Federal Reserve Bank of Dallas

Social Distancing, Vaccination and Evolution of COVID-19 Transmission Rates in Europe

Alexander Chudik, M. Hashem Pesaran and Alessandro Rebucci

Globalization Institute Working Paper 414 **February 2022**

Research Department

https://doi.org/10.24149/gwp414

Social Distancing, Vaccination and Evolution of COVID-19 Transmission Rates in Europe*

Alexander Chudik†, M. Hashem Pesaran‡ and Alessandro Rebucci§

January 29, 2022

Abstract

This paper provides estimates of COVID-19 effective reproduction numbers worldwide and explains their evolution for selected European countries since the start of the pandemic taking account of changes in voluntary and government mandated social distancing, incentives to comply, vaccination and the emergence of mutations. Evidence based on panel data modeling indicates that the diversity of outcomes that we document resulted from the non-linear interaction of mandated and voluntary social distancing and the economic incentives that governments provided to support isolation, with no one factor independently capable of lowering the reproduction number below one. However, the importance of these factors declined over time, with vaccine uptake driving heterogeneity in country experiences in 2021. Our approach also allows us to identify the basic reproduction number, $\mathcal{R}_0$, and how it changes with mutations. It is precisely estimated and differs little across countries.

Keywords: COVID-19, multiplication factor, under-reporting, social distancing, self-isolation, SIR model, reproduction number, pandemics, vaccine.

JEL Classification: D0, F60, C4, I120, E7.

* We thank Johns Hopkins University for assistance with the data. We are grateful to Anton Korinek, Ippei Fujiwara, Piero La Mura, Ron Smith, Zheng (Michael) Song, and Cynthia Yang for helpful comments and discussions on previous drafts of this paper. We also thank Daniel Jimenez for research assistance. The views in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System. Any errors or omissions are the responsibility of the authors.

† Alexander Chudik, Federal Reserve Bank of Dallas, alexander.chudik@gmail.com.

‡ M. Hashem Pesaran, University of Southern California, USA, and Trinity College, Cambridge, UK, pesaran@usc.edu.

§ Alessandro Rebucci, Johns Hopkins University Carey Business School, CEPR and NBER, arebucci@jhu.edu.

1 Introduction

The COVID-19 pandemic has claimed millions of lives and brought about very costly government interventions to contain it, with unprecedented and widespread economic disruption worldwide. China responded to the outbreak with strict policies, including binding and mandatory social distancing. These policies have proved successful in containing the epidemic so far, but there is a considerable uncertainty about the future course of the epidemic in China. At the other end of the spectrum, for example, Sweden initially attempted to let its epidemic run its course with only minimal interventions from the government. Other countries responded by adopting a mixture of policies, either by deliberate choice or due to popular opposition to the implementation of lock-downs or even milder forms of social distancing. What drives heterogeneity in outcomes? As the COVID-19 pandemic evolved into a an endemic infection disease in the midst of heated debates on the pros and cons of social distancing, what lessons do we learn for the management of future epidemics?

In this paper, we contribute towards answer some of these questions by providing estimates of COVID-19 effective reproduction numbers worldwide and explain their evolution for selected European countries since the start of the pandemic in early 2020, taking account of changes in voluntary and government mandated social distancing, economic support to comply with isolation, vaccination, and the emergence of new mutations. Although the European countries we focus on had similar patterns of transmissions at the beginning of the pandemic, they end up with different outcomes. We show that neither social distancing nor incentives to comply, or people's voluntary responses to the epidemic can single-handedly account for the heterogeneity that we document over the period 2020-2021. Vaccine uptake, however, becomes the dominant factor accounting for country differences in outcomes in 2021.

Reproduction numbers are epidemiologic metrics to measure the spread of an infectious disease. The basic reproduction number, denoted by $\mathcal{R}_0$, is the number of new infections expected to result from one infected individual at the start of the epidemic, assuming no interventions. Within a classical susceptible-infective-removed (SIR) model the basic reproduction number is given by

$\mathcal{R}_0 = \beta_0/\gamma$, where β_0 is the initial (biological) transmission rate, and γ is the recovery rate. Since the transmissibility of a disease will vary over time due to changes in immunity, mitigation policies, or precautionary behavior, the effective reproduction number, which we denote by $\mathcal{R}_{et}$, measures the $\mathcal{R}$ number t periods after the initial outbreak. As we show in the paper, in the classical SIR model, we have $\mathcal{R}_{et} = (1 - c_t)\,\beta_t/\gamma$, where c_t is the per capita number of infected cases at time t. As a result, one can separate changes in $\mathcal{R}_{et}$ due the extent to which the susceptible population is shrinking, $1-c_t$ (which we call herding), or due to changes in the transmission rate, β_t/γ. Moreover, in our framework, for a given value of $\mathcal{R}_0$, social distancing, either voluntary due to precautionary behavior or mandatory due to non-pharmaceutical interventions (which we also call mitigation or containment policies for brevity), compliance with mandated measures, and immunity changes due to virus mutations or vaccination lead to time-variation in the transmission rate, β_t, driven by changes in contacts or susceptibility to infection or changes in immunity.

To estimate the COVID-19 time-varying transmission rates, β_t, we extend a moment condition derived by Pesaran and Yang (2021), henceforth PY. The framework developed by PY is an agent-based stochastic network model of epidemic diffusion, which aggregates up to the standard deterministic SIR model. We extend the model to allow for time variations in β_t, showing that approximately β_t can be written as $\tau_t k_t/\mu_t$, where k_t is the mean number of contacts during day t, τ_t is the mean exposure intensity to the virus, and $\mu_t \geq 1$ is the mean immunity level in the population. We then exploit time series and cross section variations in case data, adjusting for misreporting, to quantify the relative importance of these factors in the epidemic diffusion.

The estimation approach that we undertake is applicable to any level of jurisdiction and could provide guidance on how to measure the health impact in causal studies of specific mitigating policies. For the sake of brevity, in the paper, we only report the results for selected European countries, but compute rolling estimates of effective reproduction numbers and transmission rates for all jurisdictions for which Johns Hopkins University (JHU) reports case statistics.[1] This is important since measuring health outcomes is challenging in studies that seek to establish causal effects of policies to address COVID-19.

[1]The estimation results for all countries and regions that we compute are available on the authors' websites (sites.google.com/site/alexanderchudik/, pesaran.com, sites.google.com/site/alessandrorebucciphd/).

Our method of moment estimation requires only data on infected cases, thus complementing estimation methods based on death statistics. Both the reported number of infected cases and deaths is problematic and different countries might have better quality data on either one or the other. For example, Spain has very good death statistics. In other countries, death statistics have undergone major revisions on several occasions. For example, the United Kingdom death toll was revised downward by 5,377 on August 12, 2020 after a review concluded that daily death figures should only include deaths which had occurred within 28 days of a positive COVID-19 test.

Our estimation method is not only simple to apply, but also robust to the measurement error and under-reporting of infected cases. Following the medical evidence in Gibbons et al. (2014), we use a multiplication factor to allow for under-reporting in a way that will be elaborated in later sections of the paper.

To understand the factors behind the heterogeneous evolution of effective reproduction numbers in Europe, we separate the herding component from the *transmission rate* and empirically model the latter. In line with the agent-based stochastic epidemic network model of PY, we argue that it is the transmission rate, and not the $\mathcal{R}$ number, that depends on behavioral changes. As noted earlier, our time-varying parameter model shows that the transmission rate depends on the average number contacts *multiplied* by individual-specific susceptibility to become infected which in turn depends on average duration of contacts, wearing of face masks, and other recommended precautions, as well as the average degree of immunity in the population. Accordingly, in our empirical analysis we assume that a country's time-varying country-specific transmission rate depends on five factors: mandatory and voluntary social distancing, economic support to comply with mandatory polices, vaccine uptake and virus mutations.

Consistent with a simple decision theoretic model presented in the paper and a large body of empirical evidence, we distinguish between government-mandated social distancing policies and voluntary self-isolation. We also control for government economic support that affects the incentive to comply as reported in survey data (e.g., Papageorge et al., 2021; Hamermesh, 2020). To measure mandated-social distancing and incentives to comply with these policies we use the stringency and

support indices compiled by Oxford COVID-19 Government Response Tracker project.[2] To assess the impact of voluntary social distancing, we allow for a threshold effect capturing the impact of fear of becoming infected arising from news of rising cases on individual precautionary behavior. The importance of these factors in controlling the effective rate of transmission is jointly estimated within the context of the epidemic model, allowing for a lag of two or three weeks between the policy or behavioral changes and the infection outcomes. In 2021, virus mutations and vaccine availability and uptake became more prominent. Thus, we add to the baseline model the population share of vaccinated people and the Delta share of the confirmed sequenced cases, and we re estimate the model through November 2021.

Throughout the empirical analysis, we focus on nine European countries with a similar start to the outbreaks in March 2020, but with differing outcomes subsequently. In a pre-vaccine sample, we find that all determinants of the transmission rate are statistically highly significant and have the expected signs. Consistent with the heterogeneity documented in the first part of the empirical analysis, we find that when we control for voluntary behavior by adding the threshold effect, the magnitudes of estimated coefficients on mandatory social distancing and compliance indicators decline markedly. Our baseline results in the pre-vaccine sample thus suggest that these factors must be aligned to bring down $\mathcal{R}_{et}$ below one over a sustained period and none of them can stabilize it single-handedly. However, in the post-vaccine sample ending in November 2021, we find that vaccine uptake is the most important contributor to the decline in the effective transmission rate. In both samples, we estimate model-consistent basic reproduction numbers close to 5 for all the European countries that we consider.

In conclusion, to summarize the main message, our empirical analysis shows that mandatory and voluntary social distancing and incentives to comply were critical, but neither was alone capable of bringing the reproduction number below one alone. Similarly, vaccine uptake in 2021 is estimated to be the most important factor contributing to the reduction in effective transmission rates, but the other factors continued to be salient. As a result, epidemic curves show a great deal of heterogeneity and no silver bullet could have defeated COVID-19.

[2]Available at https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker.

Related Literature A very large body of research investigates the COVID-19 outbreak and the policies to contain its spread.[3] For example, Fang, Wang, and Yang (2020) analyze efforts to contain the COVID-19 outbreak in China, measuring the effectiveness of the lock-down of Wuhan and showing that these policies also contributed significantly to reducing the total number of infections also outside of Wuhan. Similarly, there is ample reduced form evidence on the impact of mandatory social distancing using state and county level data in the case of the United States, and for a few other countries. However, there are not many studies on the relative importance of mandatory and voluntary social distancing, especially for large cross sections of countries.

Caselli et al. (2020) find that both lock-downs and voluntary social distancing helped contain the first wave of COVID-19, but mandatory interventions have been critical. Jinjarak et al. (2020) find that more stringent policies are associated with lower mortality growth rates in a large cross section of countries but with some heterogeneity depending on demographics, the degree of urbanization and political freedom, as well as the international travel flows. In general, however, countries with more stringent policies at the onset of the epidemic realized lower peak mortality rates and exhibited lower duration during the first epidemic wave. We distinguish not only between mandatory and voluntary social distancing, but also consider the incentives to comply and herd immunity in lowering the reproduction number, as well as vaccine uptake and mutations. To our knowledge, no study which considers voluntary or government-mandated social distancing also controls for the possibility of herd immunity and distinguishes its impacts on effective reproduction numbers from the influence of policy and/or behavioral factors.

A number of studies consider the effects of different intervention strategies – such as isolating the elderly, closing schools and/or workplaces, and alternating work/school schedules – which should lower the average number of contacts of specific age groups, contact locations, or time windows relative to normal (pre-COVID) patterns using calibrated behavioral SIR or compartmental models.[4] We take an empirical/econometrics approach calibrating only the recovery rate, γ; a parameter on

[3] See Brodeur et al. (2020), Gupta, Simon, and Wing (2020), and Avery et al. (2020) for surveys of the early literature through the end 2020.

[4] See for example Acemoglu et al. (2020), Akbarpour et al. (2020), Alvarez et al. (2021), Atkeson et al. (2020b), Cakmakli et al. (2021), Cakmakli et al. (2020), Chudik et al. (2020), Favero et al. (2021), Matrajt and Leung (2020), and Toda (2020), among many others.

which we have much more precise clinical information.

Various methods are available in the epidemiological literature to estimate the reproduction numbers at the beginning and/or in real time during epidemics. Estimation of reproduction numbers based on different models are reviewed by Chowell and Nishiura (2008), Obadia, Haneef, and Boëlle (2012), and Nikbakht et al. (2019). More recent contributions, focusing on estimation of reproduction numbers for the COVID-19 pandemic based on death statistics include Atkeson, Kopecky, and Zha (2020b), Baqaee et al. (2020), Korolev (2020) and Toda (2020). As we noted earlier both death and case statistics are problematic, and we provide complementary evidence relying only on case statistics correcting for measurement errors and under-reporting. Other closely related papers are Fernández-Villaverde and Jones (2020), Atkeson, Kopecky, and Zha (2020a), and Cakmakli and Simsek (2020).

The rest of the paper is organized as follows. Section 2 discusses SIR model with time-varying transmission rates. Section 3 presents the method and estimation results for the transmission rates and reproduction numbers for selected European countries. Section 4 sets up the panel data model to assess the the relative importance of the factors driving the effective transmission rate over time and across countries and reports the estimation results. Section 5 concludes. Appendix presents an extension of PY model to allow for time variations in key parameters affecting the transmission of the virus (the mean daily contact numbers, mean exposure intensity, and latent immunity and vaccination). Estimates of transmission rates for key countries and regions globally are provided in an online supplement, with additional robustness evidence for the main results of the paper.

2 A SIR model with time-varying transmission rate

There are many approaches to modelling the spread of epidemics. The basic mathematical model widely used by researchers is the susceptible-infected-removed (SIR) model advanced by Kermack and McKendrick (1927). This model and its various extensions have been the subject of a vast number of studies, and have been applied extensively to investigate the spread of COVID-19.[5]

[5] A comprehensive treatment is provided by Diekmann and Heesterbeek (2000) with further contributions by Metz (1978), Satsuma et al. (2004), Harko et al. (2014), Salje et al. (2016), amongst many others.

The basic SIR model considers a given population of fixed size n, composed of three distinct groups, those individuals in period t who have not yet contracted the disease and are therefore susceptible, denoted by S_t; the 'removed' individuals who can no longer contract the disease, consisting of recovered and deceased, denoted by R_t; and those who remain infected at time t and denoted by I_t. Thus,

$$n = S_t + I_t + R_t. \tag{1}$$

As it stands, this is an accounting identity, and it is therefore sufficient to model two of the three variables (S_t, I_t, and R_t) to obtain the third as the remainder.

The classic SIR model is deterministic. It is cast in the following set of difference equations (for $t = 1, 2, ..., T$)

$$\begin{aligned} S_{t+1} - S_t &= -S_t \beta I_t, & (2) \\ I_{t+1} - I_t &= S_t \beta I_t - \gamma I_t, & (3) \\ R_{t+1} - R_t &= \gamma I_t. & (4) \end{aligned}$$

The parameter β is the rate of transmission, while γ is the recovery rate. For a given non-zero initial values S_1 and I_1, and the parameter values for β and γ, the evolution of the number of infected and recovered individuals is deterministic and given by the recursive solution of (2)-(4) for given initial values.

The evolution of the epidemic crucially depends on the two key parameters β and γ. It is easy to see from equation (2)-(4) that, without any mitigating intervention, the epidemic will spread if $\beta/\gamma = \mathcal{R}_0 > 1$ and will cease only after infecting $(\mathcal{R}_0 - 1)/\mathcal{R}_0$ of the population. The parameter $\mathcal{R}_0$ is the basic reproduction number, defined as "*the average number of secondary cases produced by one infected individual during the infected individual's entire infectious period assuming a fully susceptible population*" (Del Valle et al., 2013). The terminal condition $(\mathcal{R}_0 - 1)/\mathcal{R}_0$ is the herd immunity threshold. In the case of COVID-19, a number of different estimates have been suggested in the literature, initially placing $\mathcal{R}_0$ somewhere in the range of 2.4 to 3.9, with even larger numbers

for more recent variants of concern such as Delta and Omicorn.[6] So, the classical SIR model predicts that in the absence of intervention as much as 2/3-3/4 of the population could eventually become infected before herd immunity is reached. Our panel estimates reported below suggest $\mathcal{R}_0$ in excess of 5 in Europe at the beginning of the pandemic, requiring as much as 4/5 of the population to achieve immunity before COVID-19 stops from spreading.

This well understood possibility triggered unparalleled mitigation and containment interventions, first by China and South Korea, then Europe, the US and all other countries around the world. Such interventions, which broadly speaking we refer to as "social distancing" include case isolation, mandated face mask wearing, banning of gatherings, closures of schools and universities, and even local and national lock-downs; all aimed at slowing down the transmission rate of the virus. It is clear that these policies, together with voluntary changes in behavior in response to the epidemic, make it harder for the virus to transmit between individuals. To isolate and estimate the effects of different factors on the spread of the virus we treat the transmission rate as time-varying, i.e., β_t, and relate it to voluntary and mandatory social distancing, the provision of economic incentives to isolate, as well as vaccination and the emergence of new variants. However, given the clinical evidence discussed below, we assume the recovery rate γ to be fixed over time. We will refer to β_t/γ as the "effective transmission rate".

The standard SIR model (2)-(4) takes the transmission rate as given. In this paper, we extend the richer, agent-based stochastic network model proposed by Pesaran and Yang (2021), henceforth PY, which yields the standard SIR model above as an approximation for a large population of individuals. In this framework, outlined in the Appendix, the infection of individual i from a population of n susceptible individuals is modelled using a latent variable $x^*_{i,t+1}$, as $I\left(x^*_{i,t+1} > 0\right)$, where $I(\mathcal{A})$ is the indicator function that takes the value of unity if $\mathcal{A}$ holds and zero otherwise.

[6] For example, using data from Wuhan, Wang et al. (2020) report a pre-intervention reproductive rate of 3.86; Kucharski et al. (2020) estimate that, in China, the reproductive rate was 2.35 one week before travel restrictions were imposed on Jan 23, 2020. Ferguson et al. (2020) made the baseline assumption of $R_0 = 2.4$ and also examined values of 2.0 and 2.6 based on fitting the early growth-rate of the epidemic in Wuhan by Li et al. (2020) and Riou and Althaus (2020).

The latent variable, $x^*_{i,t+1}$, is given by

$$x^*_{i,t+1} = \tau_{it} \sum_{j=1}^{n} d_{ij}(t) z_{jt} - \mu_{it}\xi_{i,t+1}, \tag{5}$$

in which the first component, $\tau_{it} \sum_{j=1}^{n} d_{ij}(t) z_{jt}$, captures to the contact pattern of individual i with actively infected (contagious) individuals, denoted by the 0/1 variable z_{jt}, using the stochastic contact matrix, $\mathbf{D}(t) = [d_{ij}(t)]$, where $d_{ij}(t) = 1$ with probability k_{it}/n, and 0, otherwsie, and k_{it} is the average number of contacts of indiviudal i in day t. The coefficient $\tau_{it} > 0$ captures the degree (or intensity) of the exposure of individual i in day t to the virus upon contacts. For example, two individuals with the same contact patterns have different infection probability in day $t+1$, if they follow different types of precautions as to their mask wearing and hygiene habits. Finally, the term $\mu_{it}\xi_{i,t+1} > 0$, represents individual i^{th} immunity to becoming infected, which could depend on whether the individual is vaccinated. Two individual with the same contact patters, and mask wearing habits could have different infection probabilities due to having different levels of immunity, either natural or vaccine induced. Social distancing, be it mandated and/or voluntary, mask wearing, frequent hand washing, and other mitigating policies will decrease the transmission rate by decreasing the mean contacts per day, k_t, and/or reducing the mean exposure intensity parameter τ_t in (6). Vaccinations, on the other hand, are expected to decrease transmission rate, by increasing individual-specific resistance to the virus, which manifests in an increase in μ_t. Virus mutations can increase the transmission rate by affecting both μ_t and τ_t. PY consider a more elaborate set up where they allow for group heterogeneity and provide an individual based model of recovery, but focus on simulation properties of their model assuming fixed values of $\tau_{it} = \tau$, and set immunity coefficient, μ_{it}, to unity, since τ_{it} and μ_{it} are not separately identified.

Given the focus of our empirical analysis on the determinants of β_t, we allow for time variations in all the key parameters and approximate β_t by

$$\beta_t \approx \frac{k_t \tau_t}{\mu_t}, \tag{6}$$

where k_t is the average number of contacts during day t, τ_t is the average exposure intensity of the susceptible population to the virus, and $\mu_t \geq 1$ is the mean immunity level in the population. We exploit time series and cross section variations in factors that influence k_{jt}, τ_{jt} and μ_{jt}, across countries $j = 1, 2, ..., m$ and over time $t = 1, 2, ..., T$ (days) . We also relate β_t to the average effective reproduction number, $\mathcal{R}_{et}$, defined as the expected number of secondary cases produced by one infected individual in a population that includes *both* susceptible and non-susceptible individuals at time t. $\mathcal{R}_{et}$ depends on the effective transmission rate (β_t/γ), and the share of susceptible population $(s_t = S_t/n = 1 - c_t)$, and is given by

$$\mathcal{R}_{et} = \left(\frac{\beta_t}{\gamma}\right) s_t = \left(\frac{\beta_t}{\gamma}\right) (1 - c_t) , \tag{7}$$

where $c_t = 1 - s_t$ is the fraction of population that has been infected (cumulation of new infected cases), and $1 - c_t$ is the herd-immunity component of $\mathcal{R}_{et}$. It is also worth bearing in mind that at the outset of epidemic outbreak, assuming a fully susceptible population, we have $s_0 = 1$ $(c_0 = 0)$, which in turn ensures that $\mathcal{R}_{e0} = \beta_0/\gamma = \mathcal{R}_0$.

As the epidemic evolves, the average number of secondary cases caused by a single infected individual will vary over time as a result of decline in the number of susceptible individuals (due to immunity or death, vaccine protocols or virus mutations) and/or changes in behavior due to social distancing. In our empirical analysis, we will first provide country-specific estimates of the effective transmission rate. We then model their evolution for a selected number of European countries with similar initial epidemic trajectories.

3 Estimating time-varying transmission rates

PY show that the classic aggregate SIR model (2)-(4) with time-invariant transmission rate can be obtained as an approximation (for a large population n) to the individual-based stochastic network model of epidemic that we spell out in appendix, where individuals randomly interact with each

other.[7] In order to estimate country-specific time-varying transmission rates, β_t, we utilize the following moment condition, modified to allow for time-variation in the transmission rate (see equation (48) of PY, or equation (A.9) in Appendix):

$$E_t\left(\frac{1-c_{t+1}}{1-c_t}|\ i_t\right) = e^{-\beta_t i_t} + O(n^{-1}), \tag{8}$$

where $i_t = I_t/n$ is the per capita number of infected individuals. Since $\beta_t i_t$ is typically close to zero and n quite large (in millions) we have**:** (noting also that $ln(1+x) \approx x$ for a sufficiently small x)

$$\frac{\Delta c_{t+1}}{\gamma\,(1-c_t)} \approx (\beta_t/\gamma)\,i_t + v_{t+1}, \tag{9}$$

where $E\,(v_{t+1}\,|i_t\,) = 0$. (9) is in line with (2). The evolution of i_t and $r_t = R_t/n$ are governed as in the standard SIR model by (1) and (4), namely $i_t = (1-s_t) - r_t = c_t - r_t$, and $r_t = r_{t-1} + \gamma i_{t-1} = r_{t-1} + \gamma\,(c_{t-1} - r_{t-1})$, and since $0 < \gamma < 1$ then

$$i_t = c_t - \gamma \sum_{\ell=0}^{\infty} (1-\gamma)^{\ell} c_{t-1-\ell}, \tag{10}$$

which can be approximated well using current and past values of c_t, since $c_t = 0$, for dates before the start of the epidemic.

