The Causal Effects of Lockdown Policies on Health and Macroeconomic Outcomes

Jonas E. Arias* Jesús Fernández-Villaverde†

Juan F. Rubio-Ramírez‡ Minchul Shin§

March 7, 2022

Abstract

We assess the causal impact of epidemic-induced lockdowns on health and macroeconomic outcomes and measure the trade-off between containing the spread of an epidemic and economic activity. To do so, we estimate an epidemiological model with time-varying parameters and use its output as information for estimating SVARs and LPs that quantify the causal effects of nonpharmaceutical policy interventions. We apply our approach to Belgian data for the COVID-19 epidemic during 2020. We find that additional government-mandated mobility curtailments would have reduced deaths at a very small cost in terms of GDP.

JEL classification: C1; C5; I1.

Keywords: Causality; Policy interventions; Epidemiological models; Bayesian estimation.

We thank the editor, three referees, Kris Boudt, Frank Diebold, Lukasz Drozd, Jan Eeckhout, Lutz Kilian, Mikkel Plagborg-Møller, and Christian Wolf for useful comments. Some of the ideas in this paper came out of previous work with Chad Jones. John Cochrane was the first to point out to us the potential key importance of time-varying parameters in epidemiological models. An earlier version of this paper circulated as "Bayesian Estimation of Epidemiological Models: Methods, Causality, and Policy Trade-Offs." Disclaimer: The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Atlanta, the Federal Reserve Bank of Philadelphia, or the Federal Reserve System. Any errors or omissions are the responsibility of the authors. No statements here should be treated as legal advice.

*Federal Reserve Bank of Philadelphia. Email: jonas.arias@phil.frb.org.

†University of Pennsylvania. Email: jesusfv@econ.upenn.edu.

‡Emory University and Federal Reserve Bank of Atlanta. Email: juan.rubio-ramirez@emory.edu.

§Federal Reserve Bank of Philadelphia. Email: minchul.shin@phil.frb.org.

1 Introduction

In this paper, we assess the causal impact of epidemic-induced lockdowns, such as "shelter-in-place" orders and other nonpharmaceutical policy interventions (NPIs), on health and macroeconomic outcomes. In particular, we look at the effects of a government stringency shock (where mobility curtailments are stricter than those determined by the systematic component of the public health policy) and we measure how a reproduction shock (such as the spread of a new, more contagious variant of the SARS-CoV-2 virus) affects the rest of the variables depending on the level of government limitations on mobility.

Assessing causality is challenging for two reasons. First, because the levels of activity and mobility, as measured by indexes such as the Google COVID-19 Community Mobility Reports, change due to both voluntary precautionary behavior and government orders. Thus, we need some identification assumptions. Second, because we lack some of the required data. Some of the inputs one would need for a causality exercise, such as the effective reproduction number of a disease, are time-varying and not directly observed. Other data, such as new cases, are observed subject to large, biased, and time-varying measurement errors due to problems like testing constraints, unwillingness to test, etc.[1]

To tackle these challenges, we proceed in two steps. In the first step, we estimate, using the Bayesian approach, an epidemiological model with time variation in the parameters controlling an infectious disease's dynamics. Time variation in the parameters of the model allows us i) to capture changes in the behavior of individuals as they respond to public health conditions (either voluntarily or forced by government mandates); and ii) to include shifts in the transmission and clinical outcomes of the epidemic, such as virus variants, new medical treatments, or better organization of social distancing.[2]

We estimate the model with a sequential Monte Carlo to account for the nonlinearities present in the model equations while relying on a Markov chain Monte Carlo to sample from its posterior. The former is necessary to capture the fast variations in the spread of infectious diseases and the turning points of the different waves of an epidemic. The latter allows us to incorporate information in the prior of the parameters from laboratory results, clinical studies, seroprevalence surveys, and the experience of other regions or countries.

In the second step, we show how to use outputs from the epidemiological model to conduct a causal assessment of NPIs. We illustrate how the model takes observations of recorded new

[1]This point has been recognized in the epidemiological literature for a long time; for example, see O'Neill and Roberts (1999). This is why some researchers, such as Manski and Molinari (2021) and Toulis (2021), try to estimate COVID-19 prevalence based on partial identification.

[2]We do not claim that these time-varying parameters are structural in the sense of being invariant to policy interventions à la Hurwicz (1966). We consider them only as behavioral parameters, which might be complex functions of preferences, technology, and information processes. See Fernández-Villaverde and Rubio-Ramírez (2007) for details.

cases, recorded hospitalizations, and recorded new deaths in hospitals and at home, and results from infrequent seroprevalence studies and, using the cross-equation restrictions determined by the model's dynamics (e.g., the relation between the infectious population and the new cases), produces useful outputs. In particular, we can estimate the effective reproduction number and the true new cases. We employ these outputs to ascertain the causal effects of government mandates using standard causal analysis techniques such as structural vector autoregressions (SVARs), to assess the effects of a government stringency shock, and local projections (LPs), to measure the consequences of a reproduction shock.

In comparison, a reduced-form model –such as those in Gostic et al. (2020)– that does not use the structure of an epidemiological model cannot exploit the cross-equation restrictions and the biological and clinical information they bring to the estimation.[3] Interestingly, our methodology goes well beyond epidemiological models: it can be applied to any causal inference problem that relies on obtaining good estimates of unobserved variables.[4]

We illustrate these two steps using data from the COVID-19 epidemic in Belgium during 2020. However, our framework is much more general than the details of the model we specify or the data we use. It is is applicable to a wide range of epidemiological models, other diseases, and alternative techniques to ascertain causality in time series beyond SVARs and LPs.

We pick Belgium for five reasons. First, it is one of the countries that suffered the most from the COVID-19 epidemic during 2020. Second, Belgium is a small country, 30,689 km^2 (roughly 20% larger than Maryland), and geographically and climatically quite homogeneous in the area where most of the population is concentrated (that is, in the coastal plain and central plateau outside the Ardennes). That allows us to consider Belgium as a single unit for our analysis.[5] Third, Belgium has high-quality national data, including reported new cases, hospitalizations, deaths in hospitals and at home, and several national seroprevalence studies that allow us to efficiently implement the cross-equation restrictions implied by our model. Fourth, Belgium experienced, within our sample, three waves of the epidemic, which would demonstrate how our approach can handle intricate data patterns. Fifth, during 2020, the Belgian government passed mandates curtailing mobility on several occasions, which will give us the identification of the effects of these mandates.

We estimate the model using data from Sciensano, a public institution in Belgium, on COVID-

[3]The cross-equations discipline could come at a cost if the model is misspecified: the resulting filtered and smoothed estimates might be biased. We could detect misspecification by comparing our model against flexible reduced-form time-series models such as the one proposed by Ho et al. (2021).

[4]In a different environment, Gilchrist and Zakrajšek (2012) follow a similar approach by using a pricing model to recover a measure of sentiment in the corporate credit market. They show, using reduced-form regressions and SVARs, how such a measure can affect the economy and asset prices.

[5]The incidence of COVID-19 in 2020 was around 10% higher in Wallonia than in Flanders in per capita terms (with Flanders having a worst first wave and Wallonia a more damaging third wave) and about average in Brussels. To keep our analysis transparent, we ignore these small regional differences.

19 per capita deaths in hospitals and at home, per capita hospitalizations due to COVID-19, the observed per capita new cases, and the seroprevalence rates. The sample starts on March 15, 2020, the first day with available data on COVID-19 hospitalizations, and ends on November 30, 2020, right before the arrival of the first vaccines in Belgium.[6] We did not consider later data, as they would incorporate a growing share of the vaccinated population (furthermore, with a strong age component as older people were vaccinated first). This would require complicating the epidemiological model and adding cases by age, for which data are very limited. Finally, it would change the causality effect of lockdown policies. Hence, while it would be possible to handle the 2021 observations using our approach, it is beyond the scope of our paper.

We show how the data are informative about the parameters of the model, how the point estimates are in line with other evidence, and how the model fits the data well (including tracking the three waves of the epidemic in the data). Also, we recover the smoothed states such as time-varying reproduction numbers, time-varying death rates, and new cases.

Using the SVAR, we estimate that a positive stringency shock — such that, upon impact, the posterior median increase is equivalent to a one-unit increase in one of the ordinal components of the Oxford Stringency Index of mobility curtailments— leads to roughly $1,000$ fewer deaths in Belgium after 2 months, or around 6% fewer deaths in the sample at a negligible cost to output. In fact, our point estimate implies a small economic cost of about €1 per capita (although it is hard to distinguish it from zero cost). The intuition is that, by controlling the epidemic, a positive stringency shock brings higher economic activity after a few weeks that more than compensates for the very short-run losses. Using the LPs, we find that high government stringency —a level of stringency above the median in our sample— could save up to about 250 deaths in the first two weeks after the reproduction shock, or around 1.5% of the deaths in our sample, compared with low government stringency, at the small cost of between €2 and €4 per capita.

The rest of the paper is structured as follows. Section 2 presents a time-varying model of the COVID-19 epidemic. Section 3 describes the data and briefly summarizes how we estimate the model. Section 4 reports our results and Section 5 shows how we exploit these results for causality and policy trade-off analysis with SVARs and LPs. Section 6 concludes.

2 The Epidemiological Model

We specify a SIRD compartmental model allowing for time variation in i) the effective contact rate among individuals; ii) the hospitalization rates; and iii) death probabilities. Figure 1 outlines the compartments' structure.

All individuals start as susceptible (S) except for a small fraction of infectious individuals

[6]The Belgian vaccination campaign was announced on December 3, 2020, and started on December 28, 2020. Thus, it is likely that behavior began to change around early December 2020.

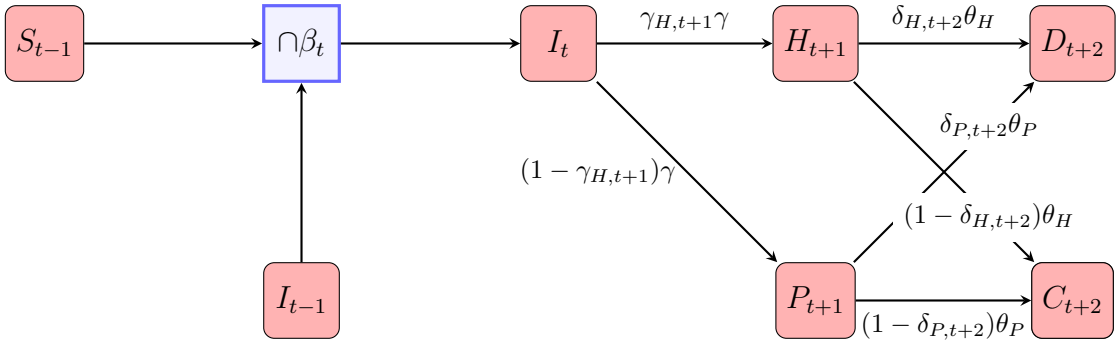


Figure 1: Outline of the model

(I). Susceptible individuals meet with infectious individuals and become, with some probability, infectious themselves. This probability is controlled by β_t, the time-varying effective contact rate. Every day, a share γ of infectious individuals can recover in a hospital (H) or at home (P). Here, "home" means all dwellings, including private residences and retirement communities, outside of hospitals (unfortunately, our data set does not distinguish between private residences and retirement communities, a potentially important distinction). Of those, the share of individuals recovering in hospitals is time varying and equal to $\gamma_{H,t+1}$. Hence, the share $1-\gamma_{H,t+1}$ recovers at home. Recovery ends with death (D) or a return to a healthy status (C). Every day a share θ_H (θ_P) of individuals in a hospital (at home) recover. Of those, $\delta_{H,t+2}$ ($\delta_{P,t+2}$) die, while $1-\delta_{H,t+2}$ ($1-\delta_{P,t+2}$) return to healthy status. These death rates are also time varying.

We assume that, once cured, an individual cannot become infectious again. We pick this specification because the evidence is that re-infections with SARS-CoV-2 during 2020 (our data sample) were possible but unlikely. However, it would be easy to extend the model to allow for re-infections, perhaps after the (stochastic) waning of immunity or the arrival of new varieties of the virus, a common concern later in 2021 and 2022. Indeed, our causal analysis in Section 5 treats the emergence of new variants as a possible interpretation of a reproduction shock.

2.1 The Transition and Measurement Equations

The transition equation of the model we will estimate is:

$$\begin{pmatrix} S_t \\ I_t \\ H_t \\ P_t \\ b_t \\ g_{H,t} \\ d_{H,t} \\ d_{P,t} \\ n_t \end{pmatrix} = \begin{pmatrix} S_{t-1} \\ I_{t-1} \\ H_{t-1} \\ P_{t-1} \\ b_{t-1} \\ g_{H,t-1} \\ d_{H,t-1} \\ d_{P,t-1} \\ n_{t-1} \end{pmatrix} + \begin{pmatrix} -\beta_t S_{t-1} I_{t-1} \\ \beta_t S_{t-1} I_{t-1} - \gamma I_{t-1} \\ \gamma_{H,t} \gamma I_{t-1} - \theta_H H_{t-1} \\ (1-\gamma_{H,t})\gamma I_{t-1} - \theta_P P_{t-1} \\ \sigma_b \varepsilon_{b,t} \\ \sigma_g \varepsilon_{g,t} \\ \sigma_h \varepsilon_{h,t} \\ \sigma_p \varepsilon_{p,t} \\ \sigma_n \varepsilon_{n,t} \end{pmatrix}, \tag{1}$$

where we have a 9×1 vector of states $\boldsymbol{X}_t = \{S_t, I_t, H_t, P_t, b_t, g_{H,t}, d_{H,t}, d_{P,t}, n_t\}$ and a 5×1 vector of innovations $\boldsymbol{\mathcal{E}}_t = \{\varepsilon_{b,t}, \varepsilon_{g,t}, \varepsilon_{h,t}, \varepsilon_{p,t}, \varepsilon_{n,t}\}$. Let us describe each of these variables.

Row one of Equation (1) tells us that the share of the population that is susceptible on day t, S_t, is equal to the share of the susceptible population the day before, S_{t-1}, minus the new infections as a share of the population, given by a matching function of the share of susceptibles and infectious, I_{t-1}, yesterday: $\beta_t S_{t-1} I_{t-1}$.

The most relevant feature of row one, and a central aspect of our model, is that the effective contact rate, β_t, in the matching function is time varying. In many epidemiological models, the effective contact rate is a constant parameter. However, individuals respond to the epidemiological situation for two reasons. First, individuals take voluntary precautionary measures (lower mobility, wearing personal protection equipment, changed business protocols). For this point to hold, we do not need to assume full agent rationality; we only require some degree of endogenous reaction. Second, individuals change how often they effectively interact with each other in any given period because governments impose NPIs in response to the health situation, such as curtailments on businesses and travel or mandatory mask-wearing. By letting the effective contact rate be time varying, the model can capture these two mechanisms.

