) (2020).

49. P. Mellacher, "COVID-Town: An Integrated Economic-Epidemiological Agent-Based Model," MPRA Paper 103661 (University Library of Munich, Germany, 2020), (https://ideas.repec.org/p/pra/mprapa/103661.html).

50. D. McFadden, A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society,* 995–1026 (1989).

51. C. Avery, W. Bossert, A. Clark, G. Ellison, S. F. Ellison, An Economist's Guide to Epidemiology Models of Infectious Disease. *Journal of Economic Perspectives* **34**, 79–104, (https://www.aeaweb.org/articles?id=10.1257/jep.34.4.79) (2020).

52. I. E. Isphording, M. Lipfert, N. Pestel, Does re-opening schools contribute to the spread of SARS-CoV-2? Evidence from staggered summer breaks in Germany. *Journal of Public Economics* **198**, 104426 (2021).

53. J. Gabler, T. Raabe, K. Röhrl, H.-M. v. Gaudecker, "Die Bedeutung individuellen Verhaltens über den Jahreswechsel für die Weiterentwicklung der Covid-19-Pandemie in Deutschland," tech. rep. (Institute of Labor Economics (IZA), 2020), (https://www.iza.org/publications/s/99/die-bedeutung-individuellen-verhaltens-uber-den-jahreswechsel-fur-die-weiterentwicklung-der-covid-19-pandemie-in-deutschland).

54. L. Morawska, J. W. Tang, W. Bahnfleth, P. M. Bluyssen, A. Boerstra, G. Buonanno, J. Cao, S. Dancer, A. Floto, F. Franchimon, C. Haworth, J. Hogeling, C. Isaxon, J. L. Jimenez, J. Kurnitski, Y. Li, M. Loomans, G. Marks, L. C. Marr, L. Mazzarella, A. K. Melikov, S. Miller, D. K. Milton, W. Nazaroff, P. V. Nielsen, C. Noakes, J. Peccia, X. Querol, C. Sekhar, O. Seppänen, S.-i. Tanabe, R. Tellier, K. W. Tham, P. Wargocki, A. Wierzbicka, M. Yao, How can airborne transmission of COVID-19 indoors be minimised? *Environment International* **142**, 105832 (2020).

55. E. L. Anderson, P. Turnham, J. R. Griffin, C. C. Clarke, Consideration of the Aerosol Transmission for COVID-19 and Public Health. *Risk Analysis* **40**, 902–907 (2020).

56. J. R. Donsimoni, R. Glawion, B. Plachter, K. Wälde, Projecting the spread of COVID-19 for Germany. *German Economic Review* **21**, 181–216, (https://www.degruyter.com/view/journals/ger/21/2/article-p181.xml) (2020).

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

58. M. J. Kühn, D. Abele, T. Mitra, W. Koslow, M. Abedi, K. Rack, M. Siggel, S. Khailaie, M. Klitz, S. Binder, L. Spataro, J. Gilg, J. Kleinert, M. Häberle, L. Plötzke, C. D. Spinner, M. Stecher, X. X. Zhu, A. Basermann, M. Meyer-Hermann, Assessment of effective mitigation and prediction of the spread of SARS-CoV-2 in Germany using demographic information and spatial resolution. (2020).

59. T. Gavenčiak, J. T. Monrad, G. Leech, M. Sharma, S. Mindermann, J. M. Brauner, S. Bhatt, J. Kulveit, 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, (https://www.medrxiv.org/content/early/2021/06/13/2021.06.10.21258647) (2021).

60. L. S. Shapley, *17. A value for n-person games* (Princeton University Press, 2016).

61. E. Petter, O. Mor, N. Zuckerman, D. Oz-Levi, A. Younger, D. Aran, Y. Erlich, Initial real world evidence for lower viral load of individuals who have been vaccinated by BNT162b2. (2021).

62. M. Levine-Tiefenbrun, I. Yelin, R. Katz, E. Herzel, Z. Golan, L. Schreiber, T. Wolf, V. Nadler, A. Ben-Tov, J. Kuint, S. Gazit, T. Patalon, G. Chodick, R. Kishony, Decreased SARS-CoV-2 viral load following vaccination. (2021).

63. E. Pritchard, P. C. Matthews, N. Stoesser, D. W. Eyre, O. Gethings, K.-D. Vihta, J. Jones, T. House, H. VanSteenHouse, I. Bell, J. I. Bell, J. N. Newton, J. Farrar, I. Diamond, E. Rourke, R. Studley, D. Crook, T. Peto, A. S. Walker, K. B. Pouwels, Impact of vaccination on SARS-CoV-2 cases in the community: a population-based study using the UK's COVID-19 Infection Survey. *medRxiv*, eprint: https://www.medrxiv.org/content/early/2021/04/23/2021.04.22.21255913.full.pdf, (https://www.medrxiv.org/content/early/2021/04/23/2021.04.22.21255913) (2021).

64. N. G. Davies, S. Abbott, R. C. Barnard, C. I. Jarvis, A. J. Kucharski, J. D. Munday, C. A. B. Pearson, T. W. Russell, D. C. Tully, A. D. Washburne, T. Wenseleers, A. Gimma, W. Waites, K. L. M. Wong, K. van Zandvoort, J. D. Silverman, K. Diaz-Ordaz, R. Keogh, R. M. Eggo, S. Funk, M. Jit, K. E. Atkins, W. J. Edmunds, and, Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. *Science* (2021).

65. J. Frisch, *Deutsche mit hoher Corona-Impfbereitschaft – aber Wissenslücken*, ed. by ÄrzteZeitung (https://www.aerztezeitung.de/Politik/Deutsche-mit-hoher-Corona-Impfbereitschaft-aber-Wissensluecken-417569.html).

66. L. E. Brümmer, S. Katzenschlager, M. Gaeddert, C. Erdmann, S. Schmitz, M. Bota, M. Grilli, J. Larmann, M. A. Weigand, N. R. Pollock, S. Carmona, S. Ongarello, J. Sacks, C. M. Denkinger, The accuracy of novel antigen rapid diagnostics for SARS-CoV-2: a living systematic review and meta-analysis. (2021).

67. R. L. Smith, L. L. Gibson, P. P. Martinez, R. Ke, A. Mirza, M. Conte, N. Gallagher, A. Conte, L. Wang, R. Fredrickson, D. C. Edmonson, M. E. Baughman, K. K. Chiu, H. Choi, T. W. Jensen, K. R. Scardina, S. Bradley, S. L. Gloss, C. Reinhart, J. Yedetore, A. N. Owens, J. Broach, B. Barton, P. Lazar, D. Henness, T. Young, A. Dunnett, M. L. Robinson, H. H. Mostafa, A. Pekosz, Y. C. Manabe, W. J. Heetderks, D. D. McManus, C. B. Brooke, Longitudinal assessment of diagnostic test performance over the course of acute SARS-CoV-2 infection. *The Journal of Infectious Diseases* (2021).

68. Robert Koch-Institut, *Laborbasierte Surveillance SARS-CoV-2* (https://ars.rki.de/Content/COVID19/Main.aspx).

69. C. Betsch, L. Wieler, M. Bosnjak, M. Ramharter, V. Stollorz, S. Omer, L. Korn, P. Sprengholz, L. Felgendreff, S. Eitze, *et al.*, Germany COVID-19 Snapshot MOnitoring (COSMO Germany): Rapid Tests. (https://projekte.uni-erfurt.de/cosmo2020/web/topic/wissen-verhalten/80-schnelltests/#best%C3%A4tigung-von-schnelltests-durch-pcr-tests-stand-23.03.21) (2021).