In principle, γ can be estimated using time series data on r_t and i_t. However, data on recoveries, R_t, either do not exist or are unreliable due to considerable measurement difficulties. For example, in Europe the recorded data on recoveries are unavailable for Spain and UK; they are of poor quality for France and Italy; and they are relatively close to our estimated recovery for Austria and Germany. To overcome this problem, we use (10) to impute data on i_t assuming a recovery rate $\gamma = 1/14$. We obtain very similar results if we use $\gamma = 1/21$.[8] The choice $\gamma = 1/14$ is consistent with the assumptions made in designing quarantine policies based on clinical evidence and also

[7] Specifically, stochastic simulation results obtained by PY show that a single group model provides a good approximation to a multi-group alternative.

[8] An alternative, pursued for example by Fernández-Villaverde and Jones (2020), is to rely only on death data. While some countries might have good death statistics, using COVID-19 death data pose challenges similar to those raised by cases. The use of death data also has the added disadvantage of being a lagging indicator and could differ across countries due to factors such as age composition, obesity, and the quality of care system.

used in calibrated behavioral epidemic models.[9]

Another data issue surrounds the measurement of confirmed cases, which are likely to be under-reported, in part due to fact that a non-negligible portion (perhaps about a half) of the cases is asymptomatic and therefore unlikely to be detected without large-scale testing. To mitigate the problem of under-reporting, we follow the epidemiological literature (see, for example, Gibbons et al., 2014) and assume that the magnitude of under-reporting is measured by the multiplication factor (MF - the ratio of true cases to reported cases). Denoting the observed values of c_t and i_t by $\tilde{c}_t$ and $\tilde{\imath}_t$, we have

$$c_t = \text{MF}\tilde{c}_t \quad \text{and} \quad i_t = \text{MF}\tilde{\imath}_t.$$

Then the moment condition in terms of observed values ($\tilde{c}_t$ and $\tilde{\imath}_t$) can be written as

$$E_t\left(\frac{1-\text{MF}\tilde{c}_{t+1}}{1-\text{MF}\ \tilde{c}_t}|\ \tilde{\imath}_t, \tilde{c}_t\right) = e^{-\beta_t \text{MF}\ \tilde{\imath}_t}. \tag{11}$$

We do not know the true MF. We allow for the multiplication factor to be larger than one. We abstract from time variation in MF, and note that the magnitude of MF will not matter in the initial stage of the pandemic, when only a relatively small fraction of the population has been infected. Estimation results in PY suggest substantial underreporting of cases in the early stage of the pandemic, with MF values declining to 2-3 at the end of their sample (April 2021) for Austria, Germany and Italy and 2-4 in the case of France and UK. Since the degree of undderreporting does not matter for the estimation of transmission rate in the early stages of pandemic, we use $MF = 3$ in figures reported below (a rather conservative value for the end of the sample, where herd immunity plays the most important role).[10] In our panel estimation results for a sample of European countries in Section 4, we report estimates for $MF = 2$ and 3; both choices provide

[9] See, for example, the medical evidence documented in Ferguson et al. (2020) which implies a value for γ in the range 0.048 to 0.071. Our results are robust to assuming $\gamma = 1/21$.

[10] In the online Appendix, we compare these results with the ones obtained for $MF = 5$. Estimates of the reproduction numbers are not sensitive to the choice of MF, whereas estimates of the effective transmission rate can be sensitive to the choice of MF towards the end of the sample in those countries where the reported share of infected population is relatively large.

similar results.[11]

One additional challenge is that the reported daily data are subject to weekly distortions (e.g., the reported number of cases on Sundays is usually lower compared with the infected cases reported for other days). To deal with this calendar distortion, as is common practice, we take seven-day moving averages of the reported data used in estimation. But again we note that our results are robust if we use reported daily cases without averaging. Using the moment condition (11), we compute rolling-window estimates of the transmission rate as

$$\hat{\beta}_t\left(W,\mathrm{MF}\right)=\mathrm{Argmin}_{\beta}\sum_{\tau=t-W+1}^{t}\left(\frac{1-\mathrm{MF}\tilde{c}_{\tau}}{1-\mathrm{MF}\tilde{c}_{\tau-1}}-e^{-\beta\mathrm{MF}\ \tilde{\imath}_{\tau-1}}\right)^{2}, \tag{12}$$

where W is the rolling window size, which we set to 14 days.[12]

3.1 Estimated transmission rates and reproduction numbers in Europe

Daily data in case numbers, and rolling estimates of effective transmission rates and reproduction numbers for selected European countries (Belgium, France, Germany, Italy, Netherlands, Poland, Portugal, Spain, and the United Kingdom) are displayed in Figures 1-2.[13] The virus outbreak in continental Europe begins with Italy in early 2020, with the recorded number of infections accelerating rapidly from February 21, 2020 onward. A rapid rise in infections takes place about one week later in Spain, Germany and France. As the rolling estimates show, the $\mathcal{R}$ number fell below one in mid- to late-April in all these countries. As lock-downs were eased during the summer, however, the transmission rates started to rise again. By the end of the 2020, the $\mathcal{R}$ numbers were much more dispersed, with some countries doing better than others. All large European countries

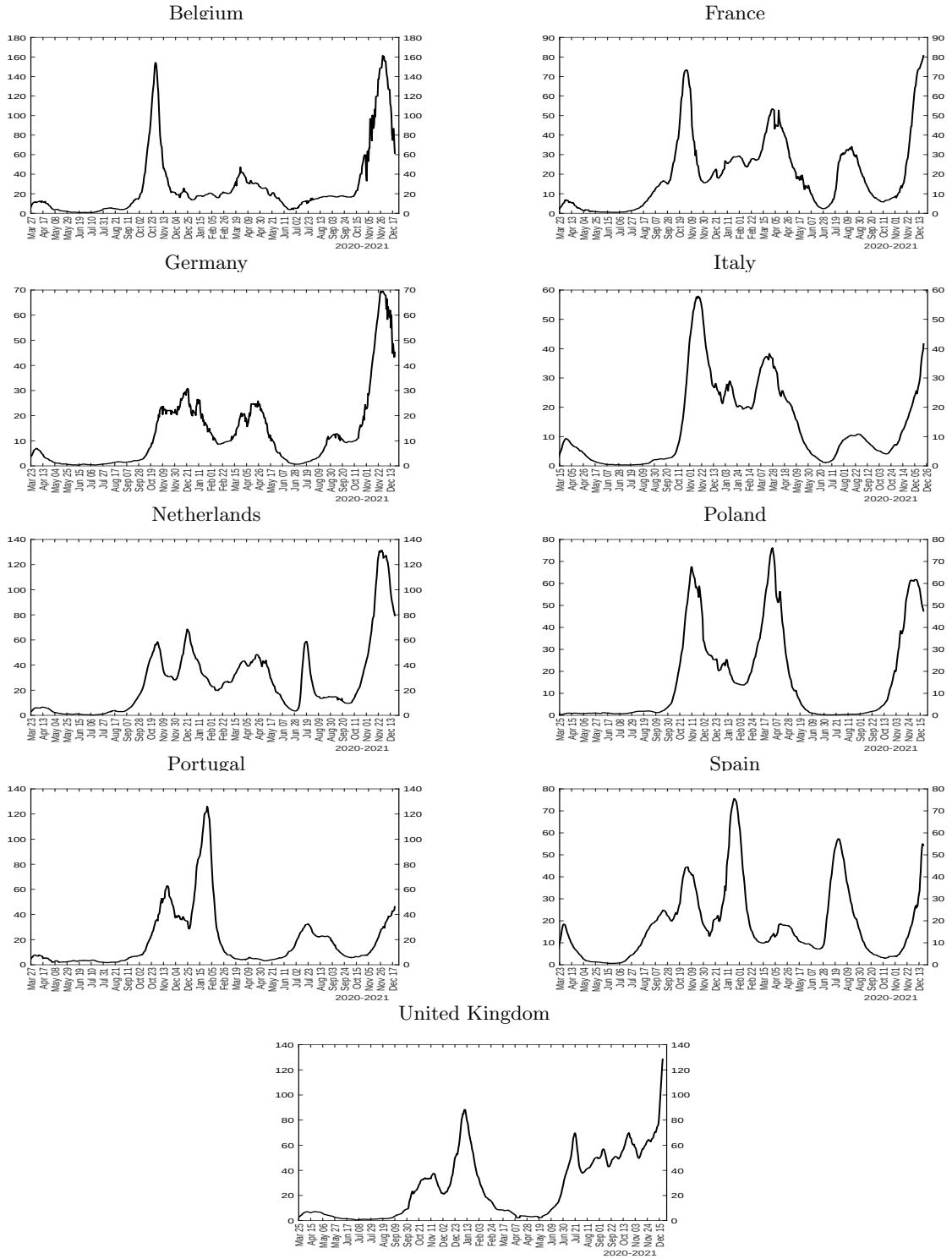
[11] The data from the Diamond Princess cruise ship reported by Moriarty et al. (2020) suggest about half of the COVID-19 cases are asymptomatic, and therefore $MF = 2$ seems to be a good lower bound.

[12] The COVID-19 data for all countries except France and Spain are sourced from the repository of the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU) –available at https://github.com/CSSEGISandData/COVID-19. We use World Health Organisation COVID-19 database (available at https://covid19.who.int/WHO-COVID-19-global-data.csv) for France and Spain due to errors in the JHU data. The population data (for year 2019) are obtained from the World Bank database, available at https://data.worldbank.org/indicator/SP.POP.TOTL.

[13] In the accompanying online supplement, we provide country specific estimates with a global coverage - providing plots for major economies and regional aggregates. Matlab codes for the estimation of transmission rates available to download from authors' websites at sites.google.com/site/alexanderchudik/, pesaran.com, and sites.google.com/site/alessandrorebucciphd/.

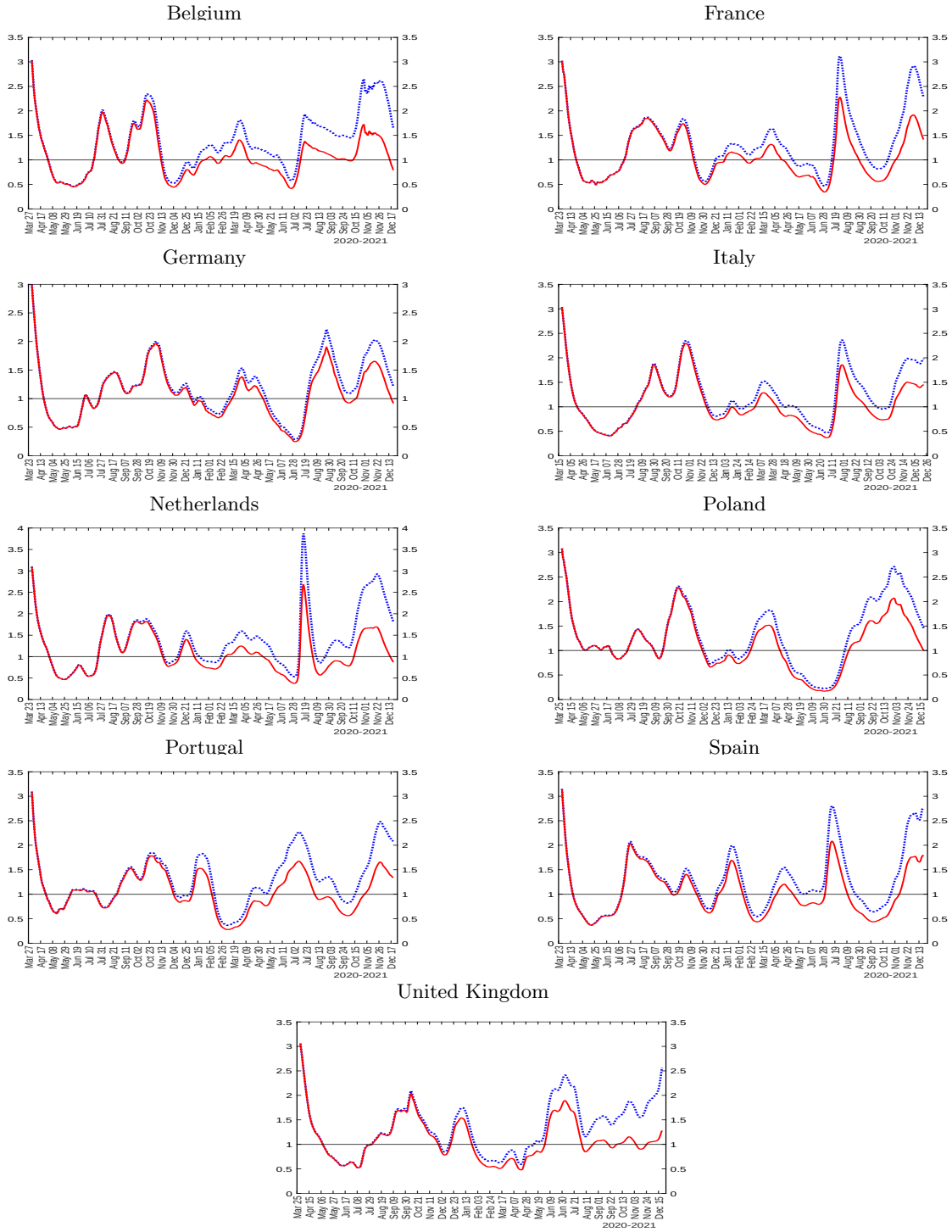
show a second wave much larger than the first one. The United Kingdom, Spain, Portugal and Netherlands exhibit distinct third waves, with larger case counts compared with their second-waves. The distinction between the effective transmission rate and the reproduction numbers in Figure 2 permits assessing the influence of herding, which become more salient toward the end of the sample period.

Figure 1: New cases for selected European countries



Notes: The figure plots seven-day moving average of the number of reported daily new confirmed cases per 100k population.

Figure 2: Estimated transmission rates and reproduction numbers for selected European countries



Notes: The figure plots the reproduction number, $\hat{\mathcal{R}}_{et}$ (solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (dotted blue line). $\hat{\mathcal{R}}_{et} = (1 - \text{MF}\tilde{c}_t)\,\hat{\beta}_t/\gamma$, where $\gamma = 1/14$, and MF=3 for each country. $\hat{\beta}_t$ is estimated using (12), where the number of active infections is computed using the data on confirmed cases minus imputed removed cases. The number of removed (recoveries + deaths) is imputed recursively using $R_t = (1-\gamma)\,R_{t-1} + \gamma C_{t-1}$ for all countries.

4 Modelling time-varying effective transmission rates in Europe

It is important to recall here that, in large populations as in our data, the transmission rate can change only if the biology of virus changes (mutations), a vaccine is introduced, people change their behaviors voluntarily, or due to policy interventions. The transmission rate β_t reflects all these factors. In this section, we focus on understanding the drivers of β_t in the selected group of European Economies discussed above, which display very similar patterns during the first wave, but diverged significantly towards the end of 2020 both in terms of epidemic peaks, level of effective reproduction numbers, and the importance of herd immunity in slowing down the spread of the virus. In particular, we develop a panel data model of the time evolution of transmission rates across the nine European countries above.

We allow for five fundamental drivers. The first three are mandatory and voluntary social distancing and policy support that affects compliance with mandatory distancing policies. These drivers are likely to be of importance throughout the pandemic. They are also consistent with a simple decision-theoretic model presented in the appendix, as well as the large literature on behavioral epidemic modeling. In the latter stage of pandemic, two additional factors are important, which we also consider: progress on vaccinations and virus mutations in 2021.

Mandatory social distancing directly reduces the number of contacts as well as the exposure intensity. A strong theoretical rationale for imposition of mandated social distancing is the presence of externalities, i.e. the fact that agents do not internalize in their cost-benefit analysis that their individual behavior contributes to the aggregate diffusion of the epidemic.[14] However, mandated social distancing imposes economic costs and infringes on individual liberty leading to personal inconveniences (Hamermesh, 2020).

Economic support to workers, households and small businesses during the pandemic can shape incentives of individuals to comply with mandatory social distancing, as it weakens the economic need to interact in work activities. Consider an individual who has a non-teleworkable job and

[14] See, for example, Bethune and Korinek (2020), Eichenbaum, Rebelo, and Trabandt (2020), and Beck and Wagner (2020) in the international context. Eichenbaum, Rebelo, and Trabandt (2020), in particular, propose a behavioral SIR-macro model in which susceptible workers and consumers react to the epidemic risk by reducing their labor supply and consumption. Infected individuals take the aggregate infection process as given. As a result mandatory social distancing is optimal even though it is extremely costly in economic terms.

is fired or furloughed. While this leads to an immediate loss of income, if economic support is adequate, individuals can weather the pandemic without needing to seek paid employment in exposed occupations continuing to interact in production activities. Lack of compliance with social distancing has been documented empirically by Wright et al. (2020). Based on survey evidence, Papageorge et al. (2021) find that higher income is associated with larger changes in self-protective behavior, particularly for individuals who cannot telework. They conclude that, both in the United States and elsewhere, policies which assume universal compliance with self-protective measures or that otherwise do not account for socio-economic differences in the costs of doing so are unlikely to be effective or sustainable.

It is also well understood that risk induces precautionary behavior. Behavioral models of COVID-19 diffusion show that, as the probability of getting infected rises, individuals lower consumption and leisure activities to avoid infection (see Eichenbaum, Rebelo, and Trabandt (2020), Toxvaerd (2020), Atkeson (2021), and Gupta, Simon, and Wing (2020)). In particular, Battiston and Gamba (2020) provide cross section evidence that the $\mathcal{R}$ number during a COVID-19 outbreak is lower the larger the size of the initial wave.

To capture voluntary as well as mandatory social distancing policies, in our statistical analyses we make use of data compiled by the Oxford COVID-19 Government Response Tracker (OxCGRT) project, which is a standard source of comparable indices measuring social distancing and other COVID-19 policies across countries.[15] In particular, we use two aggregate indices: the 'policy stringency index' (capturing the containment and closure policies) and the 'economic support index' (as a proxy variable for support to comply with the containment policies). We then model precautionary behavior leading to voluntary social distancing with a 0/1 threshold variable explained in more detail below.

4.1 Econometric Model

The econometric specification that we propose is based on augmenting the aggregate (non-linear) moment condition obtained in (8) with an equation for the evolution of the transmission rate β_t.

[15] Data available at https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker.

Taking logs on both sides of equation (8) above and recalling that n is large, we obtain

$$ln\left[(1-c_{j,t+1})/(1-c_{jt})\right] = -\beta_{jt} i_{jt} + e_{j,t+1}, \tag{13}$$

where the subscript j denotes individual countries, $j = 1, 2, ..., N$, and $e_{j,t+1}$ is an error term, assumed to be orthogonal to i_{jt}. We complement the structural equation (13) with the following specification for the transmission rate process

$$\beta_{jt}/\gamma = a_j + \boldsymbol{\psi}' \mathbf{x}_{j,t-p} + \kappa I(f_{j,t-p} > \tau_f) + v_{jt}, \tag{14}$$

where $\mathbf{x}_{j,t-p}$ is a vector of regressors, lagged p periods, $I(f_{j,t-p} > \tau_f)$ is an indicator variable which takes the value of unity if the threshold variable, f_{jt}, also lagged p periods, goes above the threshold parameter τ_f, which as a first-order approximation is assumed to be the same across countries. The choice of the threshold variable, f_{jt}, is discussed below.

As a threshold variable, we use the 7-day moving average of the reported number of new cases (per 100,000 people), denoted by f_{jt}, as this is the most commonly watched variable used in the media when reporting on the spread of COVID-19 worldwide. The idea for a threshold effect, defined in terms of the number of new cases, is to capture possible non-linearities and changes in people's willingness to isolate consistent with surveys on the role of information diffusion under COVID-19 (Bursztyn et al., 2020).

In addition to estimating the panel regressions with a common constant term, a, we also allow for country-specific constant terms by replacing a in (14) with a_j, for $j = 1, 2, ..., N$. The parameters a_j are of particular interest as they can be viewed as an estimate of the basic reproduction number, $\mathcal{R}_0$, which relates to the spread of the virus right at the start of the epidemic when $\mathbf{x}_{j,t-p} = 0$, and $I(f_{j,t-p} > \tau_f) = 0$. Note at the start of the epidemic c_t is very close to zero, and voluntary and mandatory mitigation is not yet put into effect. Inevitably, there is always some delay between the start of the epidemic and any action by government or individuals to combat the virus. It is therefore reasonable to interpret a, or a_j as measures of $\mathcal{R}_0$ which we expect to be the same across all countries.

Substituting (14) in (13), we obtain the following estimating equation:

$$y_{j,t+1} = -\frac{ln\left[(1-c_{j,t+1})/(1-c_{jt})\right]}{\gamma i_{jt}} = a_j + \boldsymbol{\psi}'\mathbf{x}_{j,t-p} + \kappa I(f_{j,t-p} > \tau_f) + u_{j,t+1}, \tag{15}$$

where $u_{j,t+1}$ is the following composite error term:

$$u_{j,t+1} = v_{jt} - \gamma^{-1}\left(\frac{e_{t+1}}{i_{jt}}\right).$$

Depending on the assumptions regarding the error terms $u_{j,t+1}$, we will report three types of standard errors for the estimates. First, we consider standard errors assuming that $u_{j,t+1}$ is cross-sectionally as well as serially uncorrelated. These assumptions are undoubtedly restrictive, and therefore we also consider more robust alternatives. Our second approach to computing the standard errors, labelled as "robust1" in the tables below, allows for serial correlation and heteroskedasticity. Our third choice, denoted as "robust2" in the tables below, allows $u_{j,t+1}$ to be correlated both over time as well as across countries (i.e., over both the t and j dimensions). Estimation with the robust standard errors is described in the online supplement. It is important to note here that the assumptions underlying the "robust2" errors are quite general, and hence the latter are quite conservative.

The reason why we can allow for arbitrary correlation of errors over cross-section is due to small number of countries in our sample ($N = 9$) relative to the much larger time dimension (with the time series dimension varying between $T = 321$ and 646). An additional advantage of our small-N and large-T panel is that the specification in equation (15) can be estimated by least squares assuming only that the regressors are weakly exogenous.[16] This is in contrast to short panels (T-small and N-large), where *strict exogeneity* is required for consistency of least squares method.

An additional concern with our specification could be omitted variables and the presence of other confounding factors. The proposed specification (15) is parsimonious and encompasses the main

[16]Typically, least squares estimates in panels with *weakly exogenous* regressors will suffer from $O(1/T)$ bias. In panels with large T relative to N, which is the case in the present application, this bias is negligible.

factors considered in the literature. As contacts and susceptibility are not separately identified in our or any other SIR model, we use average indices of mandatory, support, and voluntary distancing, and we do not attempt to disentangle measures targeting the frequency of contacts or individual vulnerability.

The parameters of interest are the country-specific intercepts, $\alpha_1, \alpha_2, ..., \alpha_N$, that measure the basic reproduction number, $\mathcal{R}_0$, and the coefficients ψ, κ and τ_f that measure the relative importance of the five factors considered in the empirical analysis as we discussed above. We estimate the parameters of interest *jointly* using (15). This is to be contrasted with a two-step procedure whereby β_{jt} is first estimated by running country-specific regressions in (13), and these estimates, say $\hat{\beta}_{jt}$, are then used in a second stage panel regression where $\hat{\beta}_{jt}$ are regressed on our five drivers of transmission rates. The joint estimation approach is likely to provide more robust inference as compared to a two-step method that does not allow for the estimation uncertainty associated with using $\hat{\beta}_{jt}$.