We assume that $b_t = \log(\beta_t)$ follows a random walk, where the innovation is a truncated standard normal distribution to avoid having fewer new cases in the model than in the data and preventing S_t from becoming negative. Later, we will argue that there will be a percentage of true cases of COVID-19 that are not reported (for instance, because the cases are asymptomatic or due to insufficient testing), but that the situation where there are *more* reported cases than true cases is not relevant empirically. The number of "false positives" is most likely trivially small and swamped, by at least an order of magnitude, by the under-reporting of "false negatives."

Row two of Equation (1) describes how the share of infectious, I_t, evolves. The new share

is equal to the share yesterday, I_{t-1}, plus the new infections, $\beta_t S_{t-1} I_{t-1}$, minus the share of infectious that move to the next compartment, γI_{t-1}.

Row three of Equation (1) determines the evolution of H_t, the share of the population hospitalized on day t. The share evolves through inflows from the compartment of infectious at a rate $\gamma_{H,t}\gamma$ and outflows at a rate θ_H. We allow $\gamma_{H,t}$ to move to reflect the changing availability of hospital beds and shifting decisions by patients (e.g., should I go to a hospital?) and clinical protocols (e.g., should this patient be hospitalized or sent back home?).

As we did with the effective contact rate, we specify a flexible law of motion for $\gamma_{H,t}$. In row six of Equation (1), we define a random walk for $g_{H,t}$, $g_{H,t} = g_{H,t-1} + \sigma_g \varepsilon_{g,t}$, where $\varepsilon_{g,t} \sim \mathcal{N}(0,1)$ and $\gamma_{H,t} = \frac{e^{g_{H,t}}}{e^{g_{H,t}}+1}$. Hence, the mapping from $\gamma_{H,t}$ to $g_{H,t}$ is $g_{H,t} = \log\left(\frac{\gamma_{H,t}}{1-\gamma_{H,t}}\right)$. This mapping will be useful for interpreting $g_{H,0}$ later on.

Row four of Equation (1) governs the evolution of the share of individuals recovering at home, P_t. In parallel to H_t, the share evolves through inflows from the compartment of infectious at a rate $(1-\gamma_{H,t})\gamma$ and outflows at a rate θ_P. Rows three and four implicitly assume that individuals recovering in a hospital or at home do not switch between these two compartments. This assumption is due to data limitations (as we do not observe how many patients previously in a hospital die while recovering at home). Fortunately, this assumption does not cause too many problems. If a patient is at home but later hospitalized, we can consider her as still being in the infectious compartment. Conversely, individuals formerly hospitalized and currently recovering at home are unlikely to be infectious and, hence, they could be considered as being in the recovered compartment. Also, the measurement error will help us tackle, empirically, deviations from this assumption in the data. Rows seven to nine of Equation (1) describe the evolution of three random walks that will be used in the measurement equation below, with innovations $\varepsilon_{i,t} \sim \mathcal{N}(0,1)$ for $i \in \{h, p, n\}$.

We observe deaths in a hospital ($D^{obs}_{H,t}$), deaths at home ($D^{obs}_{P,t}$), and hospitalized patients (H^{obs}_t) at a daily frequency; new cases (G^{obs}_t) at a weekly frequency; and the point estimates of S_t from the seroprevalence surveys for some periods. We assume that the log of each of these five variables is measured with error. The measurement error arises for many reasons, from administrative mistakes and delays in file keeping to the under-reporting of cases due to testing bottlenecks. Thus, the measurement equation is:

$$\begin{pmatrix} \log\left(\frac{1}{\text{Population}} + \Delta D^{obs}_{H,t}\right) \\ \log\left(\frac{1}{\text{Population}} + \Delta D^{obs}_{P,t}\right) \\ \log\left(H^{obs}_t\right) \\ \log\left(G^{obs}_{t+5}\right) \\ \log\left(S^{obs}_t\right) \end{pmatrix} = \begin{pmatrix} \log\left(\frac{1}{\text{Population}} + \delta_{H,t}\theta_H H_{t-1}\right) \\ \log\left(\frac{1}{\text{Population}} + \delta_{P,t}\theta_P P_{t-1}\right) \\ \log\left(H_t\right) \\ \log\left((S_{t-1} - S_t)\gamma_{n,t}\right) \\ \log(S_t) \end{pmatrix} + \begin{pmatrix} \sigma_{D_H} u_{D_H,t} \\ \sigma_{D_P} u_{D_P,t} \\ \sigma_H u_{H,t} \\ \sigma_G u_{G,t} \\ \sigma_S u_{S,t} \end{pmatrix}, \quad (2)$$

where $\boldsymbol{y}_t$ denotes the time t observations and $\boldsymbol{y}^T = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_T\}$ denotes the sample of observables available over T periods. Row one of Equation (2) links the (log) first difference of the observed share of deaths in hospitals, $\Delta D^{obs}_{H,t}$, with the share of individuals in hospitals H_{t-1} and the time-varying death rate $\delta_{H,t} = \frac{e^{d_{H,t}}}{1+e^{d_{H,t}}}$, where $d_{H,t}$ comes from row seven of Equation (1). For the first two observed series in Equation (2) we add the term 1/Population because there are a few days with zero deaths.

Row two of Equation (2) links the (log) first difference of the observed share of deaths at home, $\Delta D^{obs}_{P,t}$, with the share of individuals recovering at home P_{t-1} and the time-varying death rate $\delta_{P,t} = \frac{e^{d_{P,t}}}{1+e^{d_{P,t}}}$, where $d_{P,t}$ is a random walk that comes from row eight of Equation (1). Time variation of the death rates in a hospital and at home captures improvements in clinical protocols that increase the survival of patients, variation in hospital congestion (which may facilitate or complicate the treatment of patients), and changes in the mix of recovering individuals across different demographic groups (in a model with several compartments for individuals indexed by age, this last effect would disappear; unforunately we lack data by age).

Row three of Equation (2) tells us that the observed share of the population in hospitals is measured with error. An interpretation of this measurement error (beyond administrative mistakes in recording) is that some patients in hospitals are still infectious because isolation measures have not been fully implemented. Thus, while measured as being in a hospital, the individuals are still in the infectious compartment for the purposes of the model.

Row four of Equation (2) gives us the observed share of cases, equal to the difference of the share of susceptible individuals times a factor $\gamma_{n,t} = \frac{e^{n_t}}{e^{n_t}+1} e^{-\mu} < 1$ that determines the percentage of cases that are reported. In the case of deaths, we do not introduce systematic under-reporting. As shown by Molenberghs et al. (2020) and Sierra et al. (2020), and in contrast with other countries, Belgian data show virtually no discrepancy between COVID-19 reported mortality (confirmed and possible cases) and excess mortality. Analysis by Sciensano https://epistat.wiv-isp.be/momo/ confirms this point. The rate of under-reporting depends on n_t, which follows the random walk defined in row nine of Equation (1). The parameter μ ensures that under-reporting never goes to zero, for example, due to false negatives. We let cases be under-reported by the $\gamma_{n,t}$ factor because this was a key factor at the start of the COVID-19 epidemic and we want to recover an estimate of the rate of reporting. At the same time, we assume that cases cannot be over-reported. As we will explain in Section 3, we only use row four every Friday by matching it to the new cases reported the following Wednesday.

Row five of Equation (2) links the point estimates of the susceptible share of the population from Herzog et al. (2020) to the corresponding share from our model. As we will describe below in more detail, this study is a prospective serial cross-sectional nationwide seroprevalence evaluation conducted in Belgium using blood samples collected during five different periods. We only use this equation for the dates for which the study is available. This row forces our

smoothed states to incorporate the high-quality information from seroprevalence surveys, but allowing for differences due to sampling uncertainty and possible measurement errors.

Rows four and five of Equation (2) illustrate how state-space representations parsimoniously incorporate observables at different frequencies.

2.2 The Time-Varying Reproduction Numbers

A key state in our model is the effective contact rate β_t, the time-varying parameter that determines the speed of contagion in the matching function of susceptible and infectious individuals. When β_t is high, the infection spreads quickly. When β_t is low, the infection spreads slowly and might eventually abate. The basic reproduction number, $\mathcal{R}_{0,t}$, is a closely related concept: the expected number of cases generated by one case in a population where all individuals are susceptible (the expectation is computed assuming that the transmission rate remains constant). In our model, $\mathcal{R}_{0,t} = \frac{\beta_t}{\gamma}$. Thus, the basic reproduction number inherits the time variability encoded in β_t. We also have the effective reproduction number, $\mathcal{R}_{e,t}$, which considers the changes in the share of the susceptible population over time $\mathcal{R}_{e,t} = \mathcal{R}_{0,t} S_t = \frac{\beta_t S_t}{\gamma}$. By doing so, $\mathcal{R}_{e,t}$ offers a better measure of the instantaneous speed of the spreading of the infection. In Section 4, we will report the estimated evolution of $\mathcal{R}_{0,t}$ and $\mathcal{R}_{e,t}$.

We summarize all the variables and parameters of the model in Table 1. For ease of exposition, we partition Table 1 into five parts. The first part outlines the model's compartments, the second part summarizes the model's time-varying parameters, the third part presents observed variables, the fourth part shows the constant parameters of the model, and the fifth part lists the initial states. We assume that the initial states' distribution is degenerate and treat them as additional parameters with their corresponding priors. In addition, we let $\boldsymbol{\Theta}$ denote a vector containing all the time-invariant parameters of the state and measurement equations and the initial states $\boldsymbol{X}_0$.

3 Data and Estimation Procedure

We build daily frequency data, including data on deaths in hospitals, total deaths, hospitalizations, and new cases from Sciensano, a public institution recognized as a research institution by the Belgian Science Policy Office. All the data except for deaths in hospitals, which will be explained below, were downloaded from Sciensano on December 17, 2020. Our data span the period from March 15, 2020, to November 30, 2020. The starting date corresponds to the first day for which we have data on COVID-19 hospitalizations, and the ending date is right before the announcement of the Belgian vaccination campaign. Deaths at home are computed as the difference between total deaths and deaths in hospitals. The latter information was obtained upon request from Sciensano and it is dated December 15, 2020. In a few days, the number of

Table 1: Model Road Map

Model Compartments	
S_t	Share of the population that is susceptible at time t.
I_t	Share of the population that is infectious at time t.
H_t	Share of the population that is hospitalized at time t.
P_t	Share of the population that is recovering at home (outside the hospital) at time t.
$D_{H,t}$	Share of the population that has died in a hospital as of time t.
$D_{P,t}$	Share of the population that has died at home as of time t.
C_t	Share of the population that has recovered as of time t.
Model Time-Varying Parameters	
β_t	Effective contact rate at time t.
b_t	Natural logarithm of the effective contact rate, i.e., $b_t = \log(\beta_t)$.
$\gamma_{H,t}$	Share of the population no longer infectious at time t because they are recovering in a hospital.
$g_{H,t}$	Inverse of the standard logistic function mapping $g_{H,t}$ to $\gamma_{H,t}$, i.e., $g_{H,t} = \log(\gamma_{H,t}/(1-\gamma_{H,t}))$.
$\delta_{H,t}$	Share of those leaving the hospital at time t due to death.
$d_{H,t}$	Inverse of the standard logistic function mapping $d_{H,t}$ to $\delta_{H,t}$, i.e., $d_{H,t} = \log(\delta_{H,t}/(1-\delta_{H,t}))$.
$\delta_{P,t}$	Share of those no longer recovering at home at time t due to death.
$d_{P,t}$	Inverse of the standard logistic function mapping $d_{P,t}$ to $\delta_{P,t}$, i.e., $d_{P,t} = \log(\delta_{P,t}/(1-\delta_{P,t}))$.
$\gamma_{n,t}$	Share of new cases at time t detected in the data.
n_t	Inverse of the standard logistic function mapping n_t to $\gamma_{n,t}$, i.e., $n_t = \log(\gamma_{n,t}/(1-\gamma_{n,t}))$.
Observed Variables	
$D_{H,t}^{obs}$	Observed share of the population that has died in a hospital as of time t.
$D_{P,t}^{obs}$	Observed share of the population that has died at home as of time t.
H_t^{obs}	Observed share of the population that is hospitalized at time t.
G_t^{obs}	Observed new cases at time t as a share of the population.
S_t^{obs}	Observed share of the population that is susceptible at time t.
Model Constant Parameters	
γ	Share of the population that is no longer infectious at time t.
θ_H	Share of the population that leaves the hospital at time t.
θ_P	Share of the population that is no longer recovering at home at time t.
σ_b	Standard deviation of the innovation to b_t.
σ_g	Standard deviation of the innovation to $g_{H,t}$.
σ_h	Standard deviation of the innovation to $d_{H,t}$.
σ_p	Standard deviation of the innovation to $d_{P,t}$.
σ_n	Standard deviation of the innovation to n_t.
σ_{D_H}	Standard deviation of the innovation to the measurement equation of $\Delta D_{H,t}$.
σ_{D_P}	Standard deviation of the innovation to the measurement equation of $\Delta D_{P,t}$.
σ_H	Standard deviation of the innovation to the measurement equation of H_t.
σ_G	Standard deviation of the innovation to the measurement equation of $S_{t-1} - S_t$.
σ_S	Standard deviation of the innovation to the measurement equation of S_t.
μ	Upper bound of the share of detected cases.
Initial Values	
S_0, I_0, H_0, P_0, b_0, $g_{H,0}$, $d_{H,0}$, $d_{P,0}$, n_0	

total deaths in hospitals was above the total number of deaths. In such cases, we input zero deaths at home.

The raw data on new cases exhibit weekend and holiday effects. Hence, we use the data from new cases reported every Wednesday save for Wednesday, November 11, 2020, a public holiday in Belgium (Remembrance Day). For this date, we use the reported cases on Tuesday, November 10, 2020. We consider that those cases were infected the previous Friday. Since Sciensano assigns new cases according to the date on which the sample was taken, we impute new cases in the model to 5-day-ahead reported new cases.[7] Thus, we match an average incubation period of 5 days. As the CDC reports, symptoms may appear 2-14 days after exposure to the virus and some studies consider day 5 as the typical day of the onset of symptoms (Kucirka et al., 2020).

We picked this specification after extensive testing of alternatives. In those preliminary tests, we found that alternative approaches that seasonally adjust for nonbusiness days distorted the number of reported cases and the nonlinear nature of the time series.[8] Hence, we use the data for new cases at a weekly frequency; this is the most transparent solution that we found.

Concerning the share of the susceptible population, we use the point estimates from Herzog et al. (2020). This study is a prospective serial cross-sectional nationwide seroprevalence evaluation conducted in Belgium using blood samples collected during five periods: March 30-April 5, April 20-April 26, May 18-May 25, June 8-June 13, and June 29-July 3. These five collection periods include a total of 33 days or about 15% of our sample. To express variables in per capita terms, we use the population from 2019 obtained from the World Bank's website.

We combine the particle filter to evaluate the likelihood with a Metropolis-Hastings algorithm to draw from the posterior (the joint algorithm is often called the PFMH). We tune the MH proposal so that the resulting acceptance rate of the algorithm is close to 20% based on the adaptation strategy of Roberts and Rosenthal (2001). We take $100,000$ draws with a burn-in of $10,000$. Using the draws from the algorithm, we approximate the posterior moments of interest such as the mean, and the 5% and 95% percentiles of $\boldsymbol{\Theta}$. We approximate the posterior distribution of states using the particle smoother as in Doucet and Johansen (2009).

4 Estimation

4.1 Priors

Table 2 reports the prior distributions for the parameters and initial states of the model. We select well-known parametric families commonly applied to similar problems and that capture,

[7] According to the data documentation from Sciensano, the diagnostics include molecular techniques (i.e., polymerase chain reaction or PCR) and rapid antigen tests.

[8] Similarly, the nonlinearity of the model made the use of cumulators to aggregate cases over the week challenging to implement and generated results that were not transparent.

when available, previous existing knowledge about COVID-19.

Table 2: Prior and Posterior Moments

Parameter/ Initial State	Prior			Posterior			
	Dist.	Mean	Std	Mode	Mean	5%	95%
γ	B	0.20	0.05	0.072	0.076	0.055	0.106
θ_H	B	0.10	0.02	0.211	0.195	0.161	0.232
θ_P	B	0.15	0.05	0.131	0.139	0.083	0.207
σ_b	IG	0.20	0.20	0.159	0.182	0.123	0.253
σ_g	IG	0.20	0.20	0.166	0.175	0.138	0.220
σ_h	IG	0.20	0.20	0.085	0.090	0.063	0.121
σ_p	IG	0.20	0.20	0.129	0.147	0.116	0.182
σ_n	IG	0.20	0.20	0.264	0.343	0.170	0.730
S_0	B	0.98	0.01	0.973	0.977	0.960	0.991
I_0	B	0.001	0.001	0.0008	0.0014	0.0006	0.0026
b_0	N	log(0.60)	0.50	log(0.349)	log(0.371)	log(0.237)	log(0.595)
$g_{H,0}$	N	logit(0.01)	2.00	logit(0.236)	logit(0.148)	logit(0.065)	logit(0.316)
$d_{H,0}$	N	logit(0.15)	2.00	logit(0.144)	logit(0.136)	logit(0.098)	logit(0.187)
$d_{P,0}$	N	logit(0.15)	2.00	logit(0.043)	logit(0.025)	logit(0.009)	logit(0.069)
n_0	N	-1.00	1.00	-1.028	-1.483	-2.585	-0.571
μ	G	0.175	0.05	0.119	0.123	0.079	0.172
Mean $\widehat{p}(\boldsymbol{\Theta}\|\boldsymbol{y}^T)$ (Std)				281.24 (1.34)	280.82 (1.11)		

Note: This table is based on 90,000 posterior draws obtained after a burn-in period of 10,000 draws. The number of particles used in the estimation is 50,000. The acceptance rate is about 16%. The mean and standard deviation of the log posterior density are based on 500 evaluations at the posterior mean and mode, respectively. $\widehat{p}(\boldsymbol{\Theta}|\boldsymbol{y}^T)$ denotes the estimated posterior distribution of the time-invariant parameters and initial conditions given the observables.

We pick beta priors for γ, θ_H, and θ_p. The prior for γ is centered at 0.2 (i.e., the median time a person remains infectious is about five days), consistent with the evidence in Bar-On et al. (2020). The standard deviation of γ is 0.05, with an interdecile range (IDR) of 4 to 7 days.[9] The prior mean and standard deviation for θ_H are 0.10 and 0.02, respectively, implying a median length of average stays in hospitals of about 10 days and an IDR of 6 to 13 days, broadly in line with the evidence in Catteau et al. (2020), who found a median hospital stay in Belgium of 9 days with a 6– to 15-day interquartile range. We set the prior mean for θ_P to 0.15, implying a median average home recovery of between 6 and 7 days and an IDR of about 4 to 11 days. Our prior assumes that, on average, individuals recovering outside of hospitals present less severe symptoms and hence are more likely to recover faster.

We pick inverse gamma priors for σ_b, σ_g, σ_h, σ_P, and σ_n. Inverse gamma priors are popular choices for the step size changes in related applications. We set the prior mean to 0.2 for σ_b, σ_g, σ_h, σ_p, and σ_n to allow for quick changes in behavior, policies, and medical treatments. For the

[9]The standard deviation of γ embodies uncertainty about the average duration of spells of infectiousness in the population, not about their dispersion in the cross-section of individuals.

effective contact rate, our prior for σ_b implies that this rate varies, on average, by about 20% from one day to the next, in line with the fast changes recorded after stay-at-home orders. For the law of motion governing the share of those infected who are recovering in a hospital, our prior implies (up to first order) that this share changes on average by about $20 \times (1 - \gamma_{H,t-1})\%$ from $t-1$ to t.[10] Thus, when the share of the hospitalized population is small, we expect larger changes than when this share is large. Similarly, the prior mean of σ_h is 0.2, implying (up to first order) that the death rate in a hospital changes on average by about $20 \times (1 - \delta_{H,t-1})\%$ from $t-1$ to t. This specification allows death rates in hospitals to fluctuate more when the death rate is low. The persistence of the death rate implied by the prior increases with its level, as it is likely to be subject to shocks that are proportionally smaller when the death rate level is high. We impose an identical prior for σ_p. With regard to σ_n, our prior implies that (up to first order) the share of detected cases changes on average by about $20 \times (1 - \gamma_{n,t-1})\%$ from $t-1$ to t. Thus, when the detection rate is low, the day-to-day changes are larger than when the detection rate is high. The standard deviation of the prior for σ_b, σ_g, σ_h, σ_p, and σ_n is 0.2. Hence, our prior allows for either slow or fast changes in the time-varying parameters.

We set the prior mean for S_0 to 0.98. This is about one percentage point (p.p.) above the share of the population susceptible to the virus as of March 30, 2020, according to the point estimates in Herzog et al. (2020). The standard deviation is set to 0.01, about two times the standard deviation associated with the 95% confidence interval for the seroprevalence of COVID-19 in Belgium as of March 30, 2020, also from Herzog et al. (2020). The prior mean for I_0 is 0.001, which implies that 0.1% of the Belgian population was infectious on March 15, 2020. Thus, assuming a contagion rate of 0.6 (a value we will describe momentarily), the number of new cases on March 16, 2020 is 6,752, broadly consistent with the number implied by inflating the reported cases under our prior for the share of detected cases as discussed below. The standard deviation of I_0 is 0.001, also in line with our assumptions on the detection rate. Since the priors for S_0 and I_0 are independent, we discard draws of S_0 and I_0 that violate feasible initial conditions for the model.

The prior mean for b_0 is $\log(0.6)$, so that prior belief about $\mathcal{R}_{0,t}$ is around 3 on March 15, 2020, a value consistent with D'Arienzo and Coniglio (2020). The standard deviation for b_0 is 0.5, so that values of b_0 one standard deviation below and above the mean yield values of R_0 in the interval [1.8, 5.9]. The prior mean for $g_{H,0}$ is set so that when mapped to $\gamma_{H,0}$ it implies that 1% of those recovering from COVID-19 were doing so in a hospital on March 15, 2020.[11] The standard deviation for $g_{H,0}$ is 2, which is large enough to cover an interval of hospitalized individuals between $[15; 800]$, which is wide given the reported 266 hospitalizations. The prior

[10] This follows from linearizing $\gamma_{h,t} = \frac{e^{g_{H,t}}}{e^{g_{H,t}}+1}$ around $\gamma_{h,t-1}$ and dividing the resulting expression by $\gamma_{h,t-1}$.

[11] We sum the "true COVID-19" cases embedded in our prior as implied by a detection rate of 5.5% (a plausible value according to our prior) from March 11 until March 15. The result is 27,054. Since 266 individuals were hospitalized on March 15, we have that $266/27,054 \approx 0.01$.

mean for $d_{H,0}$ is log(0.15/(1-0.15)). This number is obtained by the measurement equation assuming that deaths in hospitals are measured without error on March 15, 2020. The standard deviation for $d_{H,0}$ is 2 to concentrate the prior for the probability of those leaving the hospital because of death in the interval $[0.02, 0.56]$. The prior mean and standard deviation for $d_{P,0}$ are the same as the ones for $d_{H,0}$.

The prior mean for n_0 is -1.00 and the standard deviation is 1.00. This choice, together with the prior over μ (described below), implies a fairly flat prior over the share of detected cases on March 15, 2020, as shown in Appendix A.1. We have degenerate priors for H_0 and P_0. In particular, we set $H_0 = 2.31 \times 10^{-5}$ (that is, 266 hospitalized expressed as a share of the Belgian population the day before the start of our sample) and $P_0 = 0$. The first is given to us by the data and the second by convenience, since its effects on the estimation are trivially small.

The prior mean of μ is 0.175, the mid-point of the 2 to 33% interval of false negatives in PCR tests reported by Arevalo-Rodriguez et al. (2020). The standard deviation for μ is 0.05, with an IDR range of 11 to 24% for the false negative rate. In that way, we allow for large differences in the accuracy of PCR tests as measured in a lab and when applied in practice.

We impose degenerate priors for σ_{D_H}, σ_{D_P}, σ_H, and σ_G by applying a Hodrick-Prescott filter to the corresponding observed variables. We fine-tune the filter so that the cyclical component is nearly serially uncorrelated (i.e., we pick the filter's smoothing parameter λ so that the cyclical component has little serial correlation). In that way, the measurement error does not explain more than 5% of the variation in each individual series. Then, we set these parameters equal to the standard deviation of the serially uncorrelated cyclical component. This procedure results in $\sigma_{D_H} = 0.29$, $\sigma_{D_P} = 0.32$, $\sigma_H = 0.017$, and $\sigma_G = 0.16$.

Finally, we also impose a degenerate prior for σ_S by fixing it to 0.02, roughly about 5 times the average standard deviation across collection periods implied by Herzog et al. (2020). Serological studies are a useful guide, but likely subject to large measurement error, as is evident in the variation in the share of individuals infected by the virus throughout the collection periods.

4.2 Posteriors

We now discuss the posterior mode, mean, and the 90% posterior probability interval of the parameters and initial states, also reported in Table 2. For ease of exposition, we focus on the posterior mean, save for a few parameters where the 90% posterior probability interval is of particular interest. Appendix A.2 shows the prior and posterior distribution for all of the estimated parameters, which show that the data are informative about the estimated parameters.

The posterior mean and 90% posterior probability interval for γ are 0.076 and $[0.055; 0.106]$, respectively. This suggests that the implied posterior mean estimate of the average length a person remains infectious is between 13 and 14 days and the 90% posterior probability interval

is between 9 and 19 days (we compute the statistics regarding days by inverting every draw of γ, not by inverting the mean of γ; the same will apply to all other statistics regarding time). The posterior mean and 90% posterior probability interval for θ_H are 0.195 and $[0.161; 0.232]$, respectively. Hence, the posterior mean of the average length of stay in hospitals is between 5 and 6 days and the 90% posterior probability interval is between 4 and 7 days. The posterior mean and 90% posterior probability interval for θ_P are 0.139 and $[0.083; 0.207]$. The implied posterior mean of the average length of stay at home is between 7 and 8 days, and the 90% posterior probability interval is between 4 and 13 days. The likelihood overturns the prior for θ_H: our point estimates suggest a longer recovery at home, perhaps due to higher-quality treatments given in hospitals.

The posterior distribution for σ_b is centered around the prior mean, 0.20, but it is more concentrated than our prior, ruling out small and abrupt changes in daily behavior. A similar conclusion emerges for σ_g. In contrast, the posterior mean for σ_h, 0.090, is much lower than its prior mean. In the case of the mortality rates at home, σ_p, the posterior mean is 0.147, which is a step-size value somewhat higher than in the case of mortality rates in hospitals. The posterior mean for changes in the share of detected cases is 0.343, which is larger than our prior mean. This is consistent with expedited increases in testing capacity.

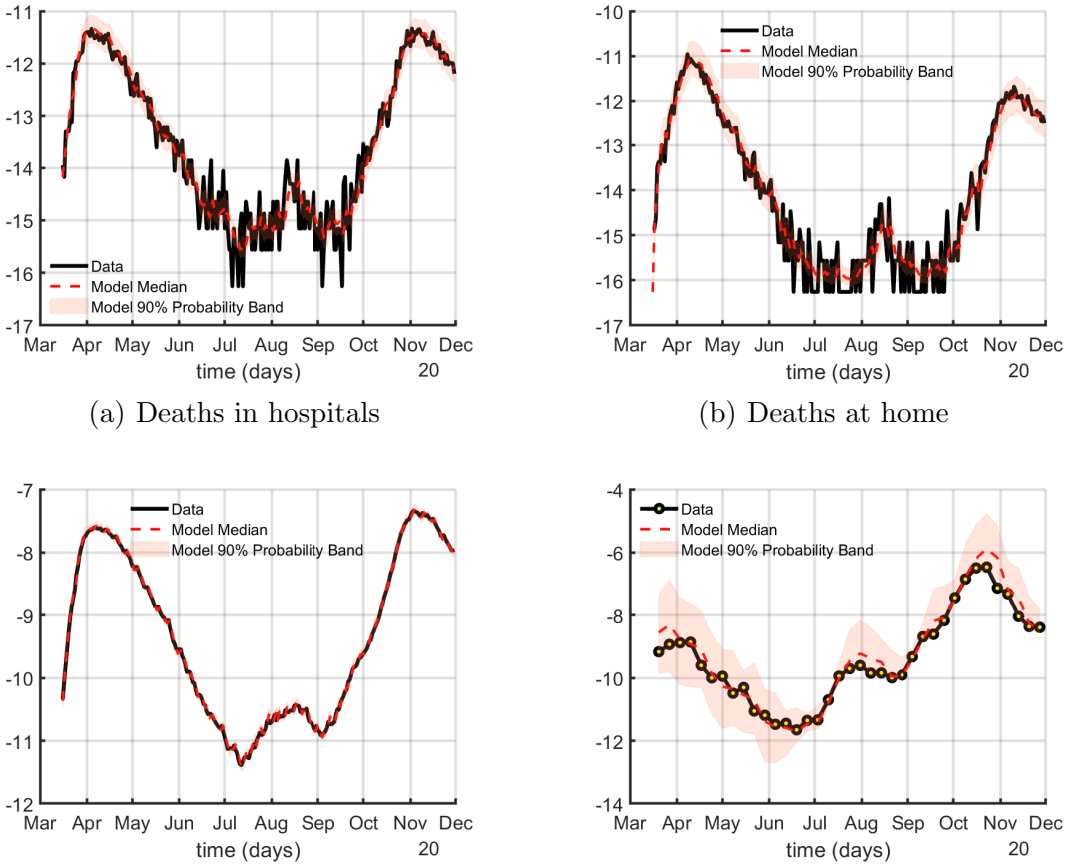
The posterior mean and 90% posterior probability bands for the initial share of the population susceptible to the virus are 0.977 and $[0.960; 0.991]$, respectively. Thus, our posterior estimates indicate that, with high probability, about 2% of the Belgian population had come in contact with the virus by mid-March 2020. The posterior mean and posterior probability bands for the initial share of the population that was infectious are 0.0014 and $[0.0006; 0.0026]$, respectively. Hence, about 0.1% of the Belgian population was infectious with high probability by around March 15. The posterior mean of the initial log of the transmission rate is $\log(0.371)$, implying an $\mathcal{R}_{0,t}$ of around 5.3.