We focus on the nine European countries: Belgium, France, Germany, Italy, Netherlands, Poland, Portugal, Spain, and the United Kingdom. As we noted already, the reason for focusing on these 9 countries is the fact that they experienced a similar trajectory at the start of the COVID-19 outbreak, but had quite differing outcomes subsequently. In this way we are able to exploit the cross country, as well as time series variations in the case data to quantify the effects of social distancing, vaccine uptake, and mutations on the effective transmission rates, β_{jt}/γ. Recall here that β_{jt}/γ differs from the effective reproduction number, $\mathcal{R}_{j,et}$, given by $\mathcal{R}_{j,et} = (1 - c_{jt})\left(\beta_{jt}/\gamma\right)$. As we noted earlier, $\mathcal{R}_{j,et}$ can fall below unity not because of the effectiveness of the mitigating policies, but simply because an increasingly larger fraction of the population is getting infected, the so called herd-immunity effect. To avoid the confounding effect of herding on the outcome variable that we want to explain, we focus on modeling of β_{jt}/γ and not $\mathcal{R}_{j,et}$.

4.2 Empirical Results

Consider first the panel regression (15) estimated with an unbalanced panel over the period February 23, 2020 to January 30, 2021. We allow for differences in the start dates of the outbreaks across

the countries. We initially choose the sample end date of January 31, 2021, prior the uptake of vaccination programs.[17] We consequently refer to this period as the pre-vaccination sample. Table 1 reports panel estimation results for $MF = 2$ and $MF = 3$ (the multiplication factor used to correct for under-reporting) and the lag order $p = 10$ days.[18]

As we can see from Table 1, both the social distancing policy index and the index of policy support have the expected negative signs, and are both statistically highly significant. This result confirms that both factors help contain the epidemic diffusion. Their estimated coefficients are robust to alternative MF corrections and lag lengths. The threshold effects also are highly statistically significant. Allowing for country-specific intercepts slightly increases the coefficients on the social distancing and the economic support indexes, and slightly lowers the estimates of the threshold effect, which, however, remains sizeable and highly significant. Allowing for error correlations (over time and space) increases the estimated standard errors, as to be expected, but the widening do not alter the inference that we make. All variables remain statistically significant.

Regarding the magnitude of the estimated coefficients, it is first useful to note that the intercept is the model-implied estimate of the $\mathcal{R}_0$ number. As discussed before, the definition of the basic reproduction number assumes no changes in behaviour in response to the pandemic. The estimated value of the common intercept, α, in Table 1, is 5.36, identical for both choices of MF, with a tight robust2 standard error (0.23). The country-specific estimates of the $\mathcal{R}_0$ number also are all in a surprisingly tight range from 5.03 to 5.73. The estimated threshold value is rather small, below 1 daily new confirmed case per 100,000 people, representing about 4-percent quantile of the daily new cases per 100,000 people in the pre-vacciantion sample. The result suggests that the threshold effect kicks in soon after the onset of the pandemic. In other words, people's behaviour changes soon after the onset of the pandemic, which makes the estimation of the $\mathcal{R}_0$ number from case numbers without a conditional statistical models difficult (if not impossible). Threshold effects alone significantly reduce the effective reproduction number (by more than 2), but on its own is not

[17] Although some vaccines doses were already administered in December 2020, only a very small fraction of the population was fully vaccinated by the end January 2021 – less than 0.8 percent for all countries in the sample, except Italy, where the share reached 1.07 percent. Vaccine uptake had increased considerably by late spring and summer of 2021.

[18] Results for a shorter lag of $p = 7$ days and a longer lag $p = 14$ days are provided in Table S2 in the online supplement.

enough to bring the effective $\mathcal{R}$ number below one, and other mitigating measures are needed. Both economic support and social distancing policy indexes further reduce the transmission significantly, with the social distancing index having a much larger contribution of the two factors.

Next, we extend the sample to November 30, 2021, prior to the arrival of Omicron variant in December of 2021. Throughout 2021, there was a significant progress with vaccine uptake, which is expected to mitigate the virus' transmission. We therefore include an additional variable measuring the share of population fully vaccinated. We expect the coefficient on this variable to be negative. Another factor in 2021 is the Delta mutation, which is considered to be more contagious and became dominant in Europe during the summer of 2021. Therefore, we also include a variable measuring the country-specific share of the Delta variant among the sequenced confirmed cases. Again a positive coefficient is expected on this additional control variable.

Table 2 presents estimates for the sample ending November 30, 2021. The table shows very similar findings for the estimates of the social distancing index, the economic support index, the threshold effects, as well as the $\mathcal{R}_0$ numbers, as reported in Table 1 for the pre-vaccination sample. The coefficients on the share of population fully vaccinated are all negative (as expected), with magnitude in the range -1.2 to -1.8 and statistically highly significant. Coefficient estimates on the Delta variant share are all positive, in line with our prior, suggesting the basic reproduction number of the Delta variant is about 0.9 to 1.1 larger compared with the earlier strains. It is also reasonable to expect that further positive effects can be identified once our sample is extended to cover the emergence and the spread of Omicron.

Table 3 provides summary statistics for all covariates used in the regressions. It reports the sample means and standard deviations (in brackets) for the pre-vaccination sub-sample ending January 31, 2021, and for the remainder of the sample (February 1, 2021 to November 30, 2021).[19] Multiplying the average values of the regressors in Table 3 with the estimated coefficients in Table 2 (using the specification with common intercept and $MF = 2$), we find that the threshold indicator brought down the scaled effective transmission rate, β_t/γ, by about a half.[20] The second largest

[19]Table S4 in the online supplement reports minimum and maximum ranges.

[20]Recall the effective reproduction number is $\mathcal{R}_{j,et} = (1 - c_{jt})\left(\beta_{jt}/\gamma\right)$, where $(1 - c_{jt})$ is the share of susceptible population, and we refer to β_{jt}/γ as the scaled effective transmission rate, which is in the same "units" as the

reduction comes from social distancing index (with a -1.3 average contribution to the reduction of β_t/γ in pre-vaccination subsample, and -1.2 contribution in the remainder of the sample). In contrast, the contribution from the economic support index is smaller, only by -0.2 on average in both periods. According to our estimates, vaccination contributed -0.6 on average in the latter sample (February-November, 2021), with the contribution increasing in magnitude throughout the sample to -1.7, as share of vaccinated population increased over time (See Table S4 in the online supplement). Hence, the contribution from vaccinations at the end of the sample (November 30) exceeds the average contribution coming jointly from the social distancing and economic support variables.

reproduction number, since $\mathcal{R}_{j,et} \approx \beta_{jt}/\gamma$ when $(1 - c_{jt}) \approx 0$.

Table 1: Panel regressions of effective transmission rates for panels with common and country-specific $\mathcal{R}_0$ numbers.

Pre-vaccination sample ending January 31, 2021 (all covariates are lagged $p = 10$ days)

	Pooled Estimates		Fixed Effects Estimates	
Multiplication Factor:	$MF = 2$	$MF = 3$	$MF = 2$	$MF = 3$
Stringency Index	**-2.08**	**-2.08**	**-2.21**	**-2.19**
standard s.e. (t-ratio)	0.11 (-19.6)	0.11 (-19.4)	0.11 (-20.1)	0.11 (-19.9)
robust1 s.e. (t-ratio)	0.18 (-11.6)	0.18 (-11.5)	0.18 (-12.5)	0.18 (-12.4)
robust2 s.e. (t-ratio)	0.25 (-8.3)	0.25 (-8.2)	0.34 (-6.5)	0.34 (-6.4)
Economic Support	**-0.45**	**-0.44**	**-1.02**	**-1.01**
standard s.e. (t-ratio)	0.07 (-6.6)	0.07 (-6.4)	0.09 (-11.4)	0.09 (-11.3)
robust1 s.e. (t-ratio)	0.14 (-3.3)	0.14 (-3.2)	0.17 (-5.9)	0.17 (-5.9)
robust2 s.e. (t-ratio)	0.23 (-2.0)	0.23 (-1.9)	0.23 (-4.4)	0.23 (-4.4)
Threshold Variable	**-2.46**	**-2.46**	**-2.03**	**-2.02**
standard s.e. (t-ratio)	0.09 (-26.9)	0.09 (-26.7)	0.10 (-21.0)	0.10 (-20.8)
robust1 s.e. (t-ratio)	0.25 (-9.7)	0.25 (-9.7)	0.24 (-8.4)	0.24 (-8.4)
robust2 s.e. (t-ratio)	0.43 (-5.8)	0.43 (-5.8)	0.60 (-3.4)	0.60 (-3.4)
threshold value	0.20	0.20	0.20	0.20
$\mathcal{R}_0$ **numbers (Constant Terms)**				
common [robust2 s.e.]	5.36 [0.23]	5.36 [0.23]		
specific [robust2 s.e.]:				
Belgium			5.39 [0.62]	5.39 [0.62]
France			5.40 [0.63]	5.40 [0.64]
Germany			5.06 [0.65]	5.05 [0.65]
Italy			5.53 [0.66]	5.52 [0.66]
Netherlands			5.37 [0.62]	5.37 [0.63]
Poland			5.03 [0.64]	5.03 [0.64]
Portugal			5.51 [0.64]	5.51 [0.64]
Spain			5.63 [0.64]	5.64 [0.64]
United Kingdom			5.73 [0.62]	5.73 [0.62]
R-squared	0.52	0.52	0.55	0.54

Notes: Number of observations is 2990 with $N = 9$ countries, $T_{\min} = 321$ and $T_{\max} = 343$ days. The estimation sample is unbalanced at the beginning. Starting dates of individual country samples are: 7-March-2020 (Belgium), 3-March-2020 (France), 2-March-2020 (Germany), 24-February-2020 (Italy), 7-March-2020 (Netherlands), 17-March-2020 (Poland), 15-March-2020 (Portugal), 3-March-2020 (Spain), and 3-March-2020 (United Kingdom). The last period is 31-January-2021 for all countries. "Robust1" standard errors are robust to serial correlation only (Newey-West type correction), whereas "robust2" standard errors are robust to serial correlation as well as any cross-sectional correlation. See online appendix for a description of the estimation of standard errors. Figures in parentheses are t-ratios. The figures in square brackets of the common intercept or the country-specific fixed effects are the standard errors robust to serial correlation as well as cross-sectional correlation (robust2). Oxford stringency and economic support indices are divided by 100 so that they take values between zero and one. Lag order is set to $p = 10$ days in all regressions. Estimates for lag orders 7 and 14 are provided in Table S2 of the online supplement.

Table 2: Panel regressions of effective transmission rates for panels common and country-specific $\mathcal{R}_0$ numbers.

Full sample ending November 30, 2021 (all covariates are lagged $p = 10$ days)

	Pooled Estimates		Fixed Effects Estimates	
Multiplication Factor:	$MF = 2$	$MF = 3$	$MF = 2$	$MF = 3$
Stringency Index	**-1.98**	**-2.01**	**-2.13**	**-2.12**
standard s.e. (t-ratio)	0.08 (-23.9)	0.09 (-22.8)	0.09 (-24.2)	0.09 (-22.6)
robust1 s.e. (t-ratio)	0.18 (-11.2)	0.18 (-10.9)	0.18 (-12.2)	0.18 (-11.6)
robust2 s.e. (t-ratio)	0.27 (-7.3)	0.28 (-7.3)	0.30 (-7.1)	0.31 (-6.9)
Economic Support	**-0.30**	**-0.29**	**-0.55**	**-0.59**
standard s.e. (t-ratio)	0.05 (-6.3)	0.05 (-5.6)	0.06 (-8.8)	0.07 (-8.8)
robust1 s.e. (t-ratio)	0.10 (-2.9)	0.11 (-2.7)	0.15 (-3.6)	0.15 (-3.8)
robust2 s.e. (t-ratio)	0.17 (-1.8)	0.17 (-1.7)	0.22 (-2.5)	0.21 (-2.8)
Vaccinated Share	**-1.62**	**-1.32**	**-1.82**	**-1.53**
standard s.e. (t-ratio)	0.15 (-10.6)	0.16 (-8.1)	0.16 (-11.8)	0.17 (-9.2)
robust1 s.e. (t-ratio)	0.33 (-4.9)	0.37 (-3.6)	0.33 (-5.6)	0.36 (-4.2)
robust2 s.e. (t-ratio)	0.53 (-3.1)	0.57 (-2.3)	0.51 (-3.6)	0.55 (-2.8)
Delta Variant Share	**1.04**	**1.02**	**1.12**	**1.11**
standard s.e. (t-ratio)	0.10 (10.4)	0.11 (9.6)	0.10 (11.1)	0.11 (10.3)
robust1 s.e. (t-ratio)	0.21 (4.9)	0.24 (4.3)	0.21 (5.3)	0.23 (4.7)
robust2 s.e. (t-ratio)	0.31 (3.4)	0.34 (3.0)	0.30 (3.8)	0.33 (3.4)
Threshold Variable	**-2.62**	**-2.59**	**-2.38**	**-2.33**
standard s.e. (t-ratio)	0.08 (-31.4)	0.09 (-29.1)	0.09 (-27.1)	0.09 (-25.0)
robust1 s.e. (t-ratio)	0.26 (-10.1)	0.26 (-10.0)	0.26 (-9.2)	0.26 (-9.0)
robust2 s.e. (t-ratio)	0.43 (-6.1)	0.42 (-6.1)	0.69 (-3.5)	0.69 (-3.4)
threshold value	0.20	0.20	0.20	0.20
$\mathcal{R}_0$ **numbers (Constant Terms)**				
common [robust2 s.e.]	5.33 [0.23]	5.34 [0.23]		
specific [robust2 s.e.]:				
Belgium			5.40 [0.64]	5.43 [0.65]
France			5.32 [0.66]	5.32 [0.67]
Germany			5.21 [0.68]	5.16 [0.68]
Italy			5.53 [0.66]	5.51 [0.67]
Netherlands			5.36 [0.66]	5.39 [0.67]
Poland			5.19 [0.66]	5.18 [0.66]
Portugal			5.44 [0.66]	5.44 [0.67]
Spain			5.48 [0.65]	5.49 [0.66]
United Kingdom			5.50 [0.65]	5.51 [0.66]
R-squared	0.41	0.38	0.42	0.39

Notes: Number of observations is 5717 with $N = 9$ countries, $T_{\min} = 624$ and $T_{\max} = 646$ days. The estimation sample is unbalanced at the beginning. Starting dates of individual country samples are: 7-March-2020 (Belgium), 3-March-2020 (France), 2-March-2020 (Germany), 24-February-2020 (Italy), 7-March-2020 (Netherlands), 17-March-2020 (Poland), 15-March-2020 (Portugal), 3-March-2020 (Spain), and 3-March-2020 (United Kingdom). The last period is 30-November-2021 for all countries. "Robust1" standard errors are robust to serial correlation only (Newey-West type correction), whereas "robust2" standard errors are robust to serial correlation as well as any cross-sectional correlation. See online appendix for a description of the estimation of standard errors. Figures in parentheses are t-ratios. The figures in square brackets of the common intercept or the country-specific fixed effects are the standard errors robust to serial correlation as well as cross-sectional correlation (robust2). Oxford stringency and economic support indices are divided by 100 so that they take values between zero and one. Lag order is set to $p = 10$ days in all regressions. Estimates for lag orders 7 and 14 are provided in Table S3 of the online supplement.

Table 3: Summary statistics for regressors: sample means and sample standard deviations (in brackets)

	Pre-vaccination sample ending January 31 2021				
	Stringency index	Economic support	Vaccinated share	Delta share	Threshold indicator
Belgium	0.59 (0.15)	0.77 (0.22)	0.00 (0.00)	0.00 (0.00)	0.97 (0.18)
France	0.63 (0.18)	0.67 (0.26)	0.00 (0.00)	0.00 (0.00)	0.95 (0.22)
Germany	0.61 (0.15)	0.41 (0.17)	0.00 (0.00)	0.00 (0.00)	0.94 (0.23)
Italy	0.73 (0.12)	0.62 (0.23)	0.00 (0.00)	0.00 (0.00)	0.96 (0.20)
Netherlands	0.59 (0.17)	0.76 (0.22)	0.00 (0.00)	0.00 (0.00)	0.96 (0.19)
Poland	0.58 (0.21)	0.50 (0.21)	0.00 (0.00)	0.00 (0.00)	0.95 (0.22)
Portugal	0.65 (0.12)	0.73 (0.11)	0.00 (0.00)	0.00 (0.00)	0.97 (0.17)
Spain	0.65 (0.16)	0.80 (0.23)	0.00 (0.00)	0.00 (0.00)	0.96 (0.21)
United Kingdom	0.67 (0.18)	0.92 (0.27)	0.00 (0.00)	0.00 (0.00)	0.94 (0.23)
all 9 countries	0.63 (0.17)	0.69 (0.26)	0.00 (0.00)	0.00 (0.00)	0.95 (0.21)
	1 February 2021 - 30 November 2021 sample				
Belgium	0.53 (0.09)	0.75 (0.00)	0.36 (0.30)	0.45 (0.46)	1.00 (0.00)
France	0.62 (0.09)	0.43 (0.11)	0.32 (0.26)	0.44 (0.46)	1.00 (0.00)
Germany	0.66 (0.13)	0.38 (0.00)	0.33 (0.26)	0.45 (0.47)	1.00 (0.00)
Italy	0.69 (0.09)	0.75 (0.00)	0.34 (0.27)	0.44 (0.45)	1.00 (0.00)
Netherlands	0.58 (0.17)	0.65 (0.12)	0.32 (0.27)	0.45 (0.47)	1.00 (0.00)
Poland	0.54 (0.15)	0.84 (0.12)	0.28 (0.21)	0.42 (0.46)	1.00 (0.00)
Portugal	0.63 (0.13)	0.75 (0.00)	0.40 (0.34)	0.53 (0.46)	1.00 (0.00)
Spain	0.56 (0.12)	0.88 (0.00)	0.38 (0.31)	0.43 (0.45)	1.00 (0.00)
United Kingdom	0.58 (0.16)	0.88 (0.28)	0.38 (0.26)	0.57 (0.46)	1.00 (0.00)
all 9 countries	0.60 (0.14)	0.70 (0.21)	0.35 (0.28)	0.47 (0.46)	1.00 (0.00)

Notes: This table report sample means (main entries) and sample deviations (in brackets) of the individual regressors in pooled regressions presented in Table 2. The top panel reports summary statistics for the pre-vaccination sample (ending January 31 2021), and the bottom panel reports summary statistics for the remainder of the full sample – 1 February 2021 to 30 November 2021. Table S4 in the online supplement reports additional summary statistics (minimum and maximum values).

5 Conclusions

This paper first estimates effective transmission rates and reproduction numbers for a large number of countries and regions based on a moment condition that can be derived from an agent-based stochastic network epidemic model. It then models their evolution for selected European countries with similar experiences at the outset of the pandemic but different outcomes subsequently as a function of social distancing, incentives to comply, the emergence of mutations, and vaccine uptake.

From a methodological perspective, the estimation approach that we propose permits distinguishing, at any jurisdictional level, between changes in the effective reproduction number due to herd immunity and changes due to variation in the average contact or the susceptibility to infection, which are the structural determinants of the epidemic diffusion. At the empirical level, using only daily COVID-19 case statistics, the paper provides estimates of transmission rates, allowing for the under-reporting of infected cases in available COVID case statistics.

Evidence based on panel data modeling indicates that the diversity of outcomes that we document resulted from the non-linear interaction of mandated and voluntary social distancing and the economic incentives that governments provided to support isolation through the end of 2020, with no one factor independently capable of lowering the reproduction number below one. However, the importance of these factors declines over time, with vaccine uptake driving heterogeneity in country experiences in 2021. Our estimates thus suggest that no one factor alone is sufficient to bring the $\mathcal{R}$ number below one and to keep it there without substantial contributions from the others. Our panel regressions also allow us to identify the basic reproduction number, $\mathcal{R}_0$, by the country-specific intercepts, giving estimates that are very similar across the nine European countries, and all well in excess of the values of 2.5 to 3.9 assumed in the literature. Our estimates also confirm the higher transmission rate of the Delta variant.[21] The main conclusions of the paper are robust to misreporting, lag orders, error heteroskedasticity, error serial correlation, and slope heterogeneity.

[21] Recall that our sample ends before the emergence of the Omicron variant.

References

Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston (2020). Optimal targeted lockdowns in a multi-group SIR model. NBER Working Paper No. 27102.

Akbarpour, M., C. Cook, A. Marzuoli, S. Mongey, A. Nagaraj, M. Saccarola, P. Tebaldi, S. Vasserman, and H. Yang (2020). Socioeconomic network heterogeneity and pandemic policy response. NBER Working Paper No. 27374.

Alvarez, F., D. Argente, and F. Lippi (2021). A Simple Planning Problem for COVID-19 Lock-down, Testing, and Tracing. *AER: Insights 3*, 367–82.

Atkeson, A. (2021). A parsimonious behavioral SEIR model of the 2020 COVID epidemic in the United States and the United Kingdom. NBER Working Paper No. 28434.

Atkeson, A., K. Kopecky, and T. Zha (2020a). Four stylized facts about COVID-19. NBER Working Paper No. 27719.

Atkeson, A., K. A. Kopecky, and T. A. Zha (2020b). Estimating and forecasting disease scenarios for COVID-19 with an SIR Model. NBER Working Paper No. 27335.

Avery, C., W. Bossert, A. Clark, G. Ellison, and S. F. Ellison (2020). An economist's guide to epidemiology models of infectious disease. *Journal of Economic Perspectives 34*(4), 79–104.

Baqaee, D., E. Farhi, M. Mina, and J. H. Stock (2020). Policies for a second wave. *Brookings Papers on Economic Activity (forthcoming)*.

Battiston, P. and S. Gamba (2020). COVID-19: R0 is lower where outbreak is larger. *Health Policy 125*, 141–147.

Beck, T. and W. Wagner (2020). National containment policies and international cooperation. CEPR Discussion Paper No. DP14668.

Bethune, Z. A. and A. Korinek (2020). Covid-19 infection externalities: Trading off lives vs. livelihoods. NBER Working Paper No. 27009.

Brodeur, A., D. M. Gray, A. Islam, and S. Bhuiyan (2020). A literature review of the economics of COVID-19. IZA Discussion Paper No. 13411.

Bursztyn, L., A. Rao, C. P. Roth, and D. H. Yanagizawa-Drott (2020). Misinformation during a pandemic. NBER Working Paper No. 27417.

Cakmakli, C., S. Demiralp, S. Kalemli-Ozcan, S. Yesiltas, and M. Yildirim (2021). The economic case for global vaccinations: An epidemiological model with international production networks. NBER Working Paper No. 28395.

Cakmakli, C., S. Demiralp, S. Kalemli-Ozcan, S. Yesiltas, and M. A. Yildirim (2020). COVID-19 and emerging markets: An epidemiological model with international production networks and capital flows. IMF Working Paper No. 20/133.

Cakmakli, C. and Y. Simsek (2020). Bridging the COVID-19 data and the epidemiological model using time varying parameter SIRD model. Koc University-TUSIAD Economic Research Forum Working Paper No. 2013.

Caselli, F., F. Grigoli, W. Lian, and D. Sandri (2020). The great lockdown: dissecting the economic effects. *International Monetary Fund. World Economic Outlook*, 65–84.

Chowell, G. and H. Nishiura (2008). Quantifying the transmission potential of pandemic influenza. *Physics of Life Reviews 5*(1), 50–77.

Chudik, A., M. H. Pesaran, and A. Rebucci (2020). Voluntary and mandatory social distancing: Evidence on COVID-19 exposure rates from Chinese provinces and selected countries. NBER Working Paper 27039.

Del Valle, S. Y., J. M. Hyman, and N. Chitnis (2013). Mathematical models of contact patterns between age groups for predicting the spread of infectious diseases. *Mathematical Biosciences and Engineering 10*, 1475.

Diekmann, O. and J. Heesterbeek (2000). *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. John Wiley & Son, New York, ISBN: 0471986828.

Eichenbaum, M. S., S. Rebelo, and M. Trabandt (2020). The macroeconomics of epidemics. NBER Working Paper No. 26882.

Fang, H., L. Wang, and Y. Yang (2020). Human mobility restrictions and the spread of the novel coronavirus (2019-nCoV) in China. NBER Working Paper No. 26906.

Favero, C. A., A. Ichino, and A. Rustichini (2021). Restarting the economy while saving lives under Covid-19. Available at SSRN: https://ssrn.com/abstract=3580626, February 2021.

Ferguson, N., P. Walker, C. Whittaker, and et al (2020). Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand. Imperial College London COVID-19 Reports, Report No. 9, 16 March 2020.

Fernández-Villaverde, J. and C. I. Jones (2020). Estimating and simulating a SIRD model of COVID-19 for many countries, states, and cities. NBER working paper No. 27128.

Gibbons, C. L., M.-J. J. Mangen, D. Plass, and et al (2014). Measuring underreporting and under-ascertainment in infectious disease datasets: a comparison of methods. *BMC Public Health 14:147*, 1471–2458.

Gupta, S., K. I. Simon, and C. Wing (2020). Mandated and voluntary social distancing during the covid-19 epidemic: A review.

Hamermesh, D. S. (2020). Lock-downs, loneliness and life satisfaction. NBER Working Paper No. 27018.

Harko, T., F. S. Lobo, and M. Mak (2014). Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal death and birth rates. *Applied Mathematics and Computation 236*, 184–194.

Jinjarak, Y., R. Ahmed, S. Nair-Desai, W. Xin, and J. Aizenman (2020). Accounting for global COVID-19 diffusion patterns, January–April 2020. *Economics of disasters and climate change 4*(3), 515–559.

Kermack, W. O. and A. G. McKendrick (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London 22*(772), 700–721.

Korolev, I. (2020). Identification and estimation of the SEIRD epidemic model for COVID-19. *Journal of Econometrics 220*, 63–85.

Kucharski, A. J., T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, and R. M. Eggo (2020). Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *The Lancet Infectious Diseases*.

Li, Q., X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, ..., and Z. Feng (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine 382*(13), 1199–1207.

Matrajt, L. and T. Leung (2020). Evaluating the effectiveness of social distancing interventions to delay or flatten the epidemic curve of coronavirus disease. *Emerging Infectious Diseases 26*(8), 1740–1748.

Metz, J. A. J. (1978). The epidemic in a closed population with all susceptibles equally vulnerable; some results for large susceptible populations and small initial infections. *Acta Biotheoretica 27*, 75–123.

Moriarty, L., M. Plucinski, B. Marston, and et al. (2020). Public health responses to COVID-19 outbreaks on cruise ships - worldwide, February-March 2020. Morbidity and Mortality Weekly Report (MMWR), 26 March 2020, 69:347-352.

Nikbakht, R., M. R. Baneshi, A. Bahrampour, and A. Hosseinnataj (2019). Comparison of methods to estimate basic reproduction number (R0) of influenza, using Canada 2009 and 2017-18 A (H1N1) data. *Journal of Research in Medical Sciences 24*, 24–67.

Obadia, T., R. Haneef, and P.-Y. Boëlle (2012). The R0 package: A toolbox to estimate reproduction numbers for epidemic outbreaks. *BMC Medical Informatics and Decision Making 12*(1), 1–9.

Papageorge, N. W., M. V. Zahn, M. Belot, E. Van den Broek-Altenburg, S. Choi, J. C. Jamison, and E. Tripodi (2021). Socio-demographic factors associated with self-protecting behavior during the Covid-19 pandemic. *Journal of Population Economics 34*, 691–738.

Pesaran, M. H. and C. F. Yang (2021). Matching theory and evidence on COVID-19 using a stochastic network SIR model. *Journal of applied econometrics*. Forthcomming.

Riou, J. and C. L. Althaus (2020). Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Eurosurveillance 25*(4).

Salje, H., D. A. Cummings, and J. Lessler (2016). Estimating infectious disease transmission distances using the overall distribution of cases. *Epidemics 17*, 10–18.

Satsuma, J., R. Willox, A. Ramani1, B. Grammaticos, and A. Carstea (2004). Extending the SIR epidemic model. *Physica A: Statistical Mechanics and its Applications 336*, 369–375.

Toda, A. A. (2020). Susceptible-infected-recovered (SIR) dynamics of COVID-19 and economic impact. Preprint arXiv: 2003.11221.

Toxvaerd, F. (2020). Equilibrium social distancing. Cambridge Working Papers in Economics No. 2021, Faculty of Economics, University of Cambridge.

Wang, C., L. Liu, X. Hao, H. Guo, Q. Wang, J. Huang, ..., and T. Wu (2020). Evolving epidemiology and impact of non-pharmaceutical interventions on the outbreak of coronavirus disease 2019 in Wuhan, China. *medRxiv*.

Wright, A. L., K. Sonin, J. Driscoll, and J. Wilson (2020). Poverty and economic dislocation reduce compliance with covid-19 shelter-in-place protocols. *Journal of Economic Behavior Organization 180*, 544–554.

A Appendix

In this appendix, we motivate the time-varying specification of the effective reproduction number and the transmission rate spelled out in the main text and provide the moment condition that we use for estimation based on the network-model of Pesaran and Yang's (2021, PY). We also set up a simple decision-theoretic model of social distancing to justify the choice of economic support variable and the threshold indicator used in the panel-data model of the transmission rate in Europe.

A.1 A time-varying specification of the the single-group PY model

PY considers both single- and multi-group models built from the bottom up from an individual stochastic network model of interaction and COVID-19 contagion. After spelling out the critical assumptions on the infection diffusion process, here, we focus on the moment conditions for the estimation of the single-group model that tracks very closely the multi-group model in simulations.

Consider n individuals indexed by $i = 1, 2, ..., n$. Some are initially infected while the rest is susceptible to be infected at a later date, indexed by t. An individual i infected at date $t = t_i^*$ is represented by the random variable

$$x_{it} = 0, \text{ for all } t < t_i^*; \quad \text{and} \quad x_{it} = 1, \text{ for all } t \geq t_i^*. \tag{A.1}$$

The event of recovery or death of individual i will be represented by the variable y_{it}, which is equal to zero unless the individual is "removed", i.e., recovered or dead. The indicator variable

$$z_{it} = (1 - y_{it})\, x_{it} \tag{A.2}$$

then denotes "an active" infection, namely if the individual i is infected and not yet recovered. Specifically, z_{it} takes value 1 only if the individual is infected and not yet recovered. It takes value 0 if individual i has not yet been infected, or has been infected but recovered/died. It follows that $x_{it} = z_{it} + y_{it}$.

Individual i *becoming* infected is modelled using the following Markov switching process as a

function of the latent variable $x^*_{i,t+1}$,

$$x_{i,t+1} = x_{it} + (1 - x_{it})\, I\left(x^*_{i,t+1} > 0\right), \tag{A.3}$$

where $I(\mathcal{A})$ is the indicator function that takes the value of unity if $\mathcal{A}$ holds and zero otherwise, and $x^*_{i,t+1}$ is composed of two different components:

$$x^*_{i,t+1} = \tau_{it} \sum_{j=1}^{n} d_{ij}(t)\, z_{jt} - \mu_{it}\xi_{i,t+1}. \tag{A.4}$$

Since $I\left(x^*_{i,t+1} > 0\right)$ is unaffetced by re-scaling of $x^*_{i,t+1}$, and by assumption $\mu_{it} > 0$, then the above can also be written equivalenly as

$$\frac{x^*_{i,t+1}}{\mu_{it}} = \left(\frac{\tau_{it} k_t}{\mu_{it}}\right) \sum_{j=1}^{n} \left(\frac{d_{ij}(t)}{k_t}\right) z_{jt} - \xi_{i,t+1}.$$

The first component of the latent variable $x^*_{i,t+1}$, $\tau_{it} \sum_{j=1}^{n} d_{ij}(t)\, z_{jt}$, captures to the contact pattern of individual i with actively infected individuals, z_{it}. τ_{it} is a individual measure of exposure intensity (an exposure intensity parameter). $\mathbf{D}(t) = [d_{ij}(t)]$ is a contact network matrix, such that $d_{ij}(t) = 1$ if individual i is in contact with individual j at time t. It is assumed the elements of the $n \times n$ network matrix $\mathbf{D}(t)$ are independent draws with $E[d_{ij}(t)] = p_t = k_t/n$, where k_t is the mean daily contact during day t.

The second component of the latent variable $x^*_{i,t+1}$, $\mu_{it}\xi_{i,t+1} > 0$, is an unobserved individual-specific positive random variable, which captures the unobservable characteristics that lead to different probabilities of infection, even for individuals with the same contact pattern. Following PY, we assume $\xi_{i,t+1}$ is exponentially and independently distributed over i and t with the cumulative distribution function given by

$$\Pr\left(\xi_{i,t+1} < a\right) = 1 - \exp(-a), \quad \text{for } a > 0, \tag{A.5}$$

where $E\left(\xi_{i,t+1}\right) = 1$.

Since individual i becomes infected if $x^*_{i,t+1} > 0$, what matters most for the spread of the virus is the ratio $k_t\tau_{it}/\mu_{it}$ in Equation (A.4) which captures both social distancing (voluntary or mandatory) and vaccination. Note also that, at a theoretical level, it is not possible to distinguish between reducing interactions via k_t, or reducing the intensity of these exposures via τ_{it} through hand-washing, mask-wearing, etc., or via vaccination that increases μ_{it}.

The recovery depends on the number of days since infection. Specifically, we assume that the recovery process for individual i is given by

$$y_{i,t+1} = y_{it} + z_{it}\ \zeta_{i,t+1}\left(t_i^*\right), \tag{A.6}$$

where $\zeta_{i,t+1}\left(t_i^*\right) = 1$ if individual i recovers at time $t+1$, having been infected exactly at time t_i^* and not before, and $\zeta_{i,t+1}\left(t_i^*\right) = 0$, otherwise. Furthermore, it is assumed the time to removal, denoted by $T^*_{it} = t - t_i^*$, follows the geometric distribution (for $t - t_i^* = 1, 2, \ldots$)

$$\Pr\left[\zeta_{i,t+1}\left(t_i^*\right) = 1\right] = \Pr\left(T^*_{it} = t - t_i^*\right) = \gamma\left(1-\gamma\right)^{t-t_i^*-1}, \tag{A.7}$$

where γ can be interpreted as the probability of recovery at time $t+1$ having remained infected for $t - t_i^* - 1$ days. This implies the following recovery micro-moment condition

$$E\left(y_{i,t+1}\,|y_{it}, z_{it}\right) = y_{it} + \gamma z_{it}. \tag{A.8}$$

Noting that a susceptible individual can now be denoted by $s_{it} = 1 - z_{it} - y_{it}$, and using the definitions above we have that cumulative number of infections is $C_t = \sum_{i=1}^n x_{it}$, the total number of removed (recovered or deceased) is $R_t = \sum_{i=1}^n y_{it}$, and the total number of active cases is $I_t = \sum_{i=1}^n z_{it} = C_t - R_t$, while the number of "susceptible" individuals is $S_t = \sum_{i=1}^n s_{it} = n - I_t - R_t$.

By normalizing μ_{it} to 1, and assuming that $\tau_{it} = \tau$ and $k_t = k$ are time invariant, PY derive the following approximation for the aggregate moment condition for the population share of the

total cumulative confirmed cases, $c_t = C_t/n$:

$$E\left(\frac{1-c_{t+1}}{1-c_t}\middle|\ i_t\right) = e^{-\beta i_t} + O\left(n^{-1}\right), \tag{A.9}$$

where $\beta = (1-e^{-\tau})\,k \approx \tau k$ is the transmission rate, and $i_t = I_t/n$ is the per capita number of active cases. Under these assumptions, the basic reproduction number, $\mathcal{R}_0$, is given by

$$\mathcal{R}_0 \approx \gamma^{-1} np(1-e^{-\tau}) \approx \gamma^{-1}\tau k = \gamma^{-1}\beta_0. \tag{A.10}$$

In our empirical analysis, we allow for time variations in k, τ and μ and and accordingly obtain the following mean transmission rate

$$\beta_t \approx \frac{k_t \tau_t}{\mu_t},$$

where k_t is the average number of contact during day t, τ_t is the average exposure intensity of the susceptible population to the virus, and $\mu_t \geq 1$ expected to be an increasing function of the proportion of susceptible population vaccinated, thus capturing the effects of vaccination.

A.2 A simple decision-theoretic model of voluntary social distancing

To motivate the choice of variables that we consider in the panel data model of the transmission rate in Europe, in this appendix, we introduce a simple decision-theoretic model of social distancing.

Consider an individual i from a fixed population of size n in the epidemic day t, and suppose the individual in question is faced with the voluntary decision of whether to isolate or not. Under self-isolation, an individual that does not telework incurs the loss of wages net of any COVID-19 economic support amounting to $(1-\theta_{it})w_{it}$, plus the inconvenience cost, a_{it}, of being isolated, where w_{it} is the wage and θ_{it} is the percentage of income lost which is compensated by the government support. For those individuals who can work from home θ_{it} is likely to be 1 or very close to it. But for many workers who are furloughed or become unemployed, θ_{it} is likely to be close to zero, unless they are compensated by transfers from the government.

On the other hand, if the individual decides not to self-isolate then he/she receives the uncertain pay-off of $(1-d_{it})w_{it}-d_{it}\phi_{it}$, where d_{it} is an indicator which takes the value of unity if the individual contracts the disease and zero otherwise. The parameter ϕ_{it} represents the cost of contracting the disease and is expected to be quite high. We are ruling out the possibility of death as an outcome and also assume that if the individual does not isolate and get sick does not earn the wage.

In this setting the individual decides to self-isolate if the sure loss of self-isolating is less than the expected loss of not self-isolating, namely if

$$(1-\theta_{it})\,w_{it}+a_{it}<E\left[d_{it}\phi_i-(1-d_{it})w_{it}\,|\mathcal{I}_{t-1}\right], \tag{A.11}$$

where $\mathcal{I}_{t-1}$ is the publicly available information that includes c_{t-1}, the total number of infections. Assume now for simplicity that the probability of anyone contracting the disease is uniform across the population and this is correctly perceived to be given by π_{t-1}. Hence $E\left(d_{it}\,|\mathcal{I}_{t-1}\right)=\pi_{t-1}$, and the condition for self-isolating in any day t can be written as

$$(2-\theta_{it})w_{it}+a_{it}<\pi_{t-1}(w_{it}+\phi_{it}),$$

or as

$$\frac{2-\theta_{it}+(a_{it}/w_{it})}{1+(\phi_{it}/w_{it})}=\lambda_{it}<\pi_{t-1}. \tag{A.12}$$

Since $\pi_{t-1}\leq 1$, then for individual i to self-isolate we must have $\lambda_{it}<1$ (note that $\lambda_{it}\geq 0$, with $\lambda_{it}=0$ when $\phi_{it}\to\infty$) or if

$$\phi_{it}/w_{it}>a_{it}/w_{it}+(1-\theta_{it}). \tag{A.13}$$

This condition clearly illustrates that an individual is more likely to self-isolate if the relative cost of contracting the disease, ϕ_{it}/w_{it}, is higher than the inconvenience cost of self-isolating plus the proportion of wages being lost due to self-isolation. Also, an individual is more likely to self-isolate voluntarily if the wage loss, measured by θ_{it}, is low thus showing that compensating some workers for the loss of their wages encourages a larger fraction of the population to comply with mandatory

social distancing. The above formulation could easily accommodate the differential incentive to self-isolate across different age groups and sectors of economic activity. Given that the epidemic affects the young and the old differently, with the old being more at risk as compared to the young, then $\phi_{old} > \phi_{young}$, and the old are more likely to self-isolate. Similarly, low-wage earners are more likely to self-isolate as compared to high-wage earners with the same preferences (ϕ_{it} and a_{it}), and facing the same transfer rates, θ_{it}. But the reverse outcome could occur if low-wage earner face a higher rate of transfer as compared to the high-wage earners.

According to this simple model the fraction of population that are willing to socially isolate voluntarily is given by

$$p_{n,t} = n^{-1} \sum_{i=1}^{n} I\left[(1-\theta_{it})w_{it} < (\phi_{it} - a_{it})\right],$$

where $I(A)$ is an indicator function that takes the value of 1 if A holds and zero otherwise. It is clear that the extent of voluntary social distancing, $p_{n,t}$, is positively related to the size of the economic support, θ_{it}, and the perceived net cost of contracting the virus, $(\phi_{it} - a_{it})$, which could rise sharply when epidemic surges and/or if better messaging by health authorities about the true costs of contracting the disease is provided. The inclusion of the economic support variable, and the threshold indicator in our empirical analysis are intended to capture such effects, with the indicator taking the value of unity when people become more fearful of catching the virus.

Online Supplement to "Social Distancing, Vaccination and Evolution of COVID-19 Transmission Rates in Europe"

Alexander Chudik M. Hashem Pesaran Alessandro Rebucci

January 29, 2022

This online supplement is organized as follows. Section S.1 outlines regions definitions. Section S.2 provides plots of cases and baseline estimates of the transmission rates and $\mathcal{R}$ numbers for (i) China and the rest of the world, (ii) major world regions, and (iii) selected large countries. Section S.3 provides comparisons of the estimated $\mathcal{R}$ numbers for $MF = 3$ and 5. Section S.4 provides comparisons of the estimated effective transmission rates for $MF = 3$ and 5. Section S.5 provides details regarding the estimation of standard errors in pooled regressions. Section S.6 provides pooled and fixed effects estimates of effective transmission rates using alternative choices of lag orders. Section S.7 presents summary statistics for the five factors in the panel data model of the of the transmission rates.

S.1 Regions definitions

Table S1: Regions definitions

East Asia and Pacific
Australia, Brunei, Cambodia, China, South Korea, Fiji, Indonesia, Japan, Laos, Malaysia, Marshall Islands, Mongolia, Burma, New Zealand, Papua New Guinea, Philippines, Samoa, Singapore, Solomon Islands, Taiwan, Thailand, Timor-Leste, Vanuatu, Vietnam
Eastern Europe and Central Asia
Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Georgia, Kazakhstan, Kyrgyzstan, Montenegro, Moldova, Romania, Russia, Serbia, Tajikistan, North Macedonia, Turkey, Ukraine, Uzbekistan
Western Europe
Andorra, Austria, Belgium, Cyprus, Czechia, Denmark, Estonia, Finland, France, Germany, Greece, Holy See, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Monaco, Netherlands, Norway, Poland, Portugal, San Marino, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom
North America
United States, Canada
Latin America and Caribbean
Antigua and Barbuda, Argentina, Bahamas, Barbados, Belize, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominica, Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Suriname, Trinidad and Tobago, Uruguay, Venezuela
Middle East and North Africa
Algeria, Bahrain, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, Syria, Tunisia, United Arab Emirates, Yemen
South Asia
Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan, Sri Lanka
SubSaharan Africa
Eastern and Southern Africa and West and Central Africa
Eastern and Southern Africa
Angola, Botswana, Burundi, Comoros, Djibouti, Eritrea, Ethiopia, Kenya, Lesotho, Madagascar, Malawi, Mozambique, Namibia, Rwanda, Seychelles, Somalia, South Africa, South Sudan, Sudan, Uganda, Tanzania, Zambia, Zimbabwe
West and Central Africa
Benin, Burkina Faso, Cabo Verde, Cameroon, Central African Republic, Chad, Congo (Brazzaville), Congo (Kinshasa), Cote d'Ivoire, Croatia, Equatorial Guinea, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Sao Tome and Principe, Senegal, Sierra Leone, Togo

S.2 COVID-19: A global pandemic with heterogeneous time-varying transmission

In this section of the online supplement we report country-specific estimates of the reproduction number, $\mathcal{R}_{et}$. We refer to it below as simply the "$\mathcal{R}$ number". We plot it alongside the effective transmission rate, $\beta_t \times 14 = \beta_t/\gamma$, to separately assess the influence of herding from social distancing, for a large sample of countries.

While we estimate the two parameters of interest for all jurisdictions for which JHU reports case statistics, in this section we report only the results for selected countries and regions.[22] Figures S.1-S.8 plot the results. Figure S.1 reports results for China and the rest of the world. Figures S.2-S.4 show results by geographic regions (excluding China): the Northern and Southern Hemispheres (Figure S.2) and all main regions of the world (Figures S.3-S.4), including East Asia and Pacific, South Asia, Eastern Europe and Central Asia, Western Europe, North America, Latin America and Caribbean, Middle East and North Africa, and Sub-Saharan Africa. Figures S.5-S.6 report results for selected large economies. Finally, Figures S.7-S.8 give results for the selected European economies also analyzed in Section 4 below. The estimates of transmission rates and $\mathcal{R}$-numbers for the regions are based on aggregate region-specific case statistics rather than averages across country specific estimates.

Each panel reports two sets of charts. The charts on the left-hand-side of the figures report the seven-day moving average of the number of reported new infected cases per 100,000 population. The charts on the right-hand-side report two lines. The solid (red) line is the estimated R-number, $\hat{\mathcal{R}}_{et} = (1 - c_t)\hat{\beta}_t/\gamma$. The dotted (blue) line is the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$. This is the variable that we model in Section 4. Recalling that the effective transmission rate, $\hat{\beta}_t/\gamma$, coincides with $\hat{\mathcal{R}}_{et}$ only when $c_t \approx 0$, but as the epidemic spreads more widely we have $c_t > 0$, herd immunity can eventually start to play a non-negligible role and manifests itself in later stages of the epidemic with an increasing gap between $\hat{\mathcal{R}}_{et}$ and $\hat{\beta}_t/\gamma$, depending on the magnitude of c_t. Also, we expect $\hat{\beta}_t/\gamma$ to be in the range 0 to 3 (similarly to $\hat{\mathcal{R}}_{et}$), and $\hat{\mathcal{R}}_{et}$ to be smaller or equal to

[22] The full set of estimation results is available on the authors' websites (sites.google.com/site/alexanderchudik/, pesaran.com, sites.google.com/site/alessandrorebucciphd/).

the effective transmission rate as the epidemic progresses. Thus the gap between the red and the blue lines is a function of $s_t = 1 - c_t$, the share of susceptible (not yet infected) population.

We start by estimating the effective transmission rate, $\hat{\beta}_t/\gamma$, and hence the $\mathcal{R}$ numbers, when the seven-day moving average of new cases exceeds a threshold of 50 cases to ensure a reasonably precise estimate of β_t/γ. Note that at the early stages of the spread of the infection, when both c_t and i_t are close to zero, estimation of β_t/γ becomes problematic as can be seen directly from (6). In effect it involves computing the ratio of two very small numbers, each subject to sampling errors. Note also that, since some countries (in particular China) were able to virtually eradicate the virus in some sub-periods, there will be gaps in our charts reporting the R numbers. In addition, we start to report estimated $\mathcal{R}$ numbers at the beginning of the sample from the day in which $\hat{\mathcal{R}}_{et} < 3$ for the first time. This is to avoid showing widely varying estimated values in the initial days of the epidemic driven by unusually large growth rates of new confirmed cases, which could reflect delays in reporting the number of infected cases.