We estimate that 14.8% of those infected with COVID-19 were recovering in hospitals at the start of the sample, an initial death probability of 2.7% conditional on being hospitalized, and an initial death probability of 0.39% conditional on recovering at home. These findings suggest that, by mid-March 2020, only very sick patients were being admitted to hospitals. As documented in Appendix A.2, the posterior distribution n_0 is very similar to the prior distribution, with roughly the same mean. However, our choice of prior for n_0 was an educated guess: since this parameter was hard to identify, we optimized n_0 by starting our Monte Carlo simulation at different initial values. The optimizer consistently delivered values concentrated around the mean of our prior.

The posterior mean of the share of false negatives implied by our model is 12.3%. The posterior distribution of parameters μ and n_0 imply a 90% posterior probability interval for the initial share of detected cases of between 6 and 32%. Thus, as of March 15, 7 out of 10 cases were going undetected in Belgium.

4.3 Data Fit

Figure 2 shows the fit of the model to the data at the posterior mean (these are in-sample forecasts based on filtered states). Panel (a) plots, in a continuous black line, the (log) first difference of the share of (daily) deaths in hospitals in Belgium from mid-March until late November (i.e., the share of new deaths). We see two large peaks, one in early April and one in early November, corresponding to the first and third waves of COVID-19 in Belgium, plus a smaller wave in mid-August. In red, we plot the median of the one-step-ahead forecast of the model, evaluated at the posterior mean, together with a 90% probability band. Panel (b) plots, following the same formatting as the left panel, the (log) first difference of the share of the deaths outside of hospitals, the one-step-ahead forecast, and the 90% probability band.



(a) Deaths in hospitals

(b) Deaths at home

(c) Population in hospitals

(d) New reported cases

Figure 2: Deaths, hospitalized, and new cases: One-step-ahead forecast and data, Belgium

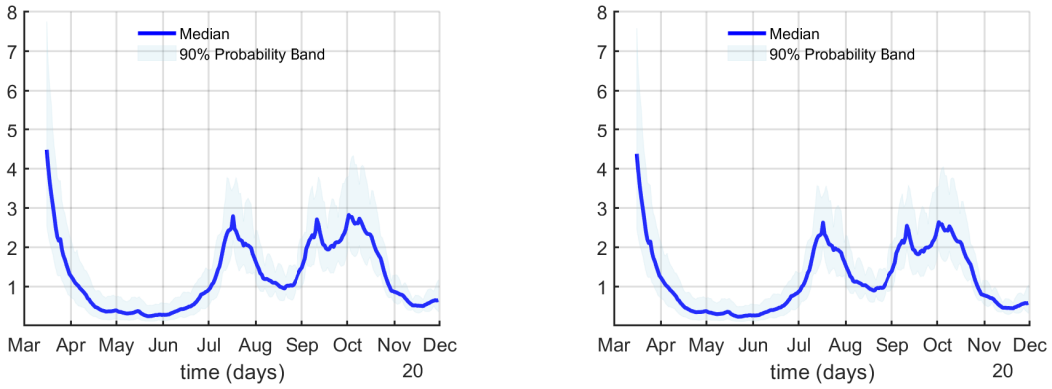
The model captures the three waves of deaths. The observations fall within the 90% probability band on most days. The only misses are during the summer of 2020, when deaths were fewer than ten a day. During this time, small random differences between one period to the next are extremely difficult to forecast. The model also accounts for the observation that deaths

in hospitals had reached, by early November, the same level as in early April, but deaths at home had not. This might reflect Belgian hospitals' better ability to cope with severe COVID-19 patients due to additional available beds.[12]

This hypothesis is supported by Panels (c) and (d), where we plot the (log) share of hospitalizations following the same formatting as above. Panel (c) shows that the model can account for hospitalization. Finally, Panel (d) draws the (log) share of newly reported cases. The data indicate many more cases in the last peak than in the first. This panel plots reported cases and our forecast of those recorded cases. We will come back to this point below.

4.4 The Estimated Time-Varying Reproduction Numbers

From our posterior distribution, we recover the smoothed value of the unobserved states (or related variables) and their probability distribution. These are, often, the relevant objects for policymaking, as they describe the epidemiological situation of a given area (country/region/...) under study and can be used, as inputs, in a loss function to pick an optimal policy.



(a) $\mathcal{R}_{0,t} = \beta_t/\gamma$ (b) $\mathcal{R}_{e,t} = \mathcal{R}_{0,t}S_t$

Figure 3: Reproduction numbers, Belgium

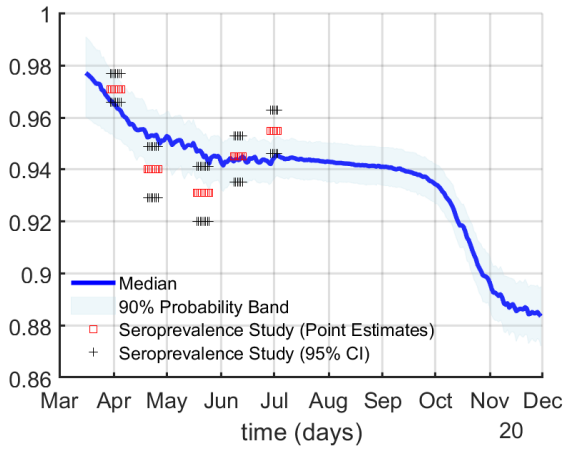
A key variable in an epidemic is $\mathcal{R}_{0,t}$. The left panel of Figure 3 plots the smoothed $\mathcal{R}_{0,t}$ and the 90% smoothed band. We see that $\mathcal{R}_{0,t}$ started in mid-March slightly above 4, around the values that several clinical studies have suggested. For example, see Table 1 in Katul et al. (2020) for a list of estimates (the authors conclude that their best estimate of an unmitigated $\mathcal{R}_{0,t}$ for the early variants of COVID-19 is 4.5, close to our result). However, $\mathcal{R}_{0,t}$ fell rapidly and, by mid-April, it was well below 1. In May and June, $\mathcal{R}_{0,t}$ stabilized around 0.5 and increased back to almost 3 during the summer of 2020, fluctuating until late October between 3 and 1. After November, $\mathcal{R}_{0,t}$ was again at a much lower level, with the whole 90% smoothed band below 1.

[12]We could also plot the one-step-ahead forecast integrating over the whole posterior. This alternative exercise makes little difference in practice, but complicates the interpretation of the figures.

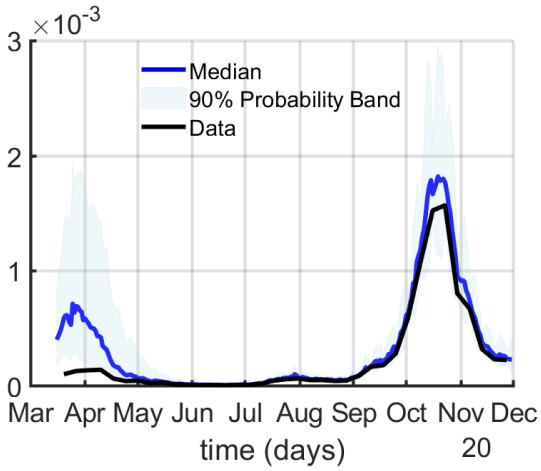
The right panel of Figure 3 reports our smoothed median and 90% probability estimates of $\mathcal{R}_{e,t}$, the effective reproduction number. While $\mathcal{R}_{e,t}$ starts at the same level as $\mathcal{R}_{0,t}$, as the susceptible population shrinks, $\mathcal{R}_{e,t}$ falls. The difference, however, between $\mathcal{R}_{0,t}$ and $\mathcal{R}_{e,t}$ is small. Panel (a) of Figure 4, our smoothed estimate of the share of susceptibles, tells us why. Even as late as November, our model estimates that only around 12% of the population had ever been infected. By early December 2020, Belgium was far away from herd immunity.

4.5 New Cases

Panel (a) of Figure 4 plots, in red squares, the point estimates of the seroprevalence studies reported by Herzog et al. (2020) and, in black crosses, the 95% confidence interval of the studies. Our model accounts for those seroprevalence studies: in all cases, our 90% probability band and the 95% confidence interval of the studies overlap, which is crucial to recover a measure of the true number of cases. Panel (b) of Figure 4 plots, in blue, the smoothed median and 90% probability bands of the true new cases, and, in black, the reported new cases in Belgium. These estimates will be the input in our causal assessment of NPIs. The first peak is much smaller, by either measure, than the third one (with a minor second peak in the middle). The fall in death probabilities explains why deaths did not reach the first wave levels.



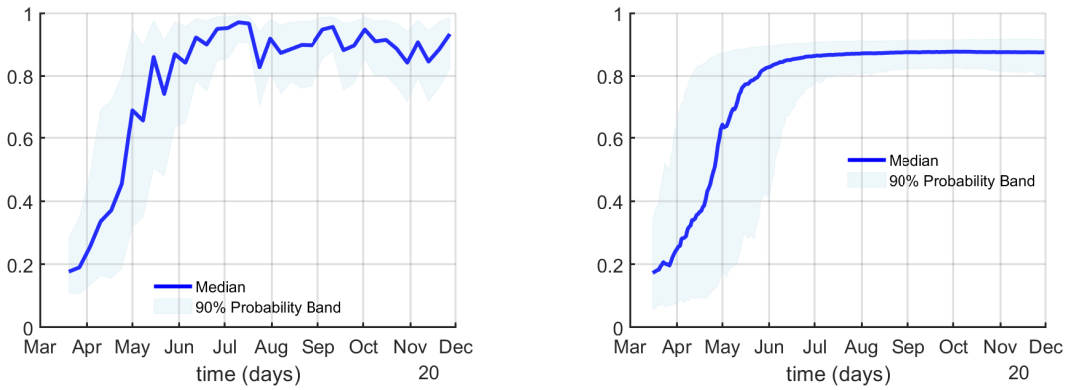
(a) Share of susceptible population, Belgium



(b) New cases, Belgium

Figure 4: New cases, Belgium

In addition, the reported cases and the estimated cases were different during the first wave. Figure 5 makes this point clearly by showing, in the left panel, our median smoothed estimate of the share of reported cases and, in the right panel, the permanent component of this share. By late March, less than 20% of all cases were being reported, while by the summer of 2020, after testing became more prevalent, around 90% of cases were being reported. The permanent component of this share suggests that this increase in reported cases is very persistent.



(a) Share of reported cases, Belgium (b) Permanent component, $\gamma_{n,t}$, Belgium

Figure 5: Share of reported cases, Belgium

For completeness, Appendix A.3 presents the smoothed estimates of the time-varying death probabilities, the share of the population that is infectious, the share of the population that is recovering, and the inflow of hospitalizations as a share of those that are no longer infectious.

5 Causality and Policy Trade-offs

This section measures, using two different approaches, the causal effects of mobility curtailment policies on the spread of the virus, the death toll, and economic activity using the smoothed estimates of variables from the epidemiological model such as the effective reproduction number, new cases, deaths in hospitals, and deaths at home. These exercises gauge the trade-off between slowing down the spread of the virus and decreasing economic activity that policymakers face when enacting shelter-in-place and/or compulsory business closure orders.

Why do we need a causality assessment? Starting in March 2020, individuals' mobility in most countries plummeted due to COVID-19, slowing down the virus's spread at the cost of lower economic activity. Some of the reductions in mobility were voluntary, as individuals took precautions to avoid getting infected (or were affected by other individuals taking such measures; for instance, a household canceling its home cleaning services to avoid having third parties come inside its dwelling reduces the mobility of the workers of the cleaning service). Some of the reductions in mobility were triggered by government mandates, such as orders to shelter-in-place or compulsory business closures.

Since both mechanisms coincided in time, we need to disentangle them to ascertain the causality effect of the government mandates on the virus's spread, the death toll, and economic activity. Knowing by how much governments can affect current and future epidemiological

conditions and gauging their cost in terms of economic activity are key factors when designing the length and severity of mobility curtailments.

Why are the outputs from the estimated epidemiological model useful? One challenge when disentangling the effects of voluntary and government-mandated changes in behavior is that some of the relevant variables required to do so, such as $\mathcal{R}_{e,t}$, are not directly observable. Furthermore, other variables, such as the number of reported cases, are subject to large, persistent, biased, and time-varying measurement errors.

The structure imposed by an epidemiological model allows us to tackle these data limitations. We can discipline the data by enforcing the cross-equation restrictions among the model's states implied by the transitions among compartments dictated by the disease's biological and clinical properties. The likelihood of the model tells us, for example, that in the case of COVID-19, relatively high seroprevalence rates and low reported new cases in Belgium during the first half of 2020 can only be reconciled with large under-reporting of cases. Likewise, the parallel evolution of hospitalizations, deaths in hospitals vs. deaths at home, and the reported new cases during 2020 informs us of the evolution of under-reporting and changes in mortality probabilities. Thanks to our use of an epidemiological model, we transform noisy and biased observations of new cases, new deaths in hospitals and at home, and seroprevalence surveys into useful unobserved outputs, such as smoothed estimates of $\mathcal{R}_{e,t}$ or the (true) number of new cases.

In contrast, one could estimate $\mathcal{R}_{e,t}$ using statistical methods (e.g., comparing newly reported cases along a moving window), but such an exercise could not correct the changing share of unreported cases. Another option would be to perform the causality and policy analyses without $\mathcal{R}_{e,t}$, but such a policy analysis would suffer from an omitted variable bias, which we know from other environments (like studying the effects of monetary policy) can be a serious flaw.

Two procedures for causality assessment We address the effects of government-mandated mobility curtailments with two popular methods for assessing causality in time series: structural vector autoregressions (SVARs) and local projections (LPs).

First, we build on the tradition of SVARs and identify a government stringency shock by restricting the systematic component of the government stringency policies rule that maps health and economic conditions into mobility curtailments. A stringency shock should be thought of an *unexpected* change in the mobility curtailment policy. The identification assumptions are motivated by the fact that health policymakers have emphasized that they follow a data-driven approach when imposing mobility curtailments. In situations where we can impose credible restrictions on the systematic component of policy, SVARs offer reliable answers (Wolf, 2020).

Second, we work with LPs to identify a reproduction shock and analyze how it affects the rest of the variables (including deaths in hospitals and at home and economic conditions) depending

on the level of government mobility curtailments. A reproduction shock can be thought of as a shock that changes the contagious properties of the virus. This can be due to biological factors (e.g., a new variant of the virus) or social mechanisms (e.g., better facial masks, improvements in ventilation in public spaces). More concretely, we want to measure whether government mobility curtailments affect the transmission of the reproduction shock. LPs are a flexible approach that allows us to address state dependencies without making strong parametric assumptions.

Many other methods for assessing causality could use the output from our estimated epidemiological model, such as a vector error correction model, regression discontinuity designs that exploit local variations in government mandates (e.g., Goolsbee and Syverson, 2020), synthetic controls (e.g., Cho, 2020), or event studies (e.g., Gupta et al., 2020). We could also generalize our SVAR analysis to a Markov-switching SVAR à la Sims and Zha (2006). We skip all those additional experiments to keep our study focused.

5.1 A Government Stringency Shock

We write our SVAR as:

$$\boldsymbol{y}_t'\boldsymbol{A}_0 = \boldsymbol{x}_t'\boldsymbol{A}_+ + \boldsymbol{\varepsilon}_t' \quad \text{for } 1 \leq t \leq T,$$

where $\boldsymbol{y}_t$ is an $n \times 1$ vector of endogenous variables, $\boldsymbol{x}_t' = \begin{bmatrix} \boldsymbol{y}_{t-1}' & \cdots & \boldsymbol{y}_{t-p}' & \boldsymbol{z}_t' & 1 \end{bmatrix}$, $\boldsymbol{z}_t$ is a $z \times 1$ vector of exogenous variables, $\boldsymbol{\varepsilon}_t$ is an $n \times 1$ vector of structural shocks, $\boldsymbol{A}_0$ is an $n \times n$ invertible matrix of parameters, $\boldsymbol{A}_+$ is an $(np + z + 1) \times n$ matrix of parameters, p is the lag length, and T is the sample size. The vector $\boldsymbol{\varepsilon}_t$, conditional on past information and the initial conditions $\boldsymbol{y}_0, ..., \boldsymbol{y}_{1-p}$, is Gaussian with mean zero and covariance matrix $\boldsymbol{I}_n$ (the $n \times n$ identity matrix). The matrices $\boldsymbol{A}_0$ and $\boldsymbol{A}_+$ are the structural parameters.

One of the equations in the SVAR characterizes the policymaker's behavior when imposing mobility curtailments aimed at slowing the transmission of the virus. Consequently, we call such an equation the government stringency policy rule.[13]

We summarize the decisions of the policymakers using a government stringency policies indicator that we will describe below. We will call this indicator the policy instrument. We assume the indicator reacts to other variables in the system such as new cases and $\mathcal{R}_{e,t}$. Without loss of generality, we assume that the first equation of the SVAR characterizes the policy rule. This implies that

$$\boldsymbol{y}_t'\boldsymbol{a}_{0,1} = \boldsymbol{x}_t'\boldsymbol{a}_{+,1} + \boldsymbol{\varepsilon}_{1t} \quad \text{for } 1 \leq t \leq T$$

is the policy equation, where $\boldsymbol{\varepsilon}_{1t}$ denotes the first entry of $\boldsymbol{\varepsilon}_t$, $\boldsymbol{a}_{+,1}$ denotes the first column of $\boldsymbol{A}_+$ for $0 \leq \ell \leq p$, and $a_{s,ij}$ for $s \in \{0, +\}$ denotes the (i, j) entry of $\boldsymbol{A}_s$ and describes the

[13] We could have several policy equations, each capturing one different containment policy as a function of public health variables, as in Chernozhukov et al. (2021). We keep our analysis to one policy equation to avoid overparameterizing the SVAR.

systematic component of the policy rule. Thus, restricting the systematic component of the policy rule is equivalent to restricting $a_{s,ij}$ for $s \in \{0, +\}$ and identifies a policy shock that we call the stringency shock.

Our baseline SVAR sample runs from March 16 through November 30 and contains seven endogenous variables. Three of these seven endogenous variables come from outside sources. First, we use the Oxford Stringency (OS) index for Belgium as our mobility curtailment policies indicator (Hale et al., 2020). The authors of the OS index compile information on when and which measures governments take. The particular index we use is a simple average of nine individual component indicators. Each component is a measure of intensity of closings of schools and universities, closings of workplaces, canceling of public events, limits on gatherings, the closing of public transportation, orders to shelter-in-place, curtailments on internal movement between cities/regions, prohibitions on international travel for non-citizens, and the presence of public information campaigns. Due to data limitations, we do not consider non-mobility-related NPIs such as face masks or improved ventilation.

Second, we use a mobility index for Belgium from the Google COVID-19 Community Mobility Reports, available at https://www.google.com/covid19/mobility/. This measure of mobility is a simple average of the measures of nonresidential mobility categories in the Google Mobility Reports (excluding parks): i) Retail and recreation; ii) Grocery and pharmacy; iii) Transit stations; and iv) Workplaces. The mobility measure is expressed in p.p. and it corresponds to daily changes in mobility relative to a baseline value for that day of the week, which is the median value observed during the 5-week period Jan 3–Feb 6, 2020.

Third, we use a daily economic news sentiment (ENS) indicator for Belgium constructed by Algaba et al. (2021) using natural language processing as a daily index of economic activity. This indicator is based on the media archive of the national Belgian News Agency. When aggregated at a monthly frequency, the index is positively correlated with the National Bank of Belgium's monthly consumer confidence survey and other measures of economic activity such as construction, manufacturing, business-related services, and industrial confidence in the euro area, among others. The daily ENS indicator was obtained in two formats: a latent daily series and a 14-day moving average of the series. We use the latter because the former is too noisy.

The other four endogenous variables come from the estimated epidemiological model. In particular, we use the model-implied point-wise median filtered estimate of the expected value of i) $\mathcal{R}_{e,t}$; ii) new cases; iii) daily deaths per capita in hospitals; iv) and daily deaths per capita at home. For each of these four series, we use a 7-day backward looking moving average because it's the most popular window for these variables.[14]

The OS index, new cases, and $\mathcal{R}_{e,t}$ are expressed in log percent. Daily deaths enter in levels

[14]Appendix A.5 shows that the main conclusions from our SVAR analysis are robust to taking into account the uncertainty associated with our filtered estimates.

so that we can compute the impulse response function (IRF) of cumulative deaths to a stringency shock. The SVAR includes 14 lags, a constant term, and, as an exogenous variable, the average daily temperature to control for the effect that weather conditions might have on the variables of interest. The observable corresponds to the temperature measured at the Brussels Airport Station and it was downloaded from https://www.wunderground.com/. Our measurement is fairly representative of the weather conditions for Belgium as a whole given that this nation's territory is only about 30,689 km^2 and most of the population is concentrated in the flat coastal plain and central plateau.

Notice that the endogenous variables are estimated with a nonlinear epidemiological model, while the SVAR is a linear structure. This does not create major problems. First, we have 14 lags and, thus, the "best linear approximation" implied by the VAR to the evolution of the endogenous variables is very flexible. Second, we selected, in our specification, endogenous variables such as the (log) $\mathcal{R}_{e,t}$ that, without being fully linear, are much closer to linearity than other variables in the epidemiological model (in fact, this was a major factor in our specification design).

The SVAR is estimated with a Bayesian approach following Arias et al. (2018). We impose a normal-generalized-normal (NGN) prior distribution over the structural parameters $(\boldsymbol{A}_0, \boldsymbol{A}_+)$.[15] The NGN prior is a conjugate prior characterized by four parameters $(\nu, \boldsymbol{\Phi}, \boldsymbol{\Psi}, \boldsymbol{\Omega})$. The parameters ν and $\boldsymbol{\Phi}$ govern the marginal prior distribution of $vec(\boldsymbol{A}_0)$: if $\nu = n$ –as will be the case in our application– $vec(\boldsymbol{A}_0)$ is normally distributed with mean zero and variance $\boldsymbol{\Phi}^{-1}$. The remaining parameters $\boldsymbol{\Psi}$ and $\boldsymbol{\Omega}$ govern the prior distribution of $vec(\boldsymbol{A}_+)$, conditional on $\boldsymbol{A}_0$. Such a distribution is normal with mean $\boldsymbol{\Psi} vec(\boldsymbol{A}_0)$ and variance $\boldsymbol{\Omega}$. We set $\nu = n$, $\boldsymbol{\Phi} = \boldsymbol{0}_{n,n}$, $\boldsymbol{\Psi} = \boldsymbol{0}_{mn,n^2}$, and $\boldsymbol{\Omega}^{-1} = \boldsymbol{0}_{mn,mn}$.

The policy equation is identified with sign restrictions on the systematic component of the policy actions in line with Arias et al. (2019) and the SVAR tradition of Leeper et al. (1996). The identification restrictions are:

Restriction 1. *The stringency index is the mobility curtailment policies indicator, and it reacts contemporaneously and positively to $\mathcal{R}_{e,t}$, mobility, new cases, deaths, and the index of economic activity.*

Restriction 1 embodies the idea that policymakers react to current public health conditions and economic activity. For example, on March 17, 2020, the Belgian government announced the first nationwide lockdown arguing that: "The situation has evolved and forced us to take severe measures to stem the spread of the virus."[16] On April 15, the government relaxed some measures since: "We are aware that the measures taken will have serious long-term consequences, both

[15] Appendix A.6 replicates our analysis using the prior robust approach in Giacomini and Kitagawa (2018).
[16] See the March 17, 2020, Reuters article "Belgium to impose coronavirus lockdown from Wednesday."

psychologically and economically. We plead that the measures last as long as necessary."[17] On October 16, the Belgian government ordered bars and restaurants to close, arguing that: "The number of confirmed cases is rising, every day, and not just by a few percentage points." Other sentences such as: "This virus is affecting our country in a very hard way," "Thirty-five people died yesterday from the effects of COVID-19," and "In the days to come, the news will be bad" make clear that policymakers were reacting to public health conditions in real time.[18]

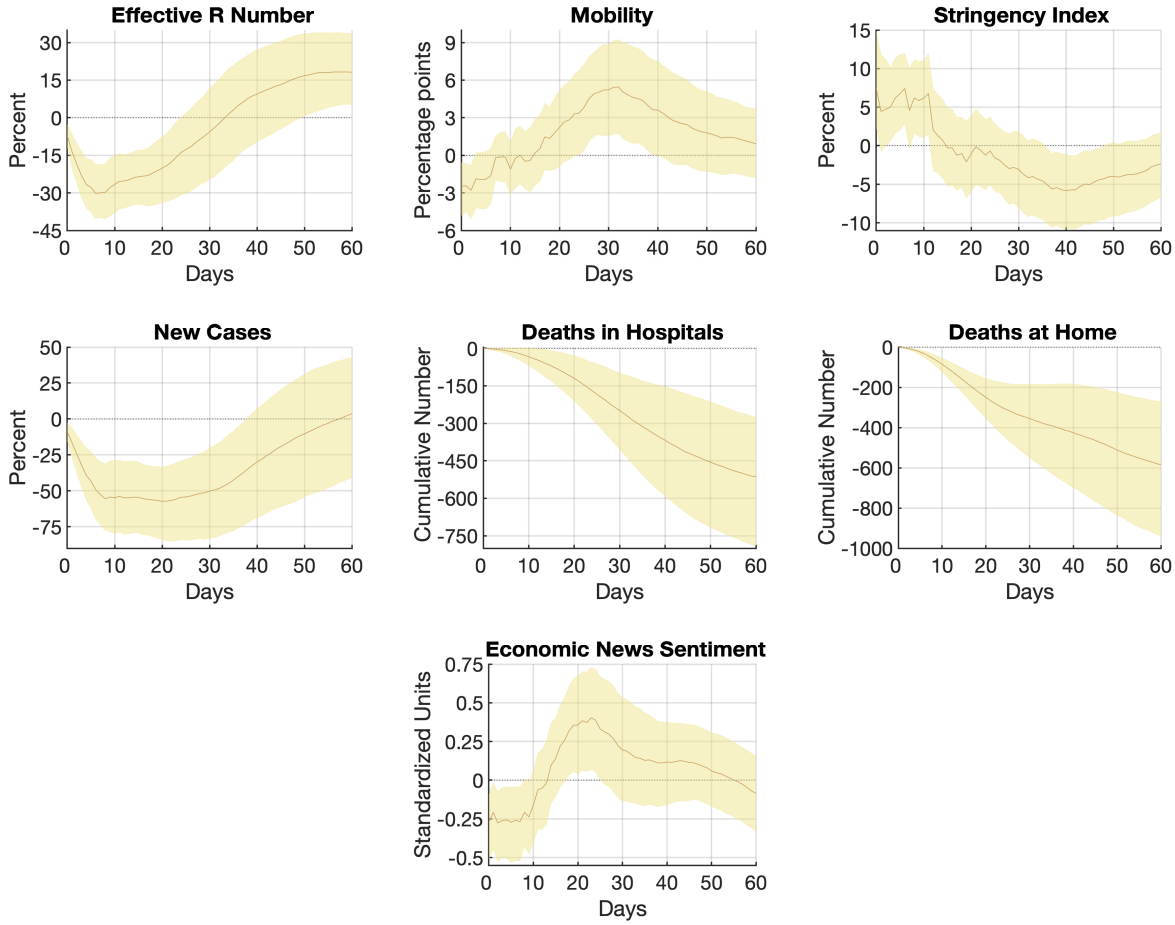


Figure 6: IRFs to a stringency shock. The solid curves represent the point-wise posterior medians, and the shaded areas represent the 68% equal-tailed point-wise probability bands. The figure is based on 10,000 independent draws.

Figure 6 plots (in brown lines) the IRFs to a positive (increase) stringency shock and (in yellow bands) the 68% point-wise posterior probability bands. The stringency shock is normalized such that, upon impact, the posterior median increase in the stringency index equals 7.4%. This is equivalent to a one-unit increase in one of the ordinal measures that composed the OS

[17] See the April 15, 2020, EURACTIV.com article "Belgium extends COVID-19 lockdown until 3 May, but relaxes some measures."

[18] See the October 16, 2020, Reuters article "Belgium to close all bars and restaurants for a month, imposes night curfew," and the AP article "Belgium imposes Covid curfew, closes bars and restaurants," on the same day.

index, on average. For instance, this change roughly corresponds to a nationwide measure that unexpectedly increases the "shelter-in-place" component of the index from 1 (which recommends not leaving the house) to 2 (which requires not leaving the house, with exceptions for daily exercise, grocery shopping, and essential trips), which would increase the OS index by about 6% relative to its median. Other examples include an unexpected nationwide increase in the limits on gatherings from 2 (curtailments on gatherings between 101-1000 people) to 4 (curtailments on gatherings between 11-100 people), which would increase the OS index by about 5% relative to its median, and an unexpected nationwide increase in the public events component from recommending cancelling to requiring cancelling, which would increase the OS index by about 10% relative to its median.