S.2.1 China and the rest of the world

China China experienced a large first wave followed by a few small and localized outbreaks (Figure S.1). Two points are worth highlighting. First the $\mathcal{R}$ number comes down very fast, in less than a month during the first wave. This is consistent with disaggregate evidence in Fang, Wang, and Yang (2020) and also clinical evidence. Second, the effective reproduction number always coincides with the effective transmission rate in the case of China, given the fact that only a very small fraction of population has been infected. The number of infected cases in China is $90,000 \times MF$ out of a population of 1.4 billion. This is a very small share even if we set MF to 20, which is at the upper end of the estimates reported for MF across many countries and reviewed in the Introduction. This confirms herd immunity had no role in the reduction of the effective reproduction number in the case of China.[23] When the epidemic resurfaces, the estimated

[23]The effective reproduction number coincides with the effective transmission rate in most other Asian countries. Nonetheless, even in Asia, we observe a great deal of heterogeneity in terms of the shape of the epidemic curve. Japan and Indonesia fared better at the start of the pandemic, but did not avoid a large second wave. South Korea, in contrast, had two waves, one in March 2020 and a second toward the end of 2020, possibly reflecting its decision to avoid China-style mandatory social distancing, embracing a strategy revolving around testing and tracing with less

effective transmission rate increases sharply, but the extremely small number of cases permitted due to aggressive containment strategies prevented any new large-scale spread of the virus. Note that under mandatory social distancing the population never reaches herd immunity. So infections recur if containment is relaxed and the virus has not yet been fully eradicated. Figure S.1 shows that China was successful not only in containing the epidemic at the start of its outbreak, but has thus far also been able to eradicate it quickly whenever it has re-surfaced through international travel. As we shall see, most other countries have not been able to accomplish this.

The Rest of the World excluding China The bottom panel of Figure S.1 reports results for the rest of the world excluding China. As we noted earlier, these estimates are based on aggregate cases, as opposed to averages of country specific estimates. In the rest of the world, the COVID-19 epidemic started later than in China and the $\mathcal{R}$ number comes down more slowly compared to China, never really falling below one until the end of 2020. The $\mathcal{R}$ number increased from May to July 2020, and then again starting at the end of August 2020. As a result, the pandemic's incidence was many, many times higher than in China in terms of cases. Indeed, our estimation results show that even an $\mathcal{R}$ number slightly above one can be devastating once the epidemic has spread widely. Overall, the rest of the world as a whole never managed to eradicate the epidemic to an extent comparable to China. Not surprisingly, as restrictions ease during the summer of 2020, the epidemic resurfaces and worsens dramatically. Moreover, some of the decline in the $\mathcal{R}$ number is due to herd immunity, which is extremely costly in terms of lives and, possibly, long term health consequences for the population.

S.2.2 Major World Regions

Comparing Northern and Southern Hemispheres reported in Figure S.2, we see that climate has made a difference to both the initial spread, which was faster in the northern winter, and the shape of the epidemic curve, which was more persistent in the southern hemisphere. It does not, however, make a significant difference in terms of the epidemic peak; the number of daily new confirmed cases

restrictive limits on mobility and interactions (results not reported but available from the authors).

peaked about 10-12 per 100k population in the Northern Hemisphere, whereas the peak number of new cases (per 100k population) was about 10 the Southern Hemisphere in January of 2021. In the South, the $\mathcal{R}$ number declined more slowly, but eventually dropped below one for several months in the middle of 2020. In both hemispheres, the estimates suggest that the COVID-19 transmission rate was falling in February 2021.

Figures S.3-S.4 report the estimates at more regional levels of disaggregation.[24] A stark difference emerges between the epidemic peaks reported in the left charts. North America reached a peak of 70 new cases per 100k population. Western Europe together with Eastern Europe and Central Asia experienced the second largest peaks at about 45-50 new cases per 100k population. Peaks in the daily new cases in Latin America and Caribbean region are also quite sizeable, but considerably smaller, staying below 20 new cases per 100k population. In contrast, the largest peak in the daily new cases is only about 7 in Middle East and North Africa, about 5 in South Asia, and even smaller peaks of less than 3 new cases per 100k population were achieved in East Asia and Pacific (excl. China) and Sub-Saharan Africa.

Large differences can be observed not only in terms of the magnitude of the peaks in new infections, but also in the trajectory of the epidemic more broadly. South Asia experienced a protracted single peak culminating in September 2021, which is reflected in the overall $\mathcal{R}$ number not falling below one from the start of the epidemic until early in September 2020. By contrast, Sub-Saharan Africa experienced two definite peaks (July 2020 and January 2021). North America and Western Europe experienced three major waves. The first wave occurred in March/April in both regions. After some significant community spread of the virus, containment policies were enacted which helped to bring the $\mathcal{R}$ number below one in a very short period of time. In North America, containment measures were relaxed quicker, and therefore the $\mathcal{R}$ number did not stay below one for long, resulting in the second wave in the summer of 2020. By contrast, the $\mathcal{R}$ number stayed below one for longer in Western Europe, until about mid-summer, when the virus began to spread exponentially again, resulting in the second (and largest) European wave in the Fall. After the new containment measures, $\mathcal{R}$ number declined again, but it did not stay below one for long,

[24] Table A1 in the online Appendix lists countries included in each region.

resulting in the third wave of infections in January 2021 in both regions.

Experience from the remaining regions is more atypical than one might expect from the epidemic models. New cases in the Middle East and North Africa and, to some extent East Asia and Pacific (excl China), exhibit a broad upward trend throughout 2020 with a number of local peaks; new cases data for Latin America and Caribbean appear to be subject to much more noise compared with any other regions, and there is an unusual jump in the daily new cases in Eastern Europe and Central Asia, driven by the data for Turkey. $\mathcal{R}$ numbers closely reflect the first derivative of the smoothed version of the new cases data in all regions; new cases subside when R falls below one and increase when R is above one.

The difference between the solid red lines ($\mathcal{R}$ numbers) and the dotted blue lines (effective transmission rate) is virtually zero in the most successful regions in terms of the total number of cases, such as Sub-Saharan Africa and South Asia, suggesting that herd immunity played no role in these regions due to the relatively small number of overall infections. On the other hand, the gap between the two lines is largest in North America, followed by Western Europe, showing that herd immunity has started to contribute more meaningfully to mitigation of the epidemic in these regions starting in December 2020.

S.2.3 Selected Large Countries

Clearly the trajectory of the epidemics has been quite heterogeneous across regions. In addition, there are considerable differences across countries within each region, to which we now turn for selected large countries. We report estimates for the United States, Brazil, India and Russia in Figure S.5, for South Africa, Australia, Iran and Turkey in Figure S.6, and nine European countries in Figures S.7-S.8—Belgium, France, Germany, Italy, Netherlands, Poland, Portugal, Spain, and UK. The selected countries include most of the G20 economies with the widest regional coverage globally.

In contrast to China's and the rest of the world, the United States (reported in the top panel of Figure S.5) stands out for the largest gap between the effective reproduction number and the effective transmission rate since the reopening of the economy in May 2020. The gap continues to

widen throughout the subsequent period, peaking at the end of the sample in February 2021. Only a few other countries in the world, including the United Kingdom, Israel and some Latin American countries, display a comparable contribution of herd immunity to the decline in the $\mathcal{R}$ number. The US case also stands out because of the three very distinct waves, with the second and the third re-emerging after a brief fall of the $\mathcal{R}$ number below one. This led to a much higher number of infections per 100,000 people compared to the rest of the world.

Like the United States, Brazil's estimates also show visible gaps between the $\mathcal{R}$ number and the effective transmission rate starting in mid-2020. The case count in Brazil is more volatile compared to the United States and the remaining countries, possibly due to differences in the data quality other than under-reporting controlled for with the multiplication factor. Unlike the US case, Brazil brought down the $\mathcal{R}$ number more gradually, falling below one for the first time only during the summer of 2020. This resulted in a protracted first wave that peaked in August. The $\mathcal{R}$ number however did not remain below one for long, and in November a second large wave took off.

India also experienced a protracted first wave. Estimates of the $\mathcal{R}$ number in India stayed above one until late September. Nevertheless, India did not experience a large number of cases per 100k population, compared with the remaining countries. As a result, herd immunity has not played a role in India. Russia, by contrast, experienced two large waves. Similarly to the western countries, Russia managed to bring the $\mathcal{R}$ number down relatively fast, but not permanently, resulting in a larger second wave at the end of 2020.

A two-wave epidemic trajectory is also observed in the case of South Africa and Australia (in Figure S.6), but with a different time profile. The first wave of the epidemic peaked in July 2020 in South Africa as authorities were unable to bring the $\mathcal{R}$ number below one quickly enough. South Africa, as the richest country in the region, stands out with much higher infection rates compared to the rest of Africa. Australia, on the other hand, managed the virus very well. We can see two small peaks, one in March and the second in July-August 2020, each followed by a rapid decline in the $\mathcal{R}$ number well below one, each time almost eradicating the virus without any discernible contribution from herd immunity.

For the two major neighboring countries, Iran and Turkey in the Middle East (the bottom two

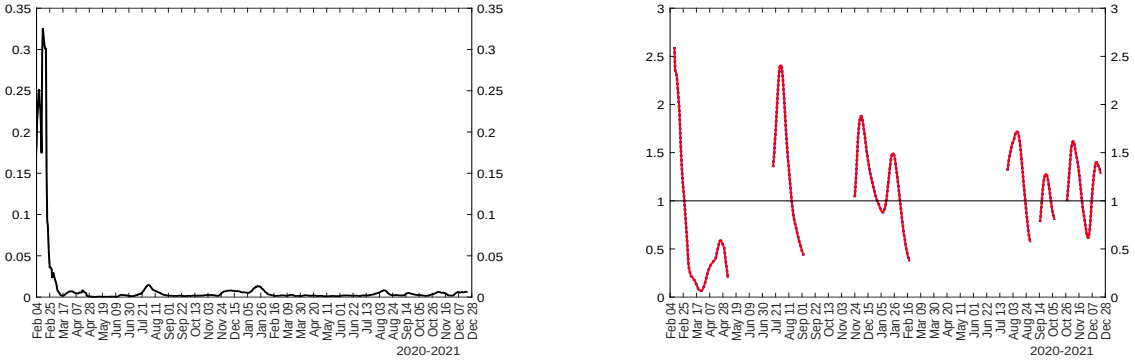
panels of Figure S.6), the trajectories of the number of new cases differ markedly, with the outbreak of the virus starting much earlier in Iran due to the close trading relations with China. The initial spread in Iran began in late February 2020 and peaked in late March after the Iranian New Year (20^{th} March) and then declined slightly before starting to move up to its second peak in November 2020. By contrast, new cases in Turkey were detected in March and remained low for quite a few months before rising dramatically to a peak of 165 per 100,000 in December 2020. The associated $\mathcal{R}$ numbers for Iran and Turkey also show very different trajectories, with Turkey's $\mathcal{R}$ number hitting the maximum value of 3 during the December 2020 peak.

The estimation results for selected European countries are reported in Figures S.7-S.8. We report the same sample of countries as the one used in the next section for panel estimation of the transmission rate determinants. The virus outbreak in continental Europe begins with Italy in early 2020, with the recorded number of infections accelerating rapidly from February 21, 2020 onward. A rapid rise in infections takes place about one week later in Spain, Germany and France, followed by Austria (not reported) at the end of February. As the rolling estimates show, the $\mathcal{R}$ number fell below one in mid- to late-April in all these countries. As lock-downs were eased during the summer, however, the transmission rates started to rise again. By the end of the 2020, the R numbers were much more dispersed, with some countries doing better than others. However, all large European countries reported in Figures S.7-S.8 show a second wave much larger than the first one. The United Kingdom, Spain, Portugal and Netherlands exhibit distinct third waves, with larger case counts compared with their second-waves.

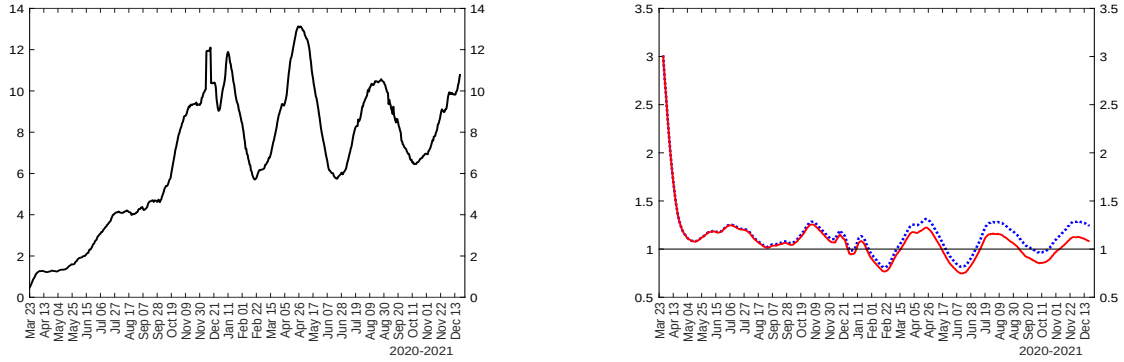
In summary, only China and a few other countries have been successful in containing the COVID-19 epidemic well. Contrary to common perception, however, not all countries accomplished this with the same draconian mandatory social distancing as in China. So we now turn to explaining the effective transmission rates to better understand the heterogeneity that we described, focusing on selected European countries reported in Figures S.7-S.8, all experiencing quite similar starting dates and the initial wave of the epidemic, but quite differing subsequent trajectories.

Figure S.1: New cases (left) and $\mathcal{R}$ numbers (right) for China and the rest of the world

China

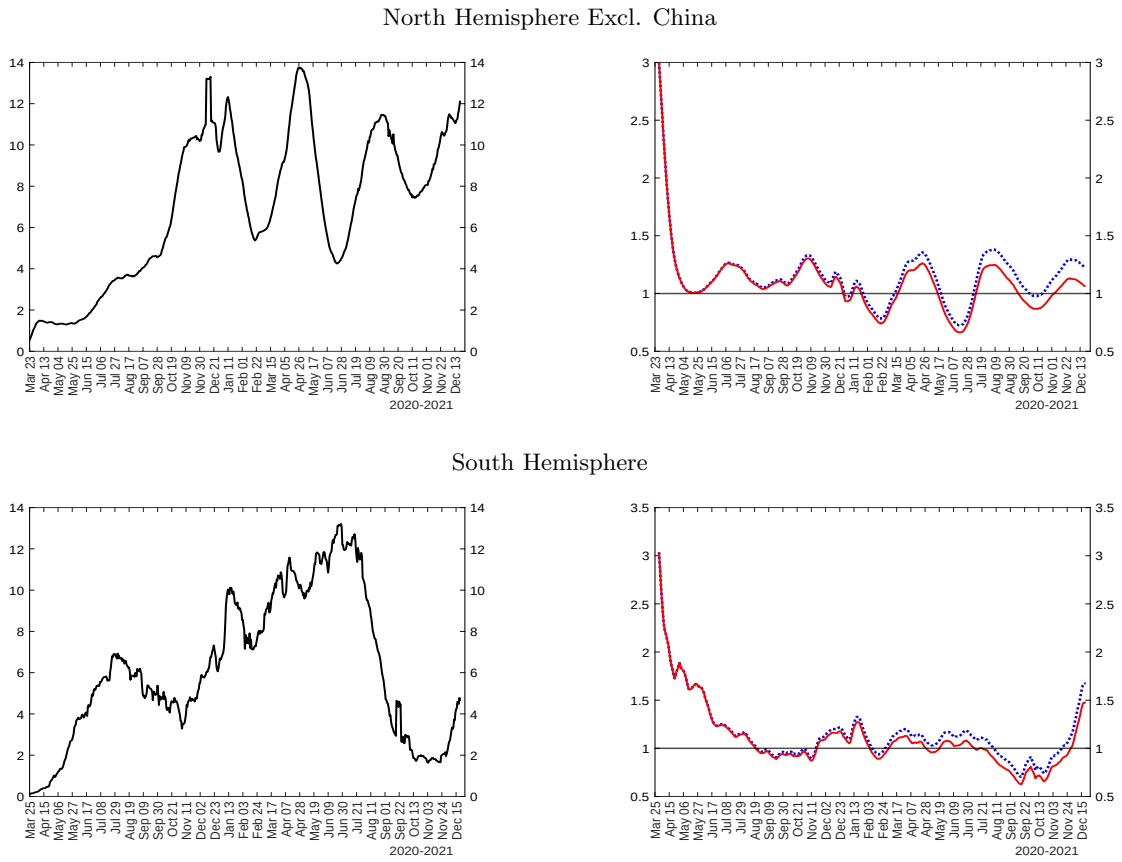


World Excl. China



Notes: The figure plots a seven-day moving average of the number of reported new cases per 100k population (left charts), the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$ (right charts, solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (right charts, dotted blue line). $\hat{\mathcal{R}}_{et} = (1 - \text{MF}\tilde{c}_t)\,\hat{\beta}_t/\gamma$, where $\gamma = 1/14$, and MF=3 for each country. $\hat{\beta}_t$ is estimated using (12), where the number of active infections is computed using the data on confirmed cases minus imputed removed cases. The number of removed (recoveries + deaths) is imputed recursively using $R_t = (1-\gamma)\,R_{t-1} + \gamma C_{t-1}$ for all countries.

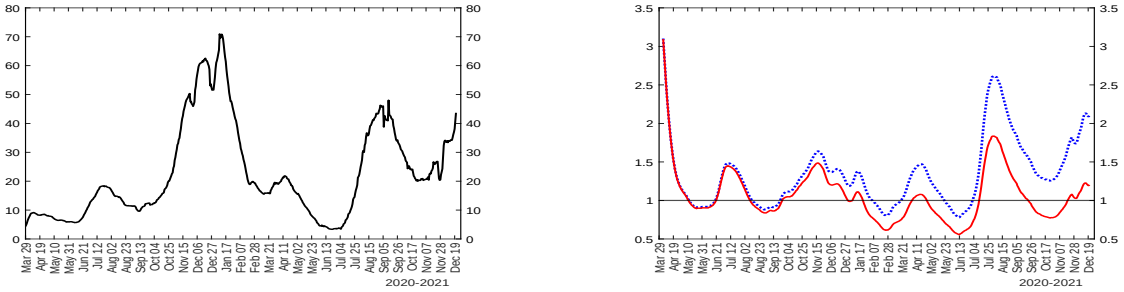
Figure S.2: New cases (left) and $\mathcal{R}$ numbers (right) for North and South Hemispheres (excl. China).



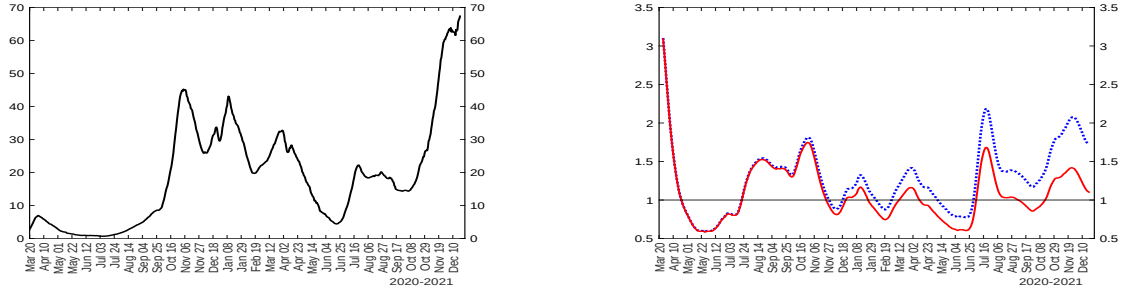
Notes: The figure plots a seven-day moving average of the number of reported new cases per 100k population (left charts), the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$ (right charts, solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (right charts, dotted blue line). See notes to Figure S.1.

Figure S.3: New cases (left) and $\mathcal{R}$ numbers (right) for selected geographic regions

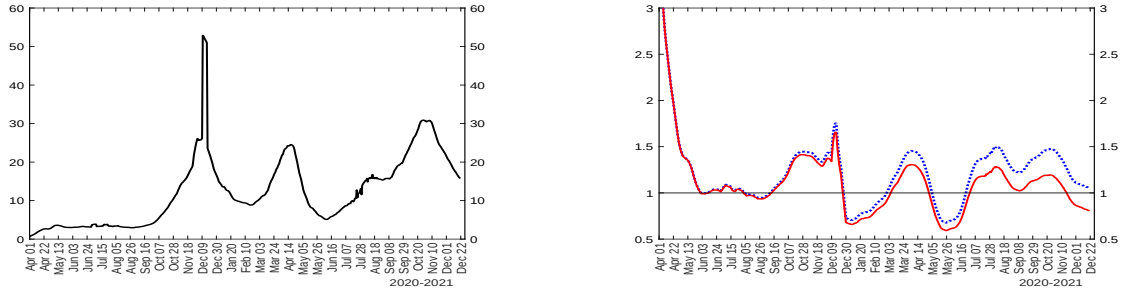
North America



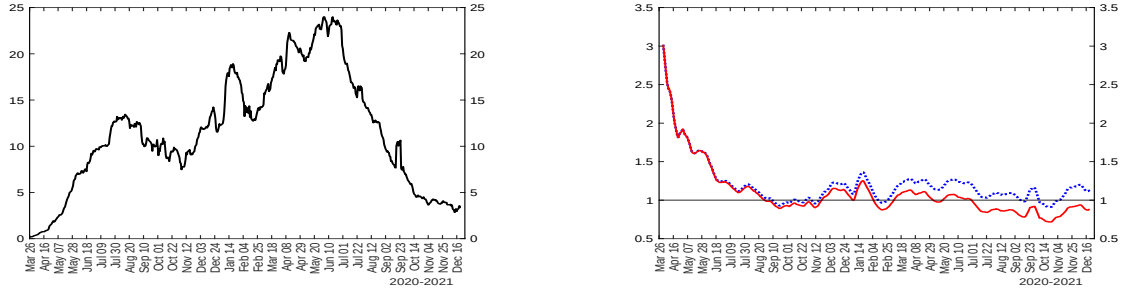
Western Europe



Eastern Europe and Central Asia



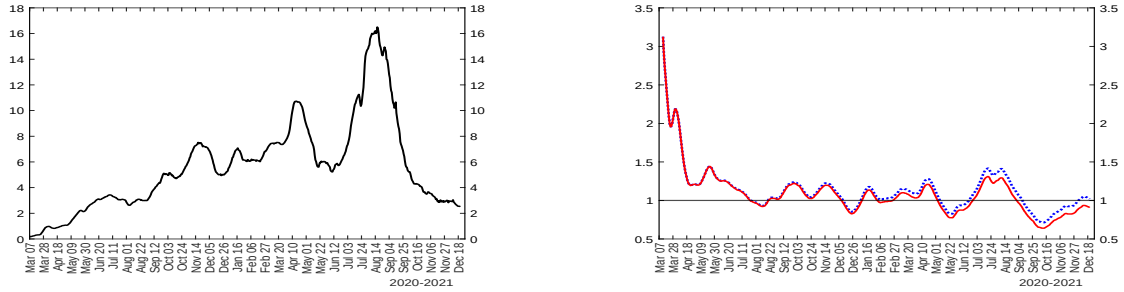
Latin America and Caribbean



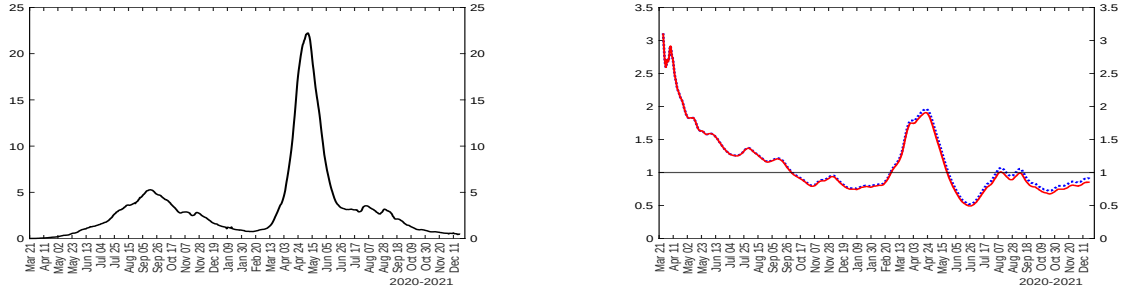
Notes: The figure plots a seven-day moving average of the number of reported new cases per 100k population (left charts), the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$ (right charts, solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (right charts, dotted blue line). See notes to Figure S.1.

Figure S.4: New cases (left) and $\mathcal{R}$ numbers (right) for selected geographic regions

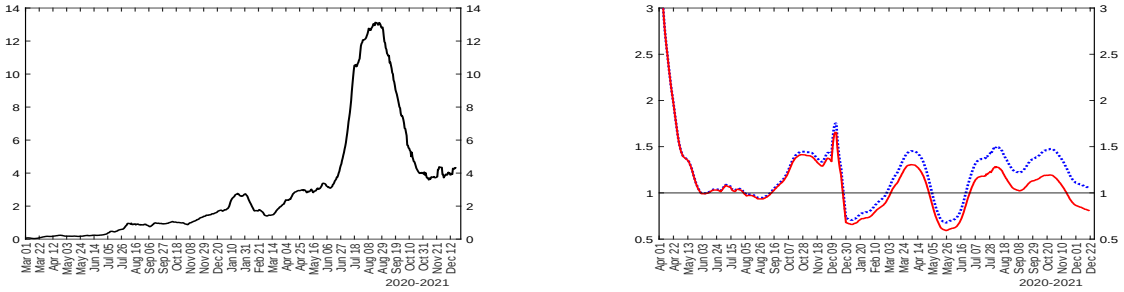
Middle East and North Africa



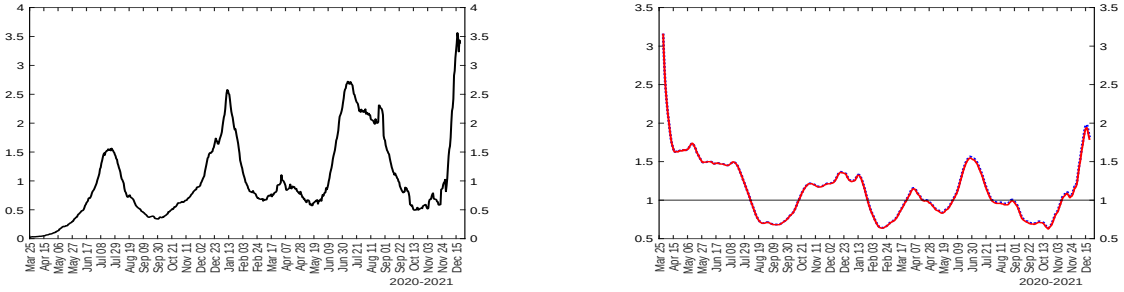
South Asia



East Asia and Pacific (Excl. China)

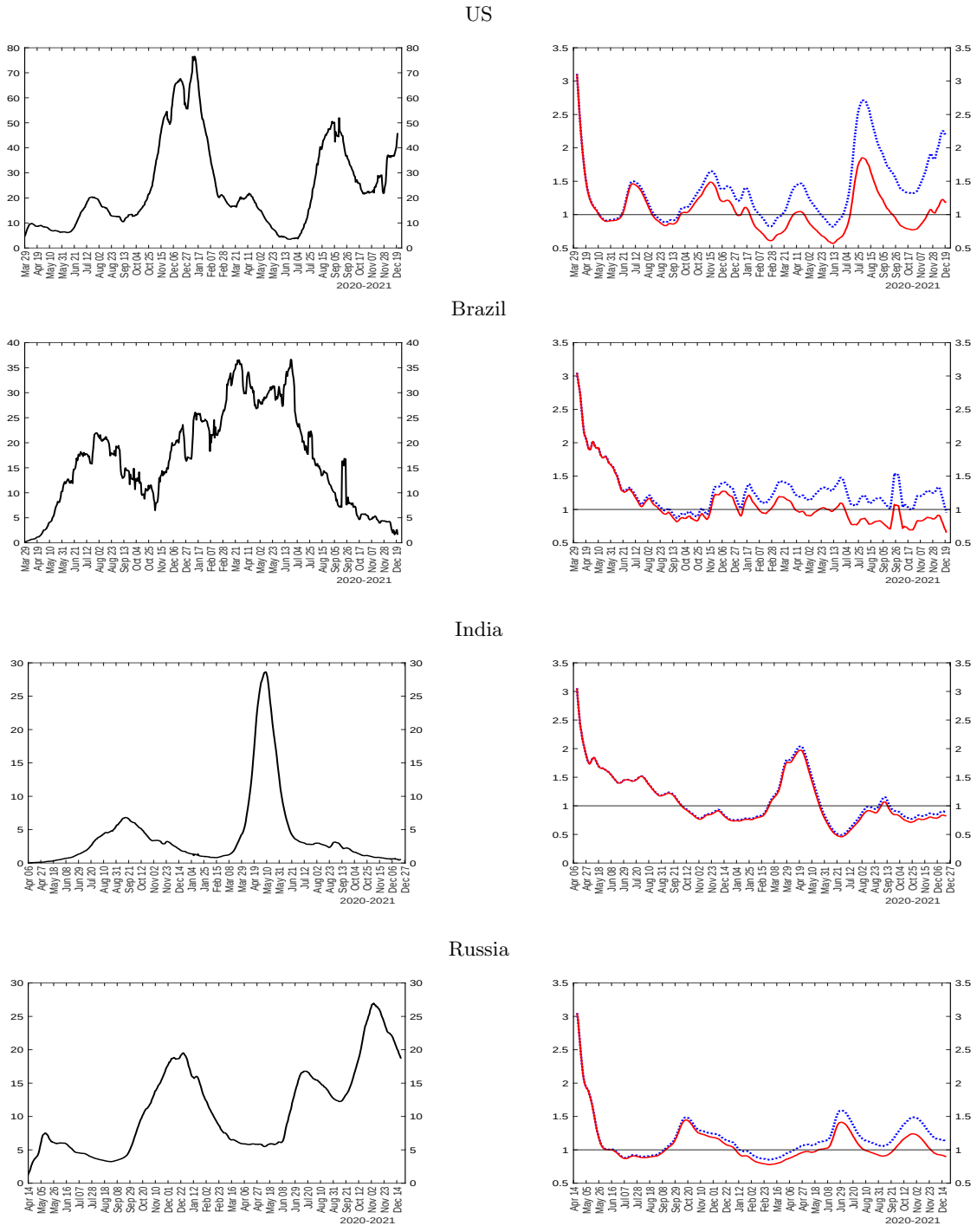


Sub-Saharan Africa



Notes: The figure plots a seven-day moving average of the number of reported new cases per 100k population (left charts), the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$ (right charts, solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (right charts, dotted blue line). See notes to Figure S.1.

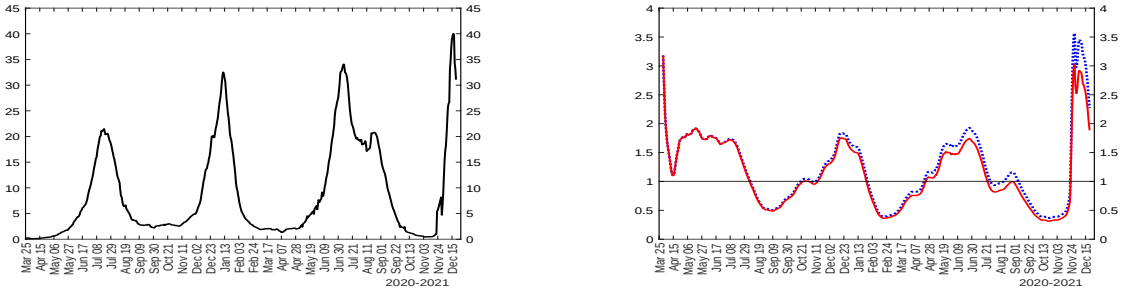
Figure S.5: New cases (left) and $\mathcal{R}$ numbers (right) for selected countries



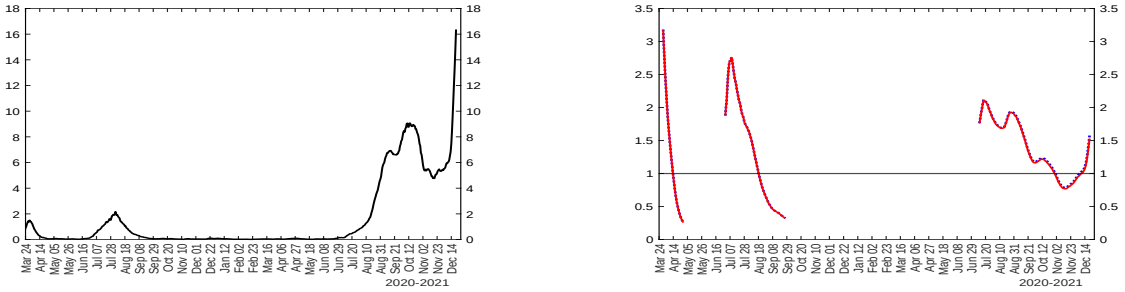
Notes: The figure plots a seven-day moving average of the number of reported new cases per 100k population (left charts), the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$ (right charts, solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (right charts, dotted blue line). See notes to Figure S.1.

Figure S.6: New cases (left) and $\mathcal{R}$ numbers (right) for selected countries

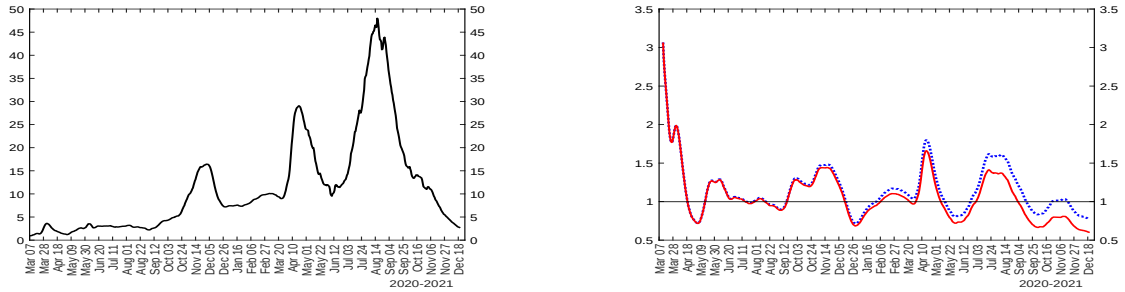
South Africa



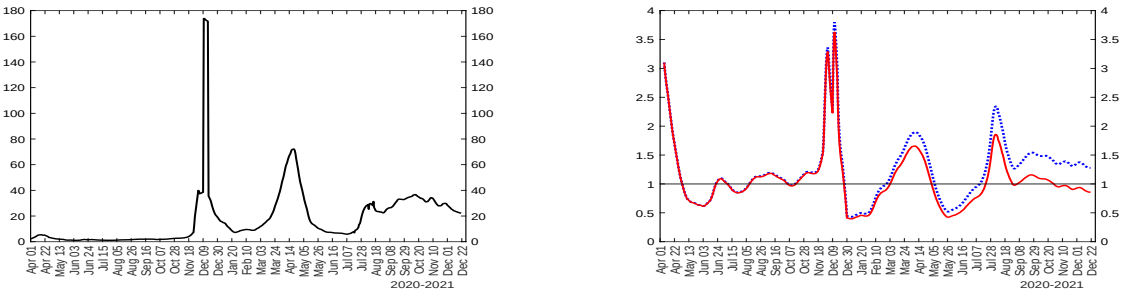
Australia



Iran

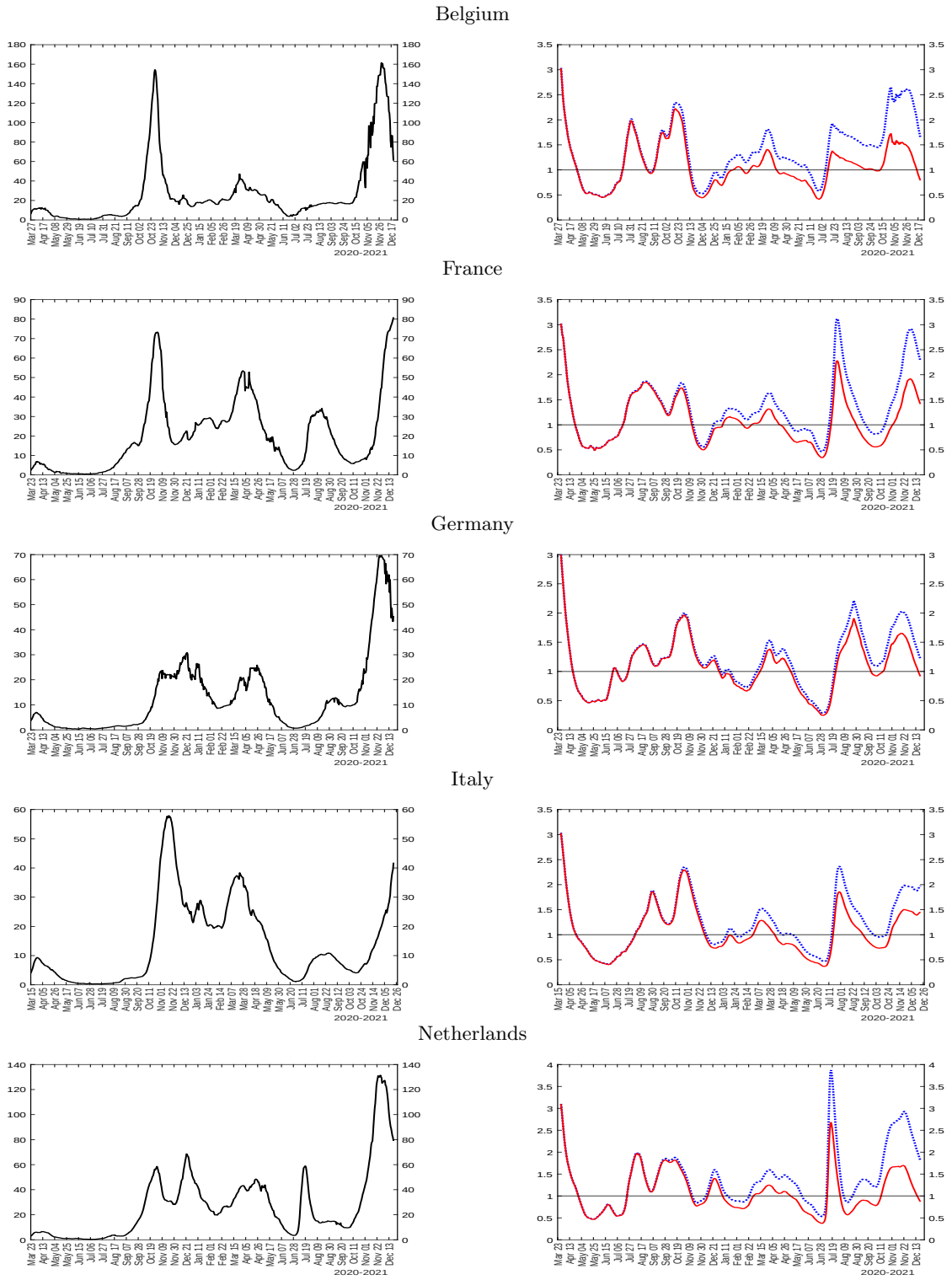


Turkey



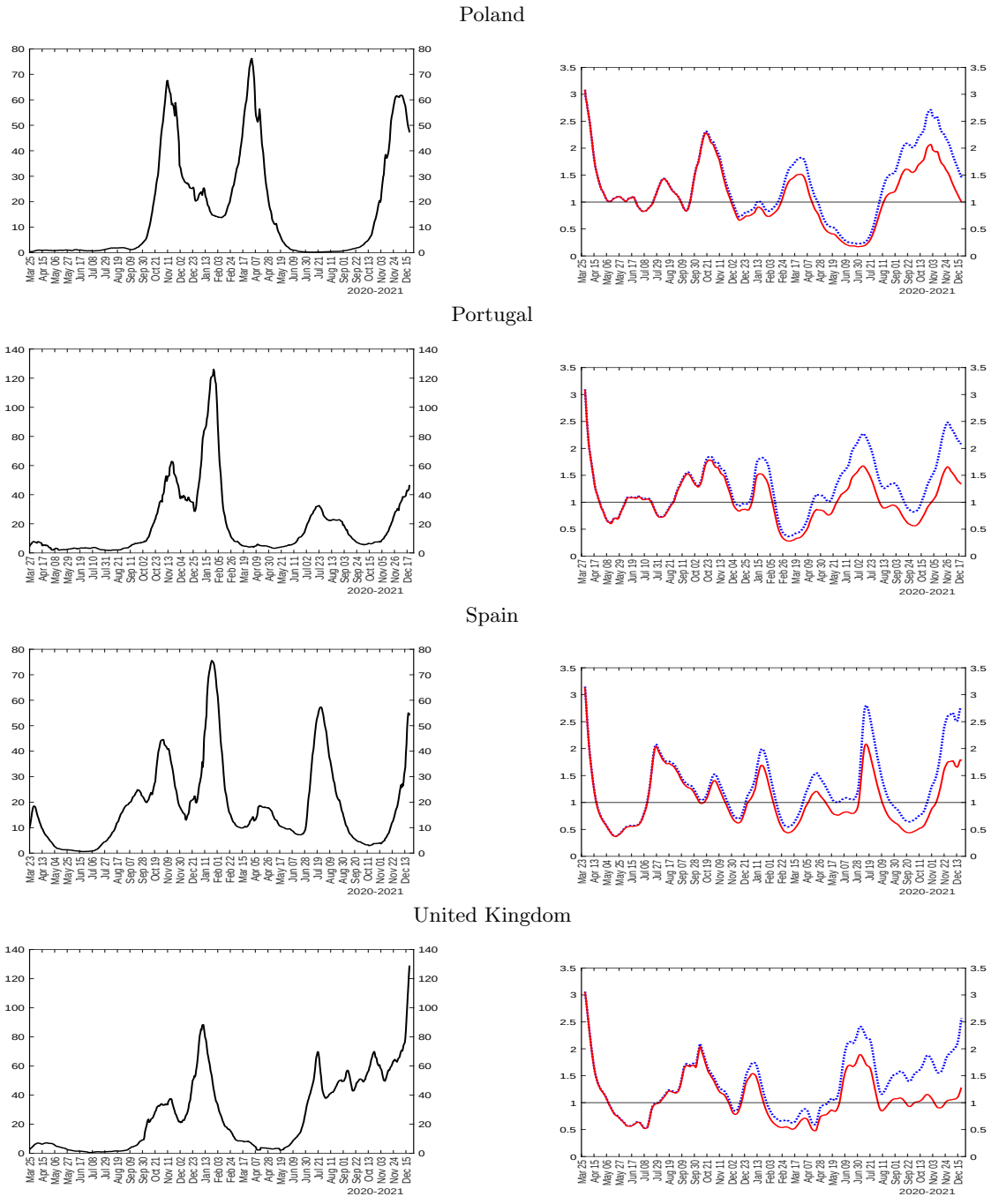
Notes: The figure plots a seven-day moving average of the number of reported new cases per 100k population (left charts), the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$ (right charts, solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (right charts, dotted blue line). See notes to Figure S.1.

Figure S.7: New cases (left) and $\mathcal{R}$ numbers (right) for sample of European countries



Notes: The figure plots a seven-day moving average of the number of reported new cases per 100k population (left charts), the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$ (right charts, solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (right charts, dotted blue line). See notes to Figure S.1.

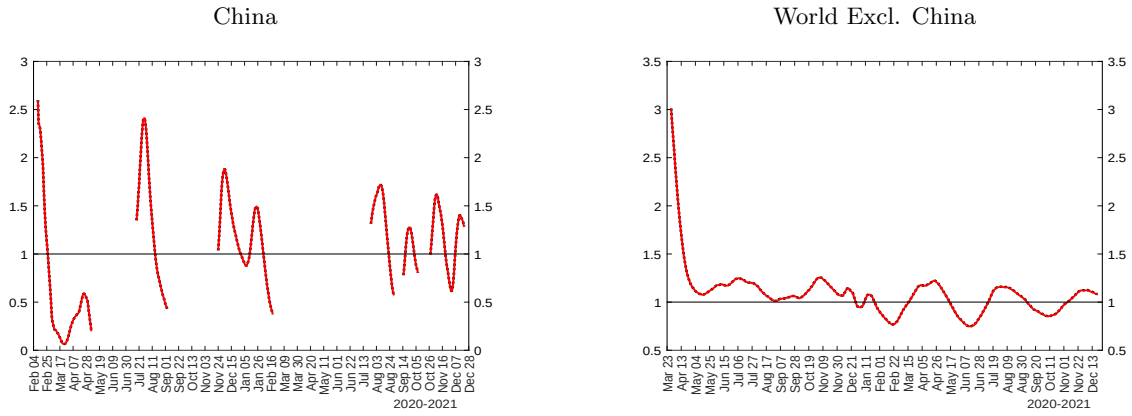
Figure S.8: New cases (left) and $\mathcal{R}$ numbers (right) for sample of European countries



Notes: The figure plots a seven-day moving average of the number of reported new cases per 100k population (left charts), the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$ (right charts, solid red line), and the effective transmission rate, $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ (right charts, dotted blue line). See notes to Figure S.1.

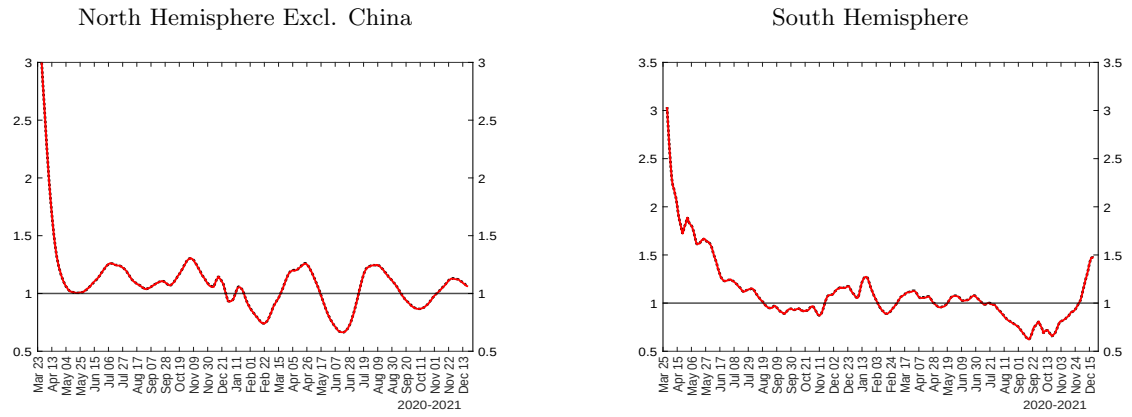
S.3 Comparison of estimated $\mathcal{R}$ numbers for selected countries and regions for two choices of multiplication factors, MF=5 and MF=3

Figure S.9: Comparisons of estimated $\mathcal{R}$ numbers for China and the rest of the world for two choices of multiplication factors, MF=5 (solid red line) and MF=3 (dotted black line)



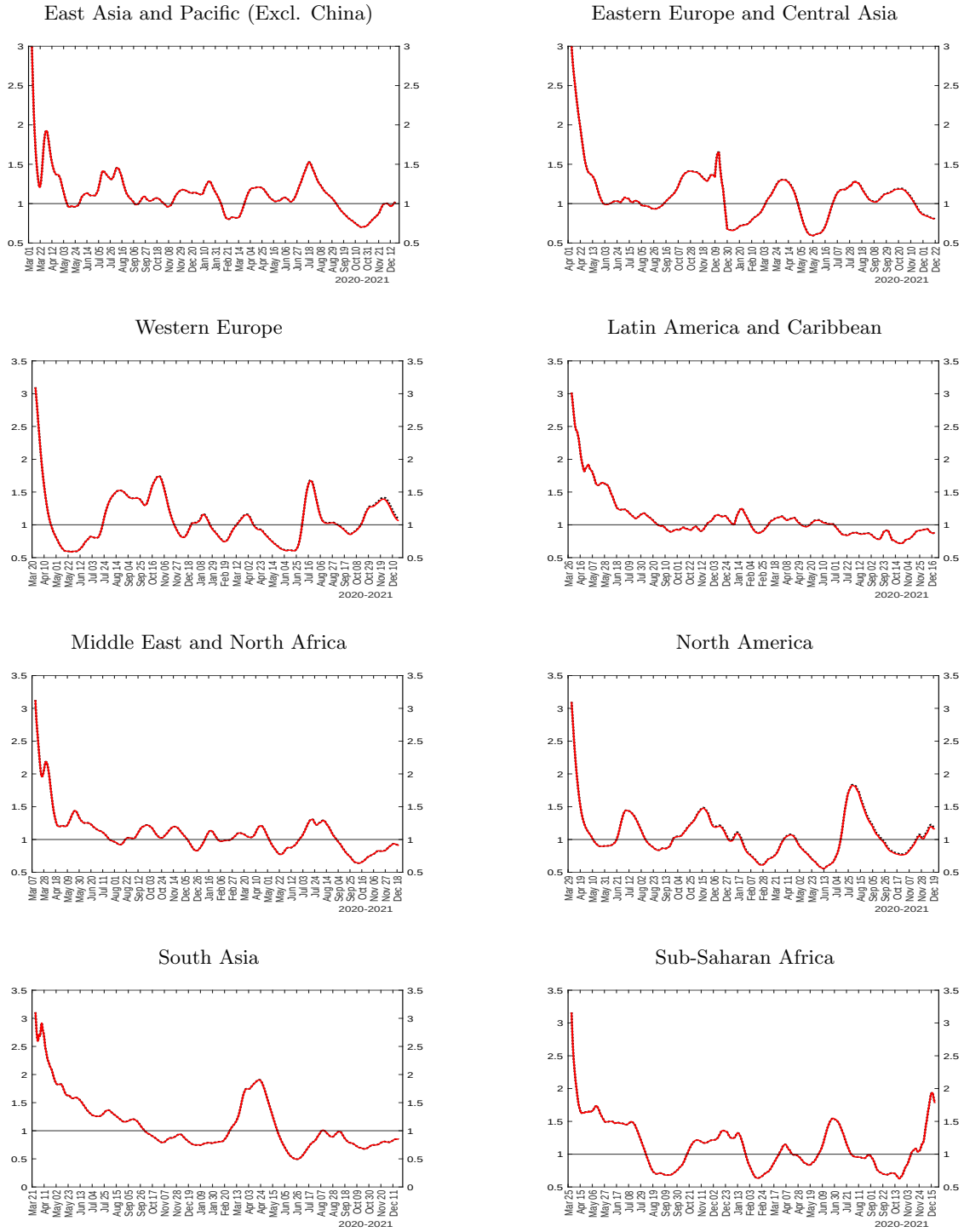
Notes: The figure plots the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$, using MF=5 (solid red line) and MF=3 (dotted black line).