A positive stringency shock leads to a drop in $\mathcal{R}_{e,t}$ for more than a month. This drop is *beyond* the fall triggered by individuals' endogenous responses due to changed risk conditions (which are controlled for by the coefficients in the SVAR). Mobility, new cases, and deaths also decline, albeit with different persistence. In the case of mobility, the decline is more transient than in the case of $\mathcal{R}_{e,t}$, suggesting an exhaustion effect among individuals: after around 8 days, mobility is back to its baseline value without a stringency shock. The decline in deaths in hospitals and at home lasts much longer, reflecting the illness's lag effects. We measure that a positive stringency shock leads to roughly $1,000$ fewer deaths (500 in hospitals and 500 at home) after 2 months. This represents about 6% of the total number of deaths in our sample, $16,840$.

To interpret the effects of a positive stringency shock in terms of output, we scale the ENS indicator following Lewis et al. (2020). We aggregate the daily ENS indicator to quarterly frequency, denote the resulting series by ENS_q, and run a regression of four-quarter real GDP growth one quarter ahead on a constant and ENS_q over the sample 2000Q1-2019Q4, i.e.,

$$\text{RGDPq4q4}_{q+1} = \alpha + \gamma \text{ENS}_q + u_t, \text{ where } q \in \{2000\text{Q1}, \ldots, 2019\text{Q3}\}.$$

We estimate this regression by OLS and obtain $\hat{\alpha} = 1.6$ and $\hat{\gamma} = 1.2$. With these point estimates, we can express the news sentiment indicator in terms of real GDP: A decline of 0.25 standardized units in sentiment implies that, if such a value were to persist for an entire quarter, we would expect (on average) real GDP in the next quarter to be about 0.3% lower than without the shock.

The pointwise posterior median IRF of the ENS indicator to a stringency shock is on average about -0.01 during the first 90 days, with a 68% probability interval of $[-0.27, 0.25]$. If we take the posterior median -0.01 and compute the equivalent in terms of the Belgian GDP per capita, we find that a stringency shock brings a net per capita cost of about €1 (roughy €11 million in the aggregate). Repeating the same exercise for the two extremes of the probability interval, we obtain the result that the per capita cost(-)/benefit(+) of a stringency shock is between −€29

and +€26 (or between −€330 million and €300 million in the aggregate).[19]

We can compare these figures with two alternative measurements. First is the large drop in observed output in Belgium in 2020 (-6.3% according to the IMF World Economic Outlook Database). This drop includes the consequences of voluntary changes in behavior and of government mandates. Researchers have accumulated much evidence that the bulk of reductions in contact rates and employment came from voluntary changes in behavior. A compelling exercise showing this result, due to the granularity of the U.S. data it uses, appears in Arnon et al. (2020). Gupta et al. (2020), Maloney and Taskin (2020a), Andersen (2020), and Maloney and Taskin (2020b) report similar results. The last paper is particularly relevant for us because it gathers findings for many countries. Furthermore, we are focusing on the effects of an *unexpected* change in mobility curtailment policies, not on the effects of the systematic component of the government policy (we will return to this point momentarily). Thus, our finding is consistent with the rest of the literature.

Second, we can compare our findings with the "value of a statistical life" (VSL). OECD (2012), a comprehensive meta-analysis of the literature and one that is often used in policy and regulatory decisions, proposes a range of the VSL for the European Union of €2.2 million-€6.5 million, with a base value of €4.3 million.[20] Since we calculate that the positive stringency shock saves around 1,000 lives, the output cost per life saved is €330,000 or less with 84% probability, well below the VSL above (even correcting by the higher age of those dying of COVID). Notice that we cannot measure the welfare effect of the positive stringency shock (which includes welfare losses, such as missing leisure activities outside the house or contacts with relatives and friends). Our statement merely compares a range of the VSL with output losses.

In summary, and given the systematic component of the health policy of the Belgian government and the voluntary changes in behavior, a marginal and unexpected tightening of mobility curtailments would have saved many lives with close to a zero impact on income per capita and a cost well below the VSL. By controlling the virus's spread, a positive stringency shock has a short-term output cost for about 15 days, but increases economic activity later. As the stringency measures improve public health conditions, the policy's systematic component drives the stringency index below zero about 15 days after the initial increase. In comparison, the reproduction number starts to increase only a month after the shock. This difference in timing is consistent with the presence of transient precautionary behavior by the public.[21]

[19]See Table A.1 in Appendix A.4 for the regression results and see Figure A.7 in Appendix A.4, which plots the ENS indicator in real GDP units along with one-quarter-ahead four-quarter real GDP growth. Given the short time span (90 days) and the low real interest rates prevailing at the moment, discounting the GDP flows to put them in present terms does not make any quantitative difference.

[20]We have transformed the figures from OECD (2012), in 2005 USD, into 2020 euros.

[21]The result that an unanticipated increase in the stringency index leads to fewer deaths at essentially no output cost is robust to not restricting the systematic component of health policy to mobility and economic activity. The only important change is that the IRFs of mobility and economic news sentiment cannot be

The systematic component of the health policy rule Beyond the analysis of IRFs, SVARs provide useful information regarding the systematic component of the government-mandates policy implied by our identification scheme.[22] More specifically, the contemporaneous coefficients correspond to ratios of entries in the vector $\boldsymbol{a}_{0,1}$. Thus, abstracting from lags and the constant term, the health policy equation can be written as:

$$\text{str}_t = \psi_R \text{re}_t + \psi_M \text{m}_t + \psi_{-\Delta S} \text{nc}_t + \psi_{D_H} \text{ndh}_t + \psi_{D_P} \text{ndp}_t + \psi_{\text{ns}} \text{ns}_t + \sigma \varepsilon_{1,t},$$

where str denotes the stringency index in log-percent, re_t denotes $\mathcal{R}_{e,t}$ expressed in log-percent, m_t denotes mobility in p.p., nc_t denotes the log of new cases in log-percent, ndh_t denotes the number of new deaths in hospitals, ndp_t denotes the number of new deaths at home, and ns_t denotes the ENS indicator. Accordingly, the coefficient $\psi_R = -\boldsymbol{a}_{0,11}/\boldsymbol{a}_{0,31}$ denotes the contemporaneous response of the stringency index to $\mathcal{R}_{e,t}$, $\psi_M = -\boldsymbol{a}_{0,21}/\boldsymbol{a}_{0,31}$ denotes the contemporaneous response of the stringency index to mobility, $\psi_{-\Delta S} = -\boldsymbol{a}_{0,41}/\boldsymbol{a}_{0,31}$ denotes the contemporaneous response of the stringency index to new cases, $\psi_{D_H} = -\boldsymbol{a}_{0,51}/\boldsymbol{a}_{0,31}$ denotes the contemporaneous response of the stringency index to new deaths in hospitals, $\psi_{D_P} = -\boldsymbol{a}_{0,61}/\boldsymbol{a}_{0,31}$ denotes the contemporaneous response of the stringency index to new deaths at home, $\psi_{\text{ns}} = -\boldsymbol{a}_{0,71}/\boldsymbol{a}_{0,31}$ denotes the contemporaneous response of the stringency index to economic conditions as measured by the ENS indicator, and $\sigma = 1/\boldsymbol{a}_{0,31}$ is the standard deviation of the health policy shock.

Table 3: Contemporaneous Coefficients in the Health Policy Equation

Coefficient	ψ_R	ψ_M	$\psi_{-\Delta S}$	ψ_{D_H}	ψ_{D_P}	ψ_{ns}
Median	0.85	2.44	0.67	8.94	5.50	20.36
68% Prob. Int.	[0.23;2.46]	[0.62;7.67]	[0.19;1.73]	[2.75;22.14]	[1.44;14.91]	[5.58;56.66]
90% Prob. Int.	[0.06;4.85]	[0.21;16.53]	[0.06;2.96]	[0.97;39.71]	[0.46;31.61]	[1.79;114.12]

Note: The table's entries denote the posterior median estimates of the contemporaneous coefficients in the health policy equation under our identification. The 68% and 90% equal-tailed posterior probability intervals are reported in brackets. The table is based on 10,000 independent draws.

Table 3 reports the posterior distribution of the contemporaneous coefficients in the health policy equation. The posterior median of ψ_R equals 0.85, which implies that the stringency index increases by about 8.3% in response to a 10% increase in the reproduction number. As a reference point, recall that a 7.4% rise in the index is equivalent to a one-unit increase in one of the ordinal measures that composed the OS index, on average. The posterior median of ψ_M

distinguished from zero during the first two weeks following the shock. We thank an anonymous referee for suggesting this check.

[22] This is a common practice when identifying monetary or fiscal policy equations in SVARs: e.g., Leeper and Zha (2003), Sims and Zha (2006), and Caldara and Kamps (2017).

equals 2.44, indicating that the stringency index increases by 12.2% (where 12.2=2.44 × 5) in response to a 5 percentage point increase in daily mobility. The posterior median of $\psi_{-\Delta S}$ equals 0.67, that is, the stringency index increases by about 7% in response to a 10% rise in new daily cases. The posterior median of ψ_{D_H}, 8.94, suggests that the stringency index increases by about 89% in response to 10 new daily deaths in hospitals. The posterior median of ψ_{D_P}, 5.5, means that the stringency index increases by about 55%. in response to 10 new daily deaths at home. The posterior median of ψ_{ns}, 20.36, means that the stringency index decreases by about 20% in response to a one-standard-deviation decrease in the ENS indicator. In terms of output, this means stringency decreases 20% (i.e., about a one-unit increase in three of the ordinal measures that compose the OS index) in response to a drop of around 1.2% in next quarter real GDP. Overall, the coefficients are plausible given the behavior of health authorities in 2020 across the advanced economies. The 68% and 90% probability intervals are wide, but one could easily extend our analysis to impose bounds on coefficients based on external evidence or judgment; see Arias et al. (2019).

5.2 A Reproduction Shock

We next identify a reproduction shock and analyze whether its effects on the virus's spread, deaths, and economic activity depend on the level of government stringency. To answer this state-dependency question we use the LP approach proposed by Jordà (2005) and developed by Ramey and Zubairy (2018), Stock and Watson (2018), and Plagborg-Møller and Wolf (2021), among others. In particular, we use LPs with interaction terms as in Ramey and Zubairy (2018) to study how a reproduction shock propagates depending on the level of government stringency. Our identification scheme consists of sign and zero restrictions implemented as described in Plagborg-Møller and Wolf (2021). Consider the following LP specification:

$$\begin{aligned} \boldsymbol{w}_{i,t+h} = I_{t-1} & \left(\boldsymbol{\mu}_{\mathbb{H},i,h} + \boldsymbol{\beta}'_{\mathbb{H},i,h} \boldsymbol{w}_t + \sum_{\ell=1}^{\nu} \boldsymbol{\delta}_{\mathbb{H},i,h,\ell} \boldsymbol{w}_{t-\ell} + \boldsymbol{\gamma}'_{\mathbb{H},i,h} \boldsymbol{s}_t \right) \\ + (1 - I_{t-1}) & \left(\boldsymbol{\mu}_{\mathbb{L},i,h} + \boldsymbol{\beta}'_{\mathbb{L},i,h} \boldsymbol{w}_t + \sum_{\ell=1}^{\nu} \boldsymbol{\delta}_{\mathbb{L},i,h,\ell} \boldsymbol{w}_{t-\ell} + \boldsymbol{\gamma}'_{\mathbb{L},i,h} \boldsymbol{s}_t \right) + \boldsymbol{\xi}_{i,h,t} \end{aligned} \tag{3}$$

for $i = 1, \ldots, n$, and $h = 0, \ldots, H$, where $\boldsymbol{w}_t$ is the $w \times 1$ vector of endogenous variables equal to $\boldsymbol{y}_t$ defined in Section 5.1 save for the OS index, $\boldsymbol{w}_{i,t+h}$ denotes the value of the i-th variable in $\boldsymbol{w}_{t+h}$, and $\boldsymbol{s}_t$ is an $s \times 1$ vector of exogenous variables.

The exogenous variables include $\boldsymbol{z}_t$ (defined in Section 5.1), the point-wise median smoothed death probability in hospitals, and the point-wise median smoothed death probability at home. I_{t-1} is a dummy variable that indicates whether Belgium is in a high government stringency regime. The high government stringency regime is determined based on whether the OS index is

above its median level over the sample (58.33).

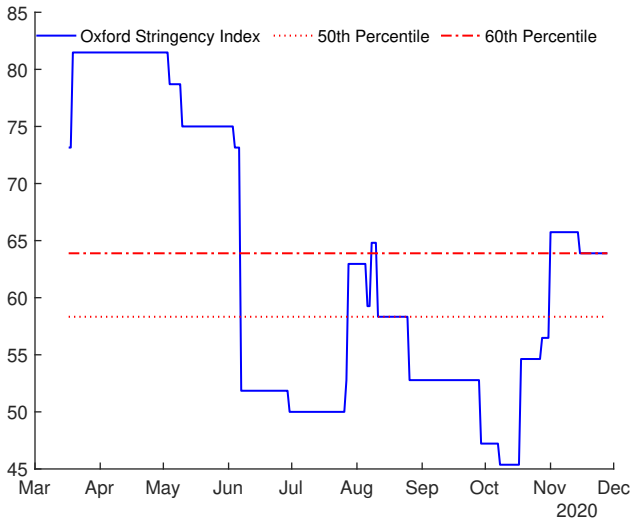


Figure 7: Stringency Index

Figure 7 shows the time series of stringency plus its 50th and 60th percentiles. A value of about 50 for the OS index corresponds to a situation where schools are open, working from home is recommended (albeit not required), there are restrictions on gatherings of more than 100 people, traveling between regions is advised to be avoided (but not prohibited), and a quarantine is required for arrivals from some or all regions. In contrast, a value of about 80 for the OS index corresponds to a situation where schools are required to close at some levels, working from home is required for some sectors, there are restrictions on gatherings of more than 10 people, restrictions on internal movement are in place, and there is a total border closure. Hence, the median threshold corresponds to a serious but not extreme lockdown.[23]

The parameters $\boldsymbol{\mu}_{\mathbb{H},i,h}$, $\boldsymbol{\beta}'_{\mathbb{H},i,h}$, $\boldsymbol{\delta}_{\mathbb{H},i,h,\ell}$, and $\boldsymbol{\gamma}'_{\mathbb{H},i,h}$ correspond to the high government stringency regime, i.e., the stringency index is above its sample median, and the parameters $\boldsymbol{\mu}_{\mathbb{L},i,h}$, $\boldsymbol{\beta}'_{\mathbb{L},i,h}$, $\boldsymbol{\delta}_{\mathbb{L},i,h,\ell}$, and $\boldsymbol{\gamma}'_{\mathbb{L},i,h}$ correspond to the low government stringency regime. The innovation term for $h = 1$, $\boldsymbol{\xi}_{1,t} = (\boldsymbol{\xi}_{1,1,t}, \ldots, \boldsymbol{\xi}_{w,1,t})'$, is assumed to be mean zero with covariance matrix $\mathbb{E}_t(\boldsymbol{\xi}_{1,t}\boldsymbol{\xi}'_{1,t}) = \boldsymbol{\Sigma}$.[24]

All told, the LP specification is similar to the SVAR specification used above with a few modifications tailored to the question at hand. First, we do not include the stringency index because it is the variable we use to split the sample. Second, we use 3 lags to reduce parameter uncertainty. Third, we add death probabilities as exogenous variables to control for any effect related to the sample split. In particular, it could be the case that high (low) government

[23] As a robustness exercise, Appendix A.7 considers a threshold at the 60th percentile. The main results are robust. We did not consider thresholds higher than the 60th percentile because, as shown by Figure 7, they would imply a very sharp division between the first three months of the pandemic and the rest of our sample.