Figure S.10: Comparisons of estimated $\mathcal{R}$ numbers for North and South Hemispheres (excl. China) for two choices of multiplication factors, MF=5 (solid red line) and MF=3 (dotted black line)



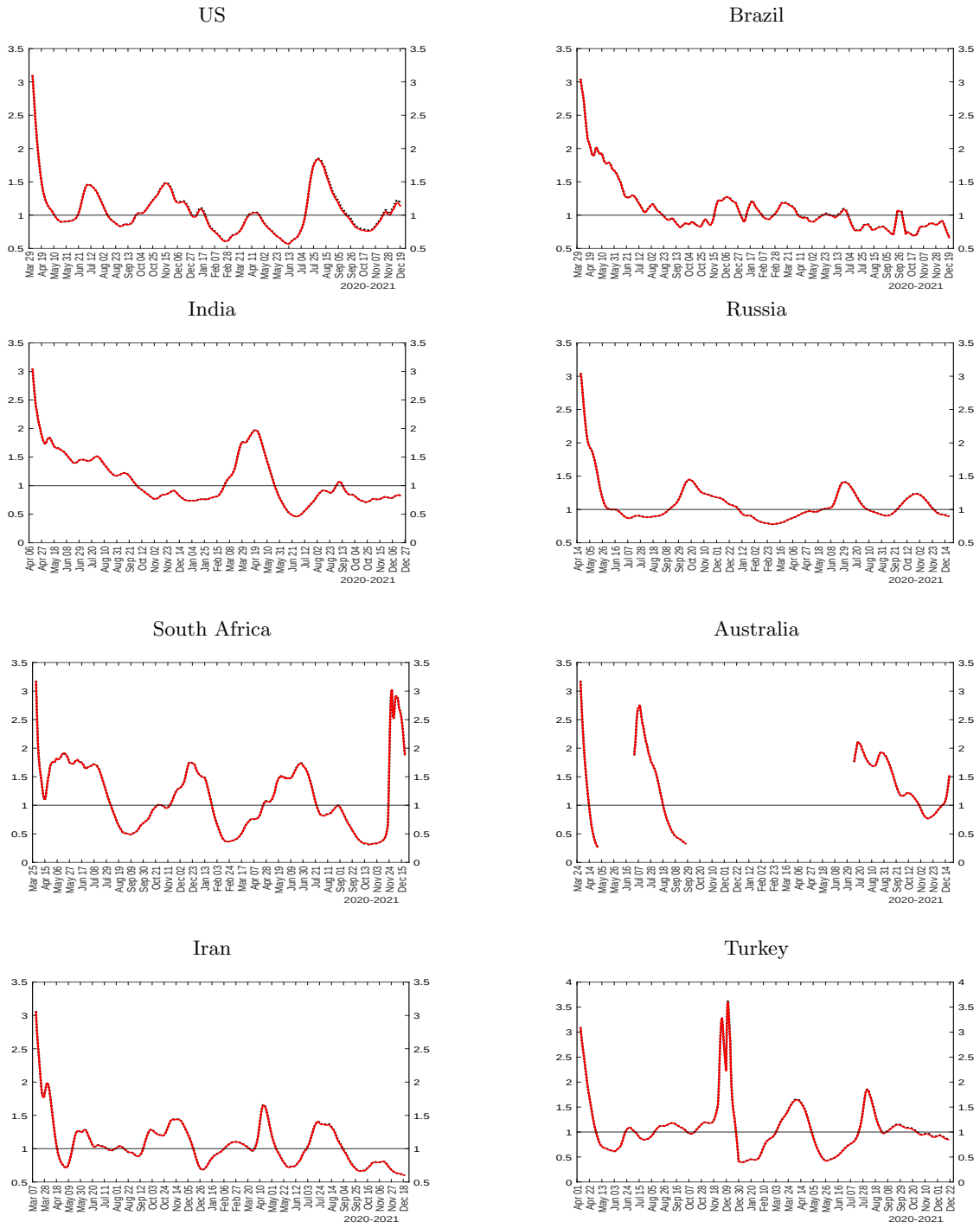
Notes: The figure plots the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$, using MF=5 (solid red line) and MF=3 (dotted black line).

Figure S.11: Comparisons of estimated $\mathcal{R}$ numbers for main geographic regions (excl. China) for two choices of multiplication factors, MF=5 (solid red line) and MF=3 (dotted black line)



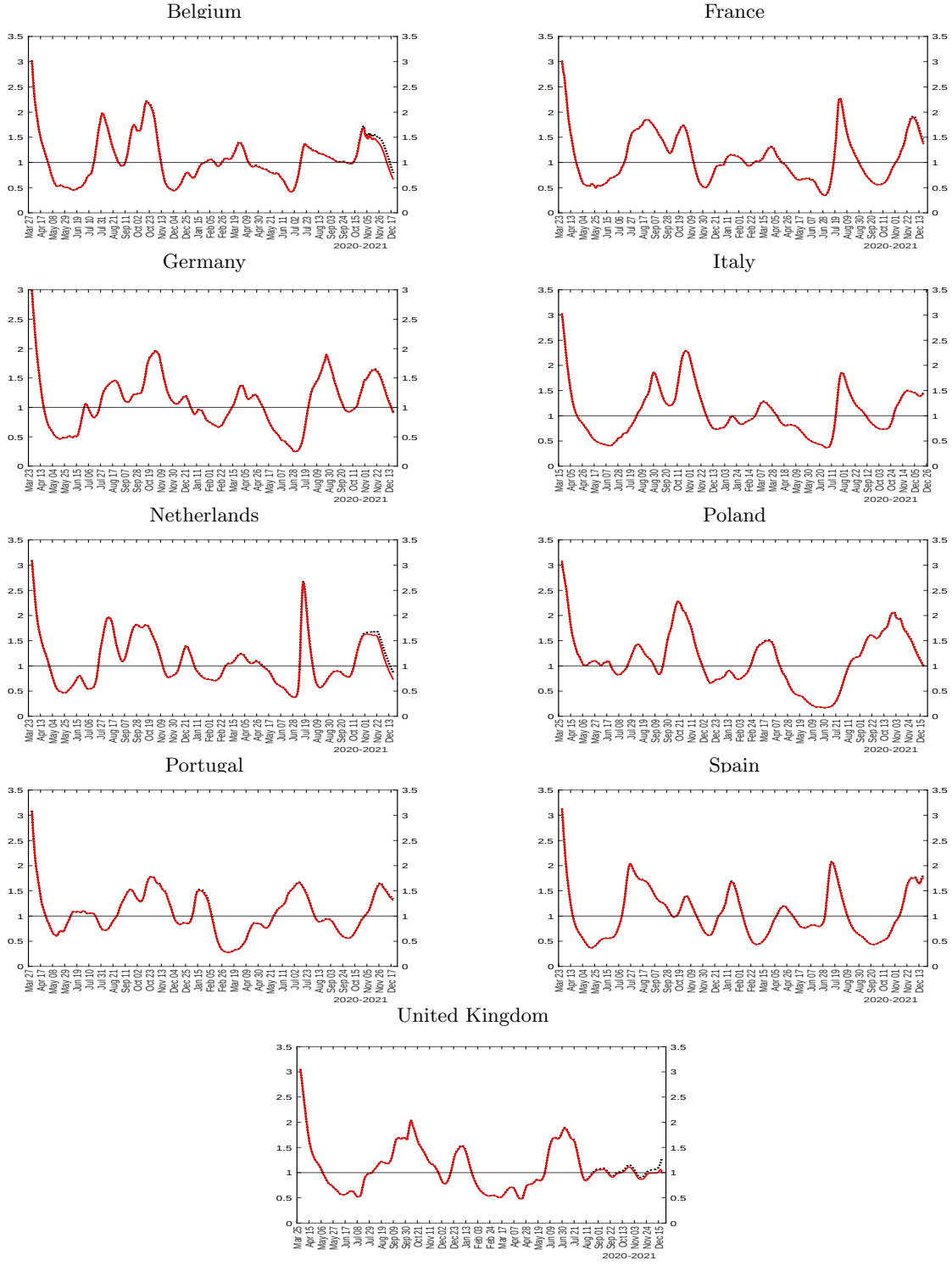
Notes: The figure plots the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$, using MF=5 (solid red line) and MF=3 (dotted black line).

Figure S.12: Comparisons of estimated $\mathcal{R}$ numbers for selected countries for two choices of multiplication factors, MF=5 (solid red line) and MF=3 (dotted black line)



Notes: The figure plots the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$, using MF=5 (solid red line) and MF=3 (dotted black line).

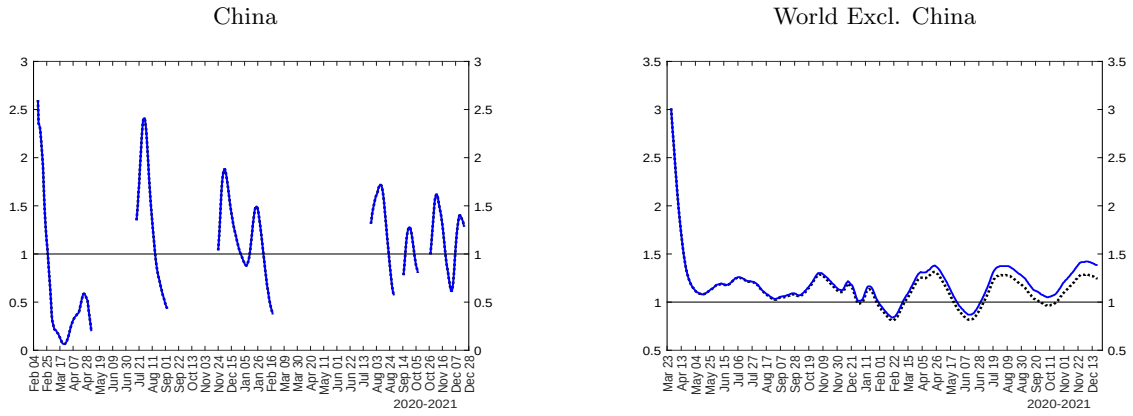
Figure S.13: Comparisons of estimated $\mathcal{R}$ numbers for sample of European countries for two choices of multiplication factors, MF=5 (solid red line) and MF=3 (dotted black line)



Notes: The figure plots the $\mathcal{R}$ number, $\hat{\mathcal{R}}_{et}$, using MF=5 (solid red line) and MF=3 (dotted black line).

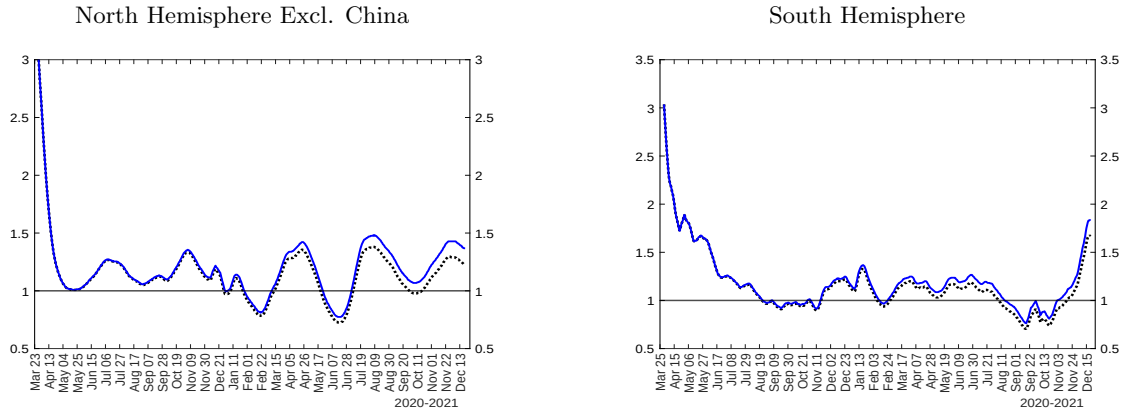
S.4 Comparison of estimated transmission rates for selected countries for two choices of multiplication factors, MF=5 and MF=3

Figure S.14: Comparison of estimated transmission rates for China and the rest of the world for two choices of multiplication factors, MF=5 (solid blue line) and MF=3 (dotted black line)



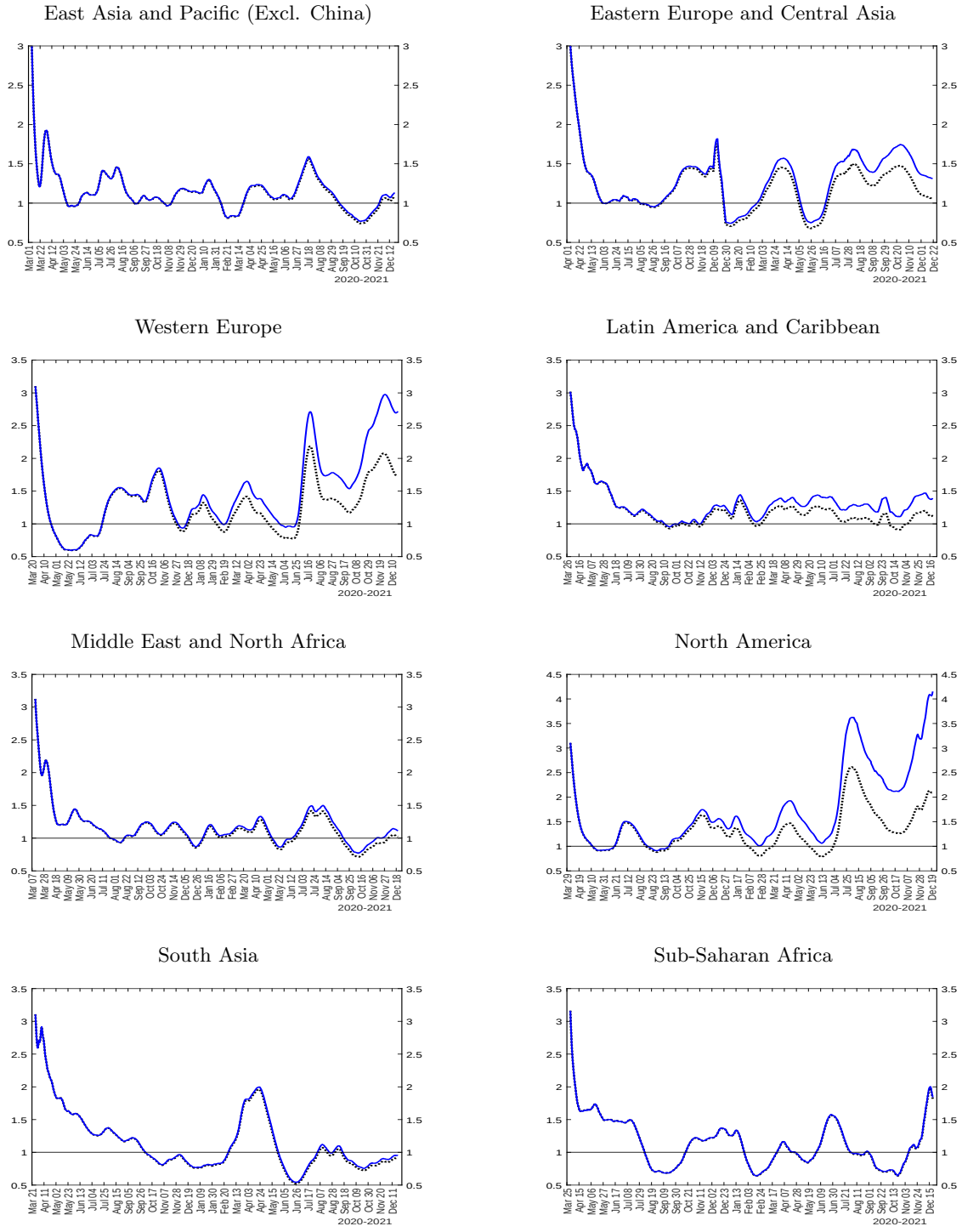
Notes: The figure plots the effective transmission rate $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ using multiplication factor MF=5 (solid blue line) and MF=3 (dotted black line).

Figure S.15: Comparison of estimated transmission rates for North and South Hemispheres (excl. China) for two choices of multiplication factors, MF=5 (solid blue line) and MF=3 (dotted black line)



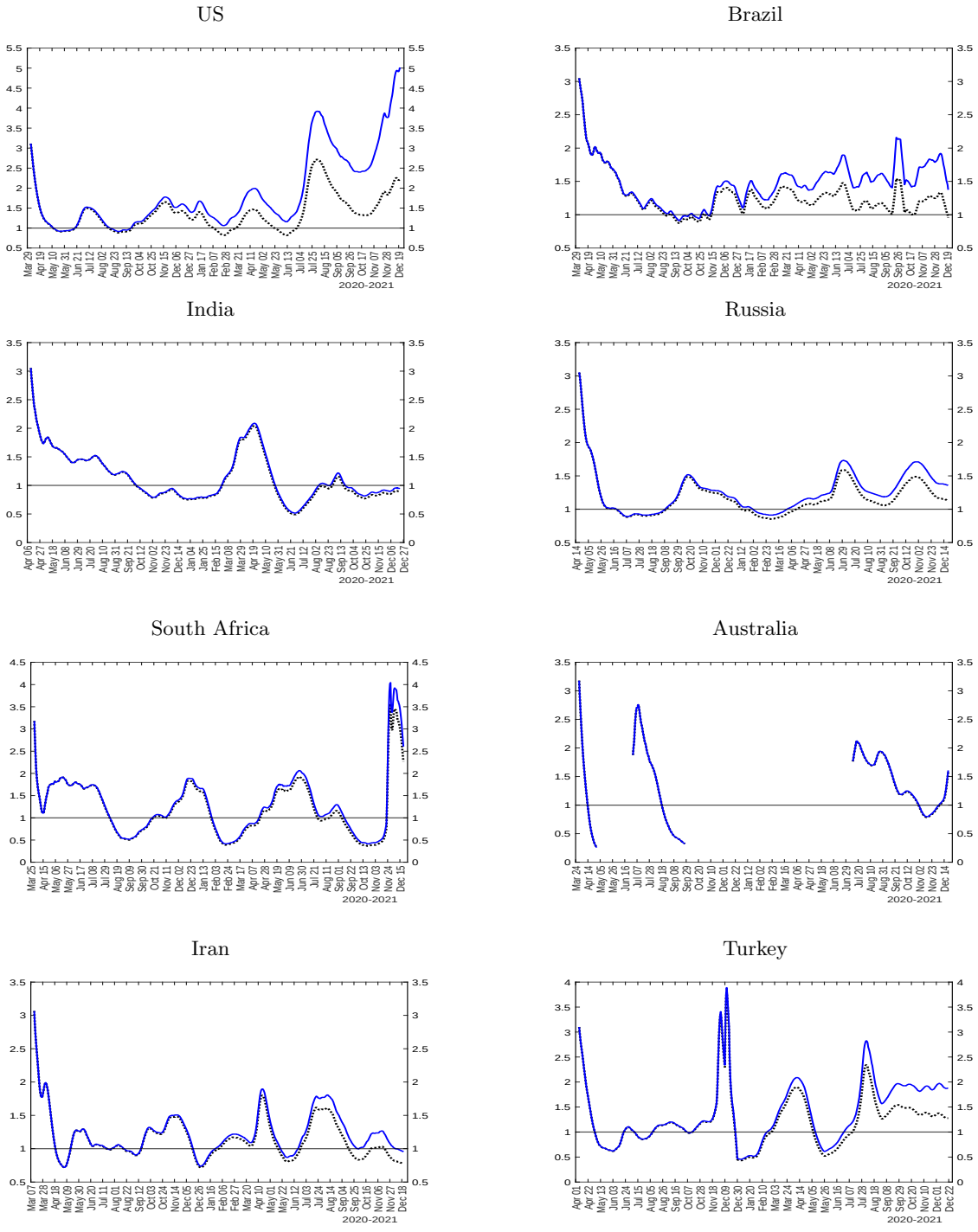
Notes: The figure plots the effective transmission rate $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ using multiplication factor MF=5 (solid blue line) and MF=3 (dotted black line).