[24] We have also experimented with the case in which the covariance matrix is regime specific. The main conclusions of this section remained unchanged.

stringency episodes are correlated with high (low) mortality rates. In fact, the death probability was higher at the beginning of our sample as better treatments and refined clinical protocols for the disease were not ready yet. The restrictions to identify the reproduction shock are:

Restriction 2. *The reproduction number increases for at least three days in response to a positive reproduction shock. Mobility decreases for at least three days in response to a positive reproduction shock. The impact response of mobility in p.p. is bounded to be smaller than the p.p. increase in the reproduction number. In addition, the reproduction shock does not affect deaths, new cases, and the ENS indicator contemporaneously.*

Restriction 2 identifies what we call a reproduction shock, that is, an exogenous variation in the transmission rate of COVID-19. As we explained above, such an exogenous variation could occur when people relax their compliance with social distancing measures or when a more contagious variant of the virus emerges. The positive sign restrictions on the reproduction number's impact response is a normalization; the positive sign restrictions on the subsequent days are imposed to sharpen identification. We impose just three days to be cautious and let the data dictate the response's shape as much as possible while keeping identification. The negative sign restriction on mobility is imposed based on the notion that, on average, people will stay at home in response to an unexpected increase in reproduction numbers.

The elasticity bound is imposed to discipline the identified set of mobility. In the absence of such a bound, the identified set would include a decline in mobility of 100 p.p. as being equally likely as no decline in mobility following an unexpected 10% increase in $\mathcal{R}_{e,t}$. Such a result is implausible. Hence, we use a bound to rule out dubious IRFs as in Kilian and Murphy (2012) and Arias et al. (2019). The zero restrictions on deaths and new cases are predicated on the CDC's reports, since it takes more than one day for symptoms to develop, and on evidence showing that it takes more than one day to die from COVID-19. This is just a matter of exposure taking time to result in a measured case and death. For example, Wortham et al. (2020) report a median clinical course of the disease of 10 days with an interquartile range of 6 to 15 days. The zero restriction on the impact response of the ENS indicator is that it takes at least one day to have broad coverage of the mobility restrictions following the shock and for the audience to process it.

Given the sign and zero restrictions described above, we compute the identified set by numerically solving the quadratic program described in the supplement to Plagborg-Møller and Wolf (2021) using Algorithm 2 of Giacomini and Kitagawa (2018). In particular, let $\boldsymbol{S}_1$ be a 7xw matrix that selects the IRFs that we restrict to be either positive or negative, and let $\boldsymbol{Z}_1$ be a 4xn matrix that selects the IRFs that we restrict to zero. Then, for each regime r $\in \{H, L\}$, we draw 1 million orthogonal matrices $\boldsymbol{Q}_r$ that satisfy:

$$\begin{bmatrix} \boldsymbol{S}_1 \hat{\boldsymbol{C}}_{0:2,r} \hat{\boldsymbol{B}} \mathbf{Q}_r \boldsymbol{e}_1 \\ \left(\boldsymbol{e}_2' \hat{\boldsymbol{C}}_{0,r} \hat{\boldsymbol{B}} \mathbf{Q}_r \boldsymbol{e}_1 / \boldsymbol{e}_1' \hat{\boldsymbol{C}}_{0,r} \hat{\boldsymbol{B}} \mathbf{Q}_r \boldsymbol{e}_1 \right) - 1 \end{bmatrix} \geq \mathbf{0} \text{ and } \boldsymbol{Z}_1 \hat{\boldsymbol{C}}_{0,r} \hat{\boldsymbol{B}} \mathbf{Q}_r \boldsymbol{e}_1 = \mathbf{0} \tag{4}$$

where $\hat{\boldsymbol{C}}_{0:2,r} = \left[\hat{\boldsymbol{C}}'_{0,r}, \hat{\boldsymbol{C}}'_{1,r}, \hat{\boldsymbol{C}}'_{2,r}\right]'$, $\hat{\boldsymbol{C}}_{h,r} = \left(\hat{\boldsymbol{\beta}}_{r,1,h}, \ldots, \hat{\boldsymbol{\beta}}_{r,w,h}\right)'$ $h \geq 0$, $\hat{\boldsymbol{\beta}}_{r,j,h}$ denotes the OLS estimate of $\boldsymbol{\beta}_{r,j,h}$ for $j = 1, \ldots, w$ and $r \in \{\mathbb{L}, \mathbb{H}\}$, $\hat{\boldsymbol{B}} = \text{chol}(\hat{\boldsymbol{\Sigma}})'$, chol is the upper triangular Cholesky decomposition of $\hat{\boldsymbol{\Sigma}}$, $\hat{\boldsymbol{\Sigma}}$ is the OLS estimate of $\boldsymbol{\Sigma}$, $\boldsymbol{e}_2'\hat{\boldsymbol{C}}_{0,r}\hat{\boldsymbol{B}}\boldsymbol{Q}_r\boldsymbol{e}_1 / \boldsymbol{e}_1'\hat{\boldsymbol{C}}_{0,r}\hat{\boldsymbol{B}}\boldsymbol{Q}_r\boldsymbol{e}_1$ denotes the ratio between the impact IRF of mobility and $\mathcal{R}_{e,t}$, and $\boldsymbol{e}_i$ denotes the i-th column of the identity matrix. Given $\hat{\boldsymbol{B}}$ and $\hat{\boldsymbol{C}}_{h,r}$, we draw $\boldsymbol{Q}_r$ K times and let $\{\boldsymbol{Q}_{r,k} : k = 1, \ldots, K\}$ be the draws that satisfy the sign and zero restrictions in Equation (4). Then, letting $\boldsymbol{q}_{r,k,1}$ denote the first column of $\boldsymbol{Q}_{r,k}$ the identified set for variable i at horizon h is given by:

$$\left[\min_k \boldsymbol{e}_i'\hat{\boldsymbol{C}}_{h,r}\hat{\boldsymbol{B}}\boldsymbol{q}_{r,k,1}\frac{10}{\boldsymbol{e}_1'\hat{\boldsymbol{C}}_{0,r}\hat{\boldsymbol{B}}\boldsymbol{q}_{r,k,1}}, \max_k \boldsymbol{e}_i'\hat{\boldsymbol{C}}_{h,r}\hat{\boldsymbol{B}}\boldsymbol{q}_{r,k,1}\frac{10}{\boldsymbol{e}_1'\hat{\boldsymbol{C}}_{0,r}\hat{\boldsymbol{B}}\boldsymbol{q}_{r,k,1}}\right].$$

The factor $\frac{10}{\boldsymbol{e}_1'\hat{\boldsymbol{C}}_{0,r}\hat{\boldsymbol{B}}\boldsymbol{q}_{r,k,1}}$ is a normalization imposed so that the reproduction shock increases $\mathcal{R}_{e,t}$ by 10% in both regimes.

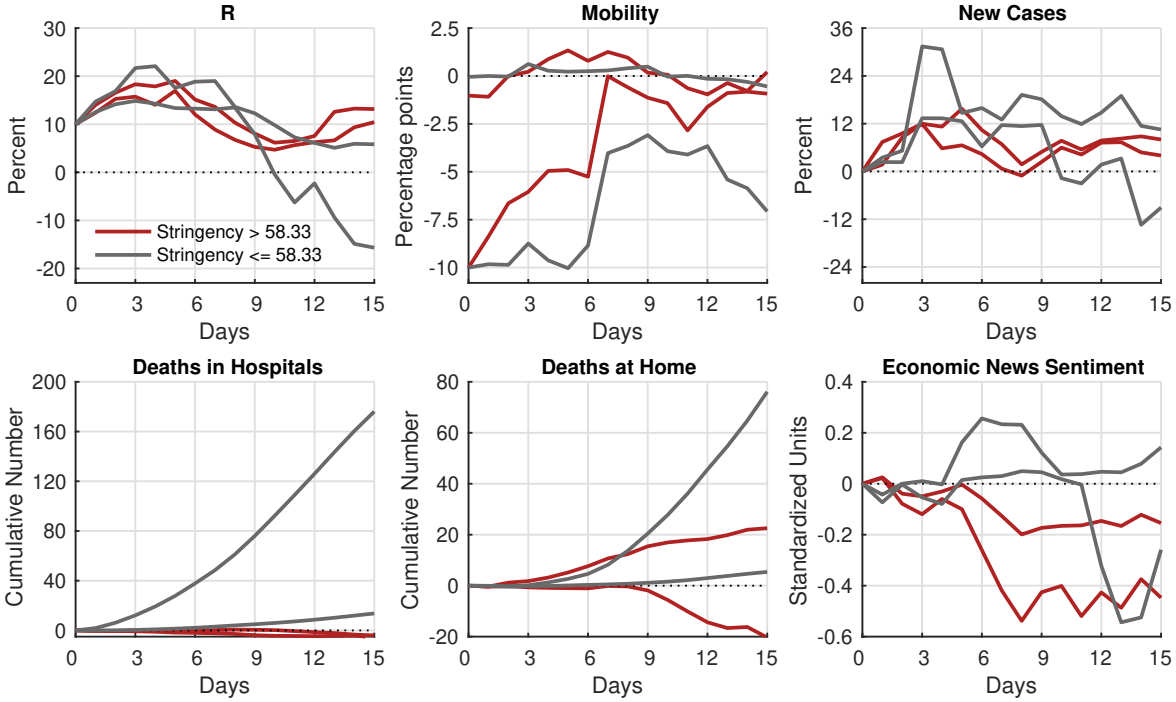


Figure 8: IRFs to a reproduction shock

Figure 8 reports the identified set following a reproduction shock. We show the IRFs from horizon 0 up to horizon 15.[25] The shorter horizon relative to the horizon of the IRFs shown in the SVAR is due to the length of our sample and the parameter uncertainty associated with LPs. Notice, in particular, that the LP looks at the variation in the systematic regime, while the SVAR focuses on the effect of a shock within the regime. Figure 8 truncates the horizon at 15

[25] One could follow Plagborg-Møller and Wolf (2020) and implement a forecast error variance decomposition (FEVD) to assess how much in-sample variation in the various epidemiological variables can be accounted for by the identified reproduction shocks.

days, which encompasses the illness duration of patients who died from COVID-19 reported in Wortham et al. (2020), who find a median duration of 10 days and an interquartile range of 6 to 15 days, which is a tad below the findings of Sousa et al. (2020), who report a median illness duration of 19 days and an interquartile range of 12 to 23 days.

In response to a reproduction shock that increases $\mathcal{R}_{e,t}$ by 10%, the mobility index in the low and high government stringency regimes decreases by a similar amount, new cases increase by less in the high government stringency regime, leading to fewer deaths (both in hospitals and at home). Another factor explaining the higher level of new cases and deaths in the low government stringency regime is that the increase in $\mathcal{R}_{e,t}$ is much more persistent in such a regime. Thus, high government stringency could save up to about 250 deaths in the first two weeks after the reproduction shock or around 1.5% of the deaths in our sample. Let us now assess the effects on economic activity of the shock to the reproduction number under the high and low stringency regimes. In response to the shock, the ENS indicator decreases by about 0.2 to 0.4 standardized units after 7 days in the high stringency regime and increases by about 0 to 0.2 standardized units in the low stringency regime.

As in Section 5.1, it is useful to map the standardized units to real per capita GDP terms. After a reproduction shock, a high stringency regime saves 250 lives at a cost of between €2 and €4 per capita (or between €91,872 and €183,745 per life saved). The cost of preventing a death is higher in the case of LPs, but notice that the reference horizon of the IRFs in the case of LPs is 7 days, while in the case of the SVARs it is 90 days, when there is more room for a recovery of economic activity. In any case, the economic cost of a high stringency regime is negligible.

Finally, while it is reasonable to assume that a reproduction shock should immediately feed through to $\mathcal{R}_{e,t}$, it is less clear that the behavioral feedback to mobility should be similarly fast. For this reason, we test the robustness of the results in Figure 8. We consider two alternatives to Restriction 2. The first alternative assumes that mobility decreases 3 days after the shock for at least 3 days in response to a positive reproduction shock. The second alternative assumes that mobility decreases 7 days after the shock for at least 3 days in response to a positive reproduction shock. In both cases, the results are unchanged.[26]

6 Conclusion

There are many environments in which to generalize our results. First, we could integrate our epidemiological model within a more economic model and estimate how individuals make decisions regarding mobility instead of imposing a random walk variation. Second, we could introduce a rich network structure between different compartments (for example, reflecting heterogeneity by age, gender, occupation, and location) and estimate how the parameters

[26]We thank an anonymous referee for suggesting this exercise.

governing the movements among those compartments evolve (do effective contact rates between regions drop more persistently than effective contact rates within regions?). Examples of these richer data structures appear in Acemoglu et al. (2020) and Aguirregabiria et al. (2020). These could be particularly helpful in linking health outcomes with economic outcomes across different population groups provided that one has access to such data. Third, we could consider how the introduction of vaccines and new medical treatments affects health and macroeconomic outcomes.

References

Acemoglu, D., V. Chernozhukov, I. Werning, and M. D. Whinston (2020). A Multi-Risk SIR Model with Optimally Targeted Lockdown. Technical report, MIT.

Aguirregabiria, V., J. Gu, Y. Luo, and P. Mira (2020, May). A Dynamic Structural Model of Virus Diffusion and Network Production: A First Report. Technical Report 14750, C.E.P.R. Discussion Papers.

Algaba, A., S. Borms, K. Boudt, and B. Verbeken (2021). Daily News Sentiment and Monthly Surveys: A Mixed–Frequency Dynamic Factor Model for Nowcasting Consumer Confidence. *Available at SSRN 3609297*.

Andersen, M. (2020). Early Evidence on Social Distancing in Response to COVID-19 in the United States. *Available at SSRN 3569368*.