Figure S.16: Comparison of estimated transmission rates for main geographic regions (excl. China) for two choices of multiplication factors, MF=5 (solid blue line) and MF=3 (dotted black line)



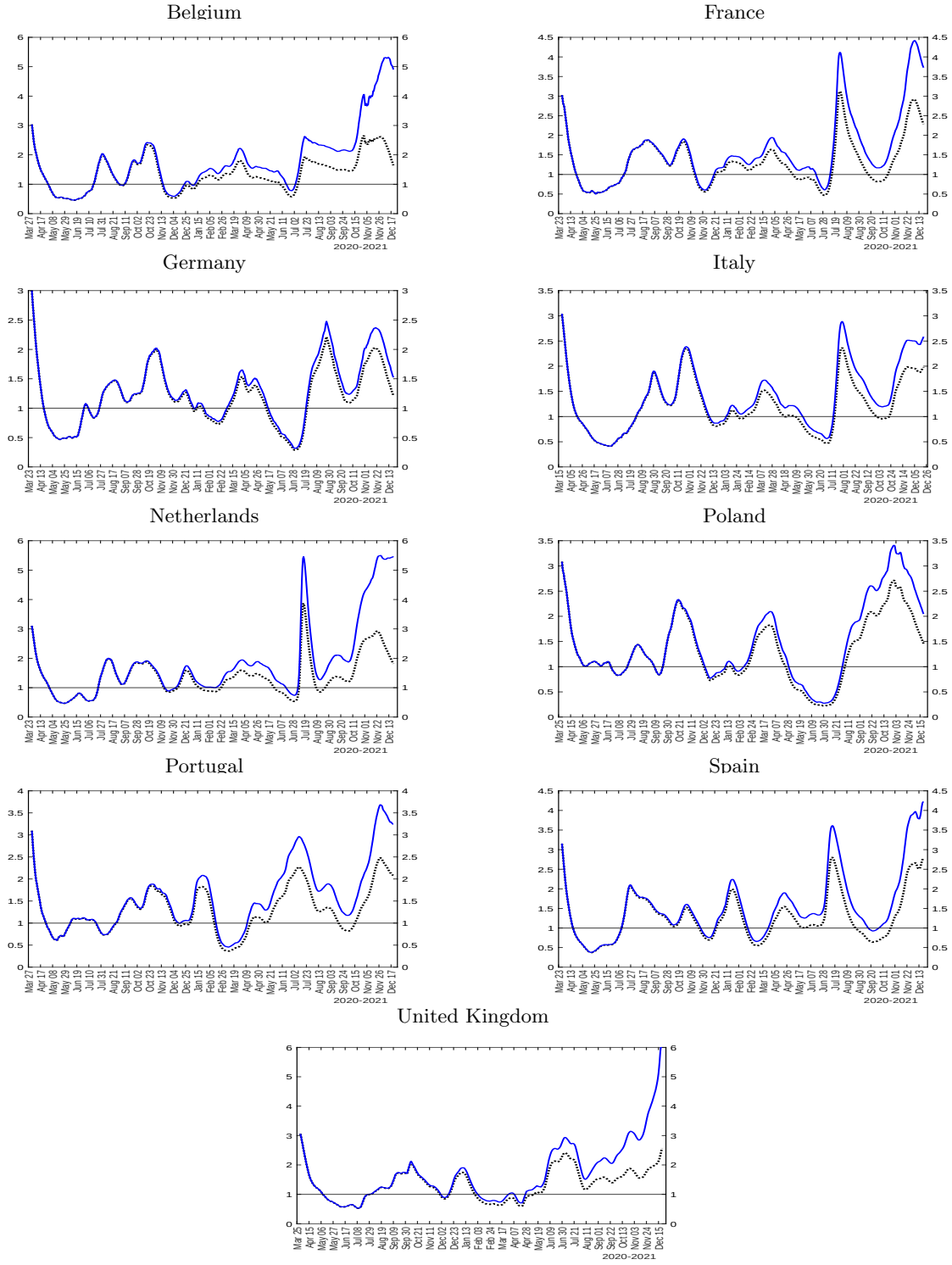
Notes: The figure plots the effective transmission rate $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ using multiplication factor MF=5 (solid blue line) and MF=3 (dotted black line).

Figure S.17: Comparison of estimated transmission rates for selected countries for two choices of multiplication factors, MF=5 (solid blue line) and MF=3 (dotted black line)



Notes: The figure plots the effective transmission rate $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ using multiplication factor MF=5 (solid blue line) and MF=3 (dotted black line).

Figure S.18: Comparison of estimated transmission rates for sample of European countries for two choices of multiplication factors, MF=5 (solid blue line) and MF=3 (dotted black line)



Notes: The figure plots the effective transmission rate $\hat{\beta}_t \times 14 = \hat{\beta}_t/\gamma$ using multiplication factor MF=5 (solid blue line) and MF=3 (dotted black line).

S.5 Conducting inference about the pooled panel results

Consider a linear panel data model, which, for convenience, can be written as

$$y_{jt} = \boldsymbol{\theta}'\boldsymbol{\zeta}_{jt} + u_{jt},$$

for $j = 1, 2, ..., N$, where $\boldsymbol{\zeta}_{jt}$ is the vector of variables (inclusive of intercept). We allow for unbalanced panel by assuming $t = 1, 2, ..., T_j$. Let $\hat{\boldsymbol{\theta}}$ be the pooled estimator. We have

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta} = \left(\sum_{j=1}^{N} \sum_{t=1}^{T_j} \boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}'_{jt} \right)^{-1} \sum_{j=1}^{N} \sum_{t=1}^{T_j} \boldsymbol{\zeta}_{jt} u_{jt}.$$

The variance of $\hat{\boldsymbol{\theta}}$ is given by

$$Var(\hat{\boldsymbol{\theta}}) = \left(\sum_{j=1}^{N} \sum_{t=1}^{T_j} \boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}'_{jt} \right)^{-1} Var \left(\sum_{j=1}^{N} \sum_{t=1}^{T_j} \boldsymbol{\zeta}_{jt} u_{jt} \right) \left(\sum_{j=1}^{N} \sum_{t=1}^{T_j} \boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}'_{jt} \right)^{-1}.$$

Assuming $E\left(\boldsymbol{\zeta}_{jt} u_{jt}\right) = 0$, we obtain

$$Var(\hat{\boldsymbol{\theta}}) = \left(\sum_{j=1}^{N} \sum_{t=1}^{T_j} \boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}'_{jt} \right)^{-1} \left(\sum_{j=1}^{N} \sum_{t=1}^{T_j} \sum_{j=1}^{N} \sum_{t'=1}^{T_j} E\left(\boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}'_{jt'} u_{jt} u_{jt'} \right) \right) \left(\sum_{j=1}^{N} \sum_{t=1}^{T_j} \boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}'_{jt} \right)^{-1}.$$

S.5.1 Inference robust to serial correlation of errors

Let

$$\mathbf{Q}_{nT} = \frac{1}{\sum_{j=1}^{N} T_j} \sum_{j=1}^{N} \sum_{t=1}^{T_j} \boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}'_{jt},$$

and

$$\mathbf{S}_{nT} = \sum_{j=1}^{N} \frac{T_j}{\sum_{h=1}^{N} T_h} \left[\frac{1}{T_j} \sum_{t=1}^{T_j} \sum_{t'=1}^{T_j} E\left(\boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}'_{jt'} u_{jt} u_{jt'} \right) \right],$$

then

$$Var(\hat{\boldsymbol{\theta}}) = \frac{1}{\sum_{j=1}^{N} T_j} \mathbf{Q}_{nT}^{-1} \mathbf{S}_{nT} \mathbf{Q}_{nT}^{-1}.$$

We estimate $\mathbf{S}_{nT}$ by the Newey-West method, extended to our panel setup:

$$\hat{\mathbf{S}}_{nT} = \sum_{j=1}^{N} \frac{T_j}{\sum_{h=1}^{N} T_h} \hat{\mathbf{S}}_j,$$

where

$$\hat{\mathbf{S}}_j = \hat{\boldsymbol{\Omega}}_{j,0} + \sum_{\ell=1}^{m_j} w\left(\ell, m_j\right) \left(\hat{\boldsymbol{\Omega}}_{j,\ell} + \hat{\boldsymbol{\Omega}}_{j,\ell}'\right), \tag{S.5.1}$$

and

$$\hat{\boldsymbol{\Omega}}_{j,\ell} = \frac{1}{T_j} \sum_{t=\ell+1}^{T_j} \boldsymbol{\zeta}_{jt} \boldsymbol{\zeta}_{j,t-\ell}' \hat{u}_{jt} \hat{u}_{j,t-\ell},$$

in which $\hat{u}_{jt} = y_{jt} - \hat{\boldsymbol{\theta}}' \boldsymbol{\zeta}_{jt}$. We set

$$w\left(\ell, m_j\right) = 1 - \frac{\ell}{m_j + 1},$$

and $m_j = m_{j,nT}$ is chosen to be a suitable increasing function of the sample size. We set $m_{j,nT}$ to be the integer part of $(T_j)^{1/3}$.

S.5.2 Inference robust to serial and cross-sectional correlation of errors

Allowing for correlation of errors over time, as well as across units (countries) requires a different estimator of $\mathbf{S}_{nT}$. It is useful to re-write $\mathbf{S}_{nT}$ as

$$\mathbf{S}_{nT} = \sum_{t=1}^{T} \sum_{t'=1}^{T} \mathbf{h}_{nt} \mathbf{h}'_{nt'},$$

where

$$\mathbf{h}_{nt} = \sum_{j \in \mathcal{N}_t} \boldsymbol{\zeta}_{jt} u_{jt},$$

in which we use $\mathcal{N}_t$ as the index set of cross-section units with available observations for a period t. $\mathbf{S}_{nT}$ is estimated as

$$\hat{\mathbf{S}}_{nT} = \hat{\boldsymbol{\Omega}}_{\cdot,0} + \sum_{\ell=1}^{m} w(\ell, m) \left(\hat{\boldsymbol{\Omega}}_{\cdot,\ell} + \hat{\boldsymbol{\Omega}}'_{\cdot,\ell} \right), \tag{S.5.2}$$

where

$$\hat{\boldsymbol{\Omega}}_{\cdot,\ell} = \frac{1}{T} \sum_{t=\ell+1}^{T} \hat{\mathbf{h}}_{nt} \hat{\mathbf{h}}_{nt-\ell}, \text{ for } \ell = 0, 1, ..., m,$$

and

$$\mathbf{h}_{nt} = \sum_{j \in \mathcal{N}_t} \boldsymbol{\zeta}_{jt} \hat{u}_{jt}.$$

$m = m_{nT}$ is chosen to be a suitable increasing function of the sample size. We set m_{nT} to be the integer part of $T^{1/3}$.

S.6 Pooled and fixed effects estimates of panel regressions of effective transmission rates using alternative choices of lag orders

Table S2: Panel regressions of effective transmission rates for pre-vaccination sample ending January 31, 2021

(all covariates are lagged $p = 7$ or 14 days)

	Estimates with Common $\mathcal{R}_0$ Numbers				Estimates with Country-Specific $\mathcal{R}_0$ Numbers			
Lag Order:	$p = 7$		$p = 14$		$p = 7$		$p = 14$	
Multiplication Factor:	$MF = 2$	$MF = 3$	$MF = 2$	$MF = 3$	$MF = 2$	$MF = 3$	$MF = 2$	$MF = 3$
Stringency Index	**-1.88**	**-1.87**	**-2.25**	**-2.24**	**-1.95**	**-1.93**	**-2.43**	**-2.42**
standard s.e. (t-ratio)	0.11 (-17.0)	0.11 (-16.8)	0.10 (-21.9)	0.10 (-21.7)	0.11 (-17.1)	0.11 (-16.9)	0.11 (-22.8)	0.11 (-22.5)
robust1 s.e. (t-ratio)	0.20 (-9.2)	0.20 (-9.2)	0.15 (-14.5)	0.16 (-14.4)	0.20 (-9.6)	0.20 (-9.6)	0.16 (-15.4)	0.16 (-15.3)
robust2 s.e. (t-ratio)	0.29 (-6.4)	0.29 (-6.4)	0.23 (-9.9)	0.23 (-9.8)	0.38 (-5.2)	0.38 (-5.1)	0.30 (-8.1)	0.30 (-8.0)
Economic Support	**-0.62**	**-0.61**	**-0.25**	**-0.23**	**-1.34**	**-1.33**	**-0.65**	**-0.64**
standard s.e. (t-ratio)	0.07 (-8.9)	0.07 (-8.6)	0.07 (-3.8)	0.07 (-3.5)	0.09 (-14.5)	0.09 (-14.4)	0.08 (-7.7)	0.08 (-7.6)
robust1 s.e. (t-ratio)	0.16 (-3.8)	0.16 (-3.8)	0.11 (-2.2)	0.11 (-2.1)	0.20 (-6.6)	0.20 (-6.6)	0.14 (-4.6)	0.14 (-4.6)
robust2 s.e. (t-ratio)	0.27 (-2.3)	0.27 (-2.3)	0.16 (-1.6)	0.16 (-1.5)	0.28 (-4.8)	0.28 (-4.9)	0.17 (-3.9)	0.17 (-3.9)
Threshold Variable	**-2.75**	**-2.75**	**-2.29**	**-2.29**	**-2.24**	**-2.23**	**-1.96**	**-1.95**
standard s.e. (t-ratio)	0.10 (-27.3)	0.10 (-27.1)	0.09 (-26.6)	0.09 (-26.4)	0.11 (-21.3)	0.11 (-21.1)	0.09 (-21.3)	0.09 (-21.1)
robust1 s.e. (t-ratio)	0.28 (-9.8)	0.28 (-9.8)	0.23 (-10.1)	0.23 (-10.0)	0.27 (-8.3)	0.27 (-8.2)	0.23 (-8.6)	0.23 (-8.6)
robust2 s.e. (t-ratio)	0.40 (-6.8)	0.40 (-6.8)	0.40 (-5.7)	0.40 (-5.7)	0.63 (-3.6)	0.63 (-3.6)	0.61 (-3.2)	0.61 (-3.2)
threshold value	0.18	0.18	0.12	0.12	0.18	0.18	0.09	0.09
$\mathcal{R}_0$ **Numbers (Constant Terms)**								
common [robust2 s.e.]	5.66 [0.18]	5.66 [0.18]	5.12 [0.26]	5.12 [0.26]				
specific [robust2 s.e.]:								
Belgium					5.73 [0.67]	5.73 [0.67]	5.13 [0.63]	5.14 [0.63]
France					5.69 [0.67]	5.69 [0.68]	5.20 [0.65]	5.20 [0.65]
Germany					5.27 [0.69]	5.26 [0.69]	4.96 [0.67]	4.95 [0.67]
Italy					5.78 [0.71]	5.77 [0.71]	5.37 [0.66]	5.36 [0.66]
Netherlands					5.71 [0.67]	5.71 [0.67]	5.12 [0.64]	5.12 [0.64]
Poland					5.29 [0.69]	5.29 [0.69]	4.89 [0.66]	4.89 [0.66]
Portugal					5.82 [0.69]	5.83 [0.69]	5.28 [0.65]	5.28 [0.65]
Spain					5.97 [0.68]	5.97 [0.69]	5.39 [0.65]	5.39 [0.66]
United Kingdom					6.10 [0.66]	6.10 [0.66]	5.44 [0.63]	5.44 [0.63]
R-squared	0.49	0.48	0.55	0.55	0.52	0.51	0.57	0.57

Notes: See notes to Table 1.

Table S3: Panel regressions of effective transmission rates for full sample ending November 30, 2021

(all covariates are lagged $p = 7$ or 14 days)

	Estimates with Common $\mathcal{R}_0$ Numbers				Estimates with Country-Specific $\mathcal{R}_0$ Numbers			
Lag Order:	$p = 7$		$p = 14$		$p = 7$		$p = 14$	
Multiplication Factor:	$MF = 2$	$MF = 3$	$MF = 2$	$MF = 3$	$MF = 2$	$MF = 3$	$MF = 2$	$MF = 3$
Stringency Index	**-1.86**	**-1.89**	**-2.04**	**-2.07**	**-1.99**	**-1.98**	**-2.21**	**-2.20**
standard s.e. (t-ratio)	0.08 (-22.0)	0.09 (-21.1)	0.08 (-25.1)	0.09 (-23.9)	0.09 (-22.1)	0.10 (-20.7)	0.09 (-25.5)	0.09 (-23.8)
robust1 s.e. (t-ratio)	0.19 (-9.8)	0.20 (-9.6)	0.15 (-13.4)	0.16 (-13.1)	0.19 (-10.5)	0.20 (-10.0)	0.15 (-14.3)	0.16 (-13.7)
robust2 s.e. (t-ratio)	0.30 (-6.3)	0.30 (-6.3)	0.24 (-8.5)	0.24 (-8.6)	0.33 (-6.0)	0.34 (-5.9)	0.26 (-8.4)	0.27 (-8.2)
Economic Support	**-0.39**	**-0.37**	**-0.20**	**-0.18**	**-0.71**	**-0.75**	**-0.36**	**-0.39**
standard s.e. (t-ratio)	0.05 (-7.9)	0.05 (-7.2)	0.05 (-4.3)	0.05 (-3.7)	0.06 (-11.2)	0.07 (-11.1)	0.06 (-5.7)	0.07 (-5.9)
robust1 s.e. (t-ratio)	0.12 (-3.2)	0.12 (-3.1)	0.09 (-2.3)	0.09 (-2.0)	0.18 (-4.0)	0.18 (-4.2)	0.12 (-2.9)	0.13 (-3.1)
robust2 s.e. (t-ratio)	0.20 (-1.9)	0.20 (-1.9)	0.12 (-1.6)	0.13 (-1.5)	0.27 (-2.7)	0.25 (-3.0)	0.15 (-2.4)	0.14 (-2.8)
Vaccinated Share	**-1.96**	**-1.69**	**-1.32**	**-0.99**	**-2.20**	**-1.94**	**-1.48**	**-1.15**
standard s.e. (t-ratio)	0.16 (-12.6)	0.16 (-10.3)	0.15 (-8.7)	0.16 (-6.1)	0.16 (-14.0)	0.17 (-11.6)	0.15 (-9.6)	0.16 (-7.0)
robust1 s.e. (t-ratio)	0.32 (-6.1)	0.35 (-4.8)	0.37 (-3.6)	0.41 (-2.4)	0.31 (-7.0)	0.34 (-5.6)	0.37 (-4.0)	0.41 (-2.8)
robust2 s.e. (t-ratio)	0.52 (-3.8)	0.56 (-3.0)	0.55 (-2.4)	0.60 (-1.6)	0.49 (-4.4)	0.54 (-3.6)	0.54 (-2.7)	0.59 (-1.9)
Delta Variant Share	**1.26**	**1.26**	**0.86**	**0.82**	**1.37**	**1.38**	**0.91**	**0.88**
standard s.e. (t-ratio)	0.10 (12.4)	0.11 (11.7)	0.10 (8.6)	0.11 (7.7)	0.10 (13.3)	0.11 (12.7)	0.10 (9.1)	0.11 (8.2)
robust1 s.e. (t-ratio)	0.20 (6.2)	0.22 (5.6)	0.24 (3.6)	0.27 (3.1)	0.20 (6.9)	0.22 (6.3)	0.24 (3.8)	0.27 (3.3)
robust2 s.e. (t-ratio)	0.29 (4.3)	0.32 (3.9)	0.33 (2.6)	0.36 (2.3)	0.27 (5.0)	0.30 (4.6)	0.33 (2.8)	0.36 (2.4)
Threshold Variable	**-2.93**	**-2.90**	**-2.42**	**-2.39**	**-2.65**	**-2.61**	**-2.23**	**-2.18**
standard s.e. (t-ratio)	0.09 (-32.2)	0.10 (-30.0)	0.08 (-30.5)	0.08 (-28.2)	0.09 (-28.0)	0.10 (-26.0)	0.08 (-26.5)	0.09 (-24.3)
robust1 s.e. (t-ratio)	0.28 (-10.6)	0.28 (-10.4)	0.24 (-10.2)	0.24 (-10.0)	0.28 (-9.4)	0.28 (-9.2)	0.23 (-9.5)	0.24 (-9.3)
robust2 s.e. (t-ratio)	0.39 (-7.5)	0.39 (-7.5)	0.42 (-5.7)	0.42 (-5.7)	0.69 (-3.8)	0.69 (-3.8)	0.71 (-3.2)	0.71 (-3.1)
threshold value	0.18	0.18	0.12	0.12	0.18	0.18	0.12	0.12
$\mathcal{R}_0$ **Numbers (Constant Terms)**								
common [robust2 s.e.]	5.65 [0.17]	5.66 [0.17]	5.08 [0.27]	5.09 [0.28]				
FE [robust2 s.e.]:								
Belgium					5.75 [0.67]	5.78 [0.68]	5.11 [0.67]	5.14 [0.68]
France					5.62 [0.68]	5.62 [0.69]	5.08 [0.70]	5.08 [0.70]
Germany					5.47 [0.69]	5.42 [0.70]	5.01 [0.71]	4.96 [0.72]
Italy					5.84 [0.69]	5.82 [0.70]	5.27 [0.69]	5.25 [0.70]
Netherlands					5.68 [0.67]	5.72 [0.69]	5.09 [0.69]	5.12 [0.70]
Poland					5.51 [0.67]	5.50 [0.68]	4.94 [0.69]	4.93 [0.70]
Portugal					5.77 [0.68]	5.78 [0.69]	5.17 [0.69]	5.17 [0.70]
Spain					5.84 [0.68]	5.84 [0.69]	5.18 [0.68]	5.18 [0.69]
United Kingdom					5.84 [0.67]	5.85 [0.68]	5.20 [0.68]	5.21 [0.69]
R-squared	0.38	0.36	0.43	0.40	0.40	0.38	0.43	0.41

Notes: See notes to Table 2.

S.7 Summary statistics

Table S4: Sample minimum and maximum

	Social Distancing index	Economic Support index	Vaccination share	Delta Variant share	Threshold indicator
	Pre-vaccination sample ending 31 January 2021				
Belgium	[0.11,0.81]	[0.00,1.00]	[0.00,0.00]	[0.00,0.00]	[0.00,1.00]
France	[0.06,0.88]	[0.00,1.00]	[0.00,0.00]	[0.00,0.00]	[0.00,1.00]
Germany	[0.11,0.85]	[0.00,0.63]	[0.00,0.00]	[0.00,0.00]	[0.00,1.00]
Italy	[0.19,0.94]	[0.00,0.75]	[0.00,0.00]	[0.00,0.00]	[0.00,1.00]
Netherlands	[0.00,0.79]	[0.00,0.88]	[0.00,0.00]	[0.00,0.01]	[0.00,1.00]
Poland	[0.11,0.87]	[0.00,0.75]	[0.00,0.00]	[0.00,0.00]	[0.00,1.00]
Portugal	[0.06,0.88]	[0.00,0.75]	[0.00,0.00]	[0.00,0.00]	[0.00,1.00]
Spain	[0.11,0.85]	[0.00,0.88]	[0.00,0.00]	[0.00,0.00]	[0.00,1.00]
United Kingdom	[0.11,0.88]	[0.00,1.00]	[0.00,0.01]	[0.00,0.00]	[0.00,1.00]
all 9 countries	[0.00,0.94]	[0.00,1.00]	[0.00,0.01]	[0.00,0.01]	[0.00,1.00]
	1 February 2021 - 30 November 2021 sample				
Belgium	[0.40,0.76]	[0.75,0.75]	[0.00,0.75]	[0.00,1.00]	[1.00,1.00]
France	[0.44,0.75]	[0.25,0.50]	[0.00,0.69]	[0.00,1.00]	[1.00,1.00]
Germany	[0.37,0.83]	[0.38,0.38]	[0.00,0.67]	[0.00,1.00]	[1.00,1.00]
Italy	[0.47,0.83]	[0.75,0.75]	[0.00,0.73]	[0.00,1.00]	[1.00,1.00]
Netherlands	[0.32,0.82]	[0.50,0.75]	[0.00,0.66]	[0.00,1.00]	[1.00,1.00]
Poland	[0.39,0.76]	[0.75,1.00]	[0.00,0.54]	[0.00,1.00]	[1.00,1.00]
Portugal	[0.41,0.88]	[0.75,0.75]	[0.00,0.88]	[0.00,1.00]	[1.00,1.00]
Spain	[0.41,0.71]	[0.88,0.88]	[0.00,0.80]	[0.00,1.00]	[1.00,1.00]
United Kingdom	[0.41,0.88]	[0.25,1.00]	[0.01,0.68]	[0.00,1.00]	[1.00,1.00]
all 9 countries	[0.32,0.88]	[0.25,1.00]	[0.00,0.88]	[0.00,1.00]	[1.00,1.00]

Notes: This table report sample maximums and minimums of the individual regressors in pooled regressions presented in Table 2. The top panel reports summary statistics for the prevaccination sample (ending January 31 2021), and the bottom panel reports summary statistics for the remainder of the full sample – 1 February 2021 to 30 November 2021.