Arevalo-Rodriguez, I., D. Buitrago-Garcia, D. Simancas-Racines, P. Zambrano-Achig, R. del Campo, A. Ciapponi, et al. (2020). False-Negative Results of Initial RT-PCR Assays for COVID-19: A Systematic Review. *medRxiv*.

Arias, J. E., D. Caldara, and J. F. Rubio-Ramírez (2019). The Systematic Component of Monetary Policy in SVARs: An Agnostic Identification Procedure. *Journal of Monetary Economics 101*, 1–13.

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2018). Inference Based on Structural Vector Autoregressions Identified with Sign and Zero Restrictions: Theory and Applications. *Econometrica 86*(2), 685–720.

Arnon, A., J. Ricco, and K. Smetters (2020). Epidemiological and economic effects of lockdown. *Brookings Papers on Economic Activity Fall 2020, Special Edition*, 61–108.

Bar-On, Y. M., A. Flamholz, R. Phillips, and R. Milo (2020). Science Forum: SARS-CoV-2 (COVID-19) by the Numbers. *eLife 9*, e57309.

Caldara, D. and C. Kamps (2017). The Analytics of SVARs: A Unified Framework to Measure Fiscal Multipliers. *The Review of Economic Studies 84*(3), 1015–1040.

Catteau, L., N. Dauby, M. Montourcy, B. E., J. Hautekiet, E. Goetghebeur, et al. (2020). Low-dose Hydroxychloroquine Therapy and Mortality in Hospitalised Patients with COVID-19: A Nationwide Observational Study of 8075 Participants. *International Journal of Antimicrobial Agents 56*(4), 106144.

Chernozhukov, V., H. Kasahara, and P. Schrimpf (2021). Causal Impact of Masks, Policies, Behavior on Early COVID-19 Pandemic in the U.S. *Journal of Econometrics 220*(1), 23–62.

Cho, S.-W. S. (2020). Quantifying the Impact of Non-Pharmaceutical Interventions during the COVID-19 Outbreak – The Case of Sweden. *The Econometrics Journal 23*(3), 323–344.

D'Arienzo, M. and A. Coniglio (2020). Assessment of the SARS-CoV-2 Basic Reproduction Number, R0, Based on the Early Phase of COVID-19 Outbreak in Italy. *Biosafety and Health 2*(2), 57 – 59.

Doucet, A. and A. M. Johansen (2009). A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later. *Handbook of Nonlinear Filtering 12*(656-704), 3.

Fernández-Villaverde, J. and J. Rubio-Ramírez (2007). How Structural Are Structural Parameters? *NBER Macroeconomics Annual 22*, 83–137.

Giacomini, R. and T. Kitagawa (2018, November). Robust Bayesian Inference for Set-identified Models. (CWP61/18).

Gilchrist, S. and E. Zakrajšek (2012, June). Credit spreads and business cycle fluctuations. *American Economic Review 102*(4), 1692–1720.

Goolsbee, A. and C. Syverson (2020, June). Fear, Lockdown, and Diversion: Comparing Drivers of Pandemic Economic Decline 2020. Working Paper 27432, National Bureau of Economic Research.

Gostic, K. M., L. McGough, E. B. Baskerville, S. Abbott, K. Joshi, C. Tedijanto, et al. (2020). Practical Considerations for Measuring the Effective Reproductive Number, R_t. *PLOS Computational Biology 16*(12).

Gupta, S., T. D. Nguyen, F. L. Rojas, S. Raman, B. Lee, A. Bento, K. I. Simon, and C. Wing (2020). Tracking public and private responses to the covid-19 epidemic: Evidence from state and local government actions. Working Paper 27027, National Bureau of Economic Research.

Hale, T., S. Webster, A. Petherick, T. Phillips, and B. Kira (2020). Oxford COVID-19 Government Response Tracker. Available at https://covidtracker.bsg.ox.ac.uk/.

Herzog, S., J. De Bie, S. Abrams, I. Wouters, E. Ekinci, L. Patteet, A. Coppens, S. De Spiegeleer, P. Beutels, P. Van Damme, N. Hens, and H. Theeten (2020). Seroprevalence of Igg Antibodies against SARS Coronavirus 2 in Belgium – A Serial Prospective Cross-Sectional Nationwide Study of Residual Samples. medRxiv.

Ho, P., T. A. Lubik, and C. Matthes (2021). How to Go Viral: A COVID-19 Model with Endogenously Time-Varying Parameters. *Journal of Econometrics*.

Hurwicz, L. (1966). On the Structural Form of Interdependent Systems. In *Studies in Logic and the Foundations of Mathematics*, Volume 44, pp. 232–239. Elsevier.

Jordà, Ò. (2005). Estimation and Inference of Impulse Responses by Local Projections. *American Economic Review 95*(1), 161–182.

Katul, G. G., A. Mrad, S. Bonetti, G. Manoli, and A. J. Parolari (2020). Global Convergence of COVID-19 Basic Reproduction Number and Estimation from Early-time SIR Dynamics. *PLOS One 16*(4), 351–367.

Kilian, L. and D. P. Murphy (2012). Why Agnostic Sign Restrictions Are Not Enough: Understanding the Dynamics of Oil Market VAR Models. *Journal of the European Economic Association 10*(5), 1166–1188.

Kucirka, L. M., S. A. Lauer, O. Laeyendecker, D. Boon, and J. Lessler (2020). Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since Exposure. *Annals of Internal Medicine*.

Leeper, E. M., C. A. Sims, and T. Zha (1996). What Does Monetary Policy Do? *Brookings Papers on Economic Activity 1996*(2), 1–78.

Leeper, E. M. and T. Zha (2003). Modest Policy Interventions. *Journal of Monetary Economics 50*(8), 1673–1700.

Lewis, D. J., K. Mertens, J. H. Stock, and M. Trivedi (2020). Measuring Real Activity Using a Weekly Economic Index. *Federal Reserve Bank of New York Staff Reports* (920).

Maloney, W. and T. Taskin (2020a). *Determinants of Social Distancing and Economic Activity during COVID-19: A Global View*. The World Bank.

Maloney, W. and T. Taskin (2020b). Social Distancing and Economic Activity during COVID-19: A Global View. *COVID Economics Issue 13*(4).

Manski, C. F. and F. Molinari (2021). Estimating the COVID-19 Infection Rate: Anatomy of an Inference Problem. *Journal of Econometrics 220*(1), 181–192.

Molenberghs, G., C. Faes, J. Aerts, H. Theeten, B. Devleesschauwer, N. Sierra, et al. (2020). Belgian COVID-19 Mortality, Excess Deaths, Number of Deaths per Million, and Infection Fatality Rates (8 March - 9 May 2020). *medRxiv*.

OECD (2012). *Mortality Risk Valuation in Environment, Health and Transport Policies*. OECD Publishing.

O'Neill, P. D. and G. O. Roberts (1999). Bayesian Inference for Partially Observed Stochastic Epidemics. *Journal of the Royal Statistical Society. Series A (Statistics in Society) 162*(1), 121–129.

Plagborg-Møller, M. and C. K. Wolf (2020). Instrumental Variable Identification of Dynamic Variance Decompositions. *arXiv preprint arXiv:2011.01380*.

Plagborg-Møller, M. and C. K. Wolf (2021). Local Projections and VARs Estimate the Same Impulse Responses. *Econometrica (Forthcoming)*.

Ramey, V. A. and S. Zubairy (2018). Government Spending Multipliers in Good Times and in Bad: Evidence from US Historical Data. *Journal of Political Economy 126*(2), 850–901.

Roberts, G. O. and J. S. Rosenthal (2001). Optimal Scaling for Various Metropolis-Hastings Algorithms. *Statistical Science 16*(4), 351–367.

Sierra, N. B., N. Bossuyt, T. Braeye, M. Leroy, I. Moyersoen, I. Peeters, et al. (2020). All-Cause Mortality Supports the COVID-19 Mortality in Belgium and Comparison with Major Fatal Events of the Last Century. *Archives of Public Health 78*(1), 1–8.

Sims, C. A. and T. Zha (2006). Were There Regime Switches in US Monetary Policy? *American Economic Review*, 54–81.

Sousa, G., T. Garces, V. Cestari, R. Florêncio, T. Moreira, and M. Pereira (2020). Mortality and Survival of COVID-19. *Epidemiology & Infection 148*.

Stock, J. H. and M. W. Watson (2018). Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments. *The Economic Journal 128*(610), 917–948.

Toulis, P. (2021). Estimation of Covid-19 Prevalence from Serology Tests: A Partial Identification Approach. *Journal of Econometrics 220*(1), 193–213.

Wolf, C. K. (2020, October). SVAR (Mis)identification and the Real Effects of Monetary Policy Shocks. *American Economic Journal: Macroeconomics 12*(4), 1–32.

Wortham, J., J. Lee, S. Althomsons, et al. (2020). Characteristics of Persons Who Died with COVID-19 — United States. *Morbidity and Mortality Weekly Report 69(45)*, 923–929.

A Online Appendix

A.1 Implied Density over the Initial Share of Detected Cases

Figure A.1 plots the prior and posterior distributions of the initial share of detected cases, $\gamma_{n,0}$, implied by our model. While the prior assigns substantial probability mass to values as high as 0.6, the posterior distribution is concentrated in values below 0.4, indicating that, according to our model, the share of detected cases on March 15, 2020, was below 40% with high probability.

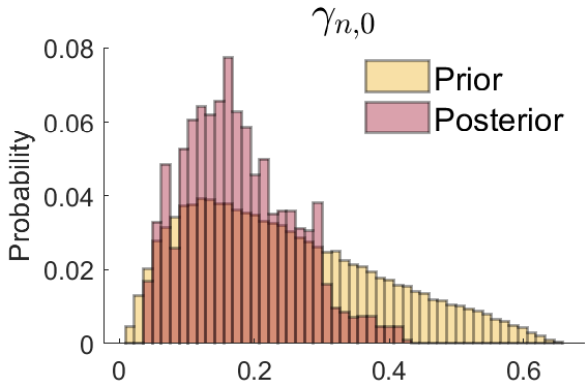


Figure A.1: Prior histograms (yellow) are based on 100,000 independent draws from the prior distribution presented in Table 2. Posterior histograms (red) are based on the MCMC chain with 90,000 posterior draws obtained after a burn-in period of 10,000 draws.

A.2 Prior versus Posterior

Figure A.2 shows the prior and posterior distributions for γ, θ_H, and θ_P. These parameters are inversely related to the average duration a person remains infectious, the average duration of stay in hospitals, and the average duration of stay at home while recovering from COVID-19. Figure A.2 reveals that the data are very informative about γ and θ_H and less informative about θ_P. Figure A.3 shows the prior and posterior distributions for σ_b, σ_h, σ_p, σ_g, and σ_n. These parameters govern the step size of the time-varying parameters of our model. Clearly, the data are informative about them. Finally, Figure A.4 shows the prior and posterior distributions for b_0, d_{H_0}, d_{P_0}, g_{H_0}, S_0, I_0, n_0, and μ. These parameters are the initial-value parameters and the share of false negatives parameter in the case of μ. Figure A.4 documents that data are informative about these 8 parameters as well. The prior for n_0 is truncated at 1 to rule out large values of the permanent component of detected cases. Even so, the posterior indicates that there is not much probability mass near the truncation, suggesting that the upper bound for the prior could be relaxed without affecting our conclusions.

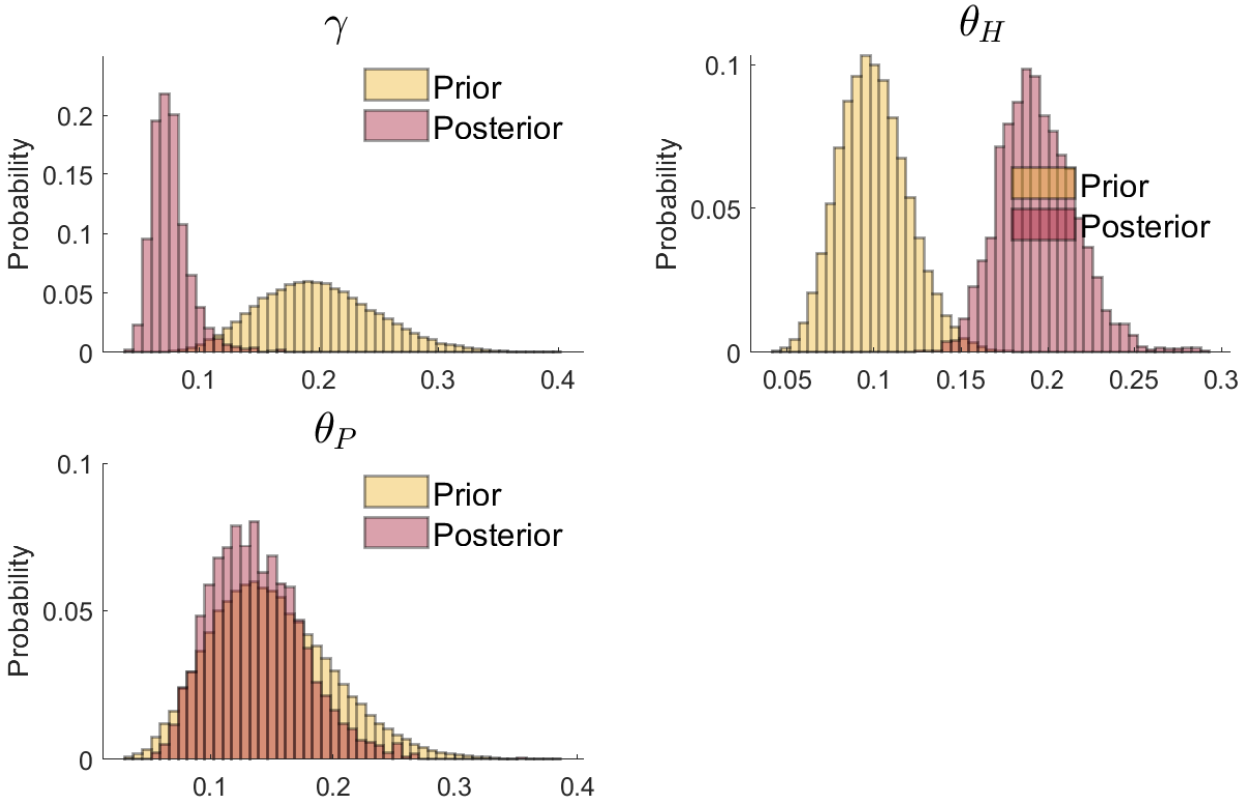


Figure A.2: Duration Parameters. Table 1 presents definitions of these parameters. Prior histograms (yellow) are based on 100,000 independent draws from the prior distribution presented in Table 2. Posterior histograms (red) are based on the MCMC chain with 90,000 posterior draws obtained after a burn-in period of 10,000 draws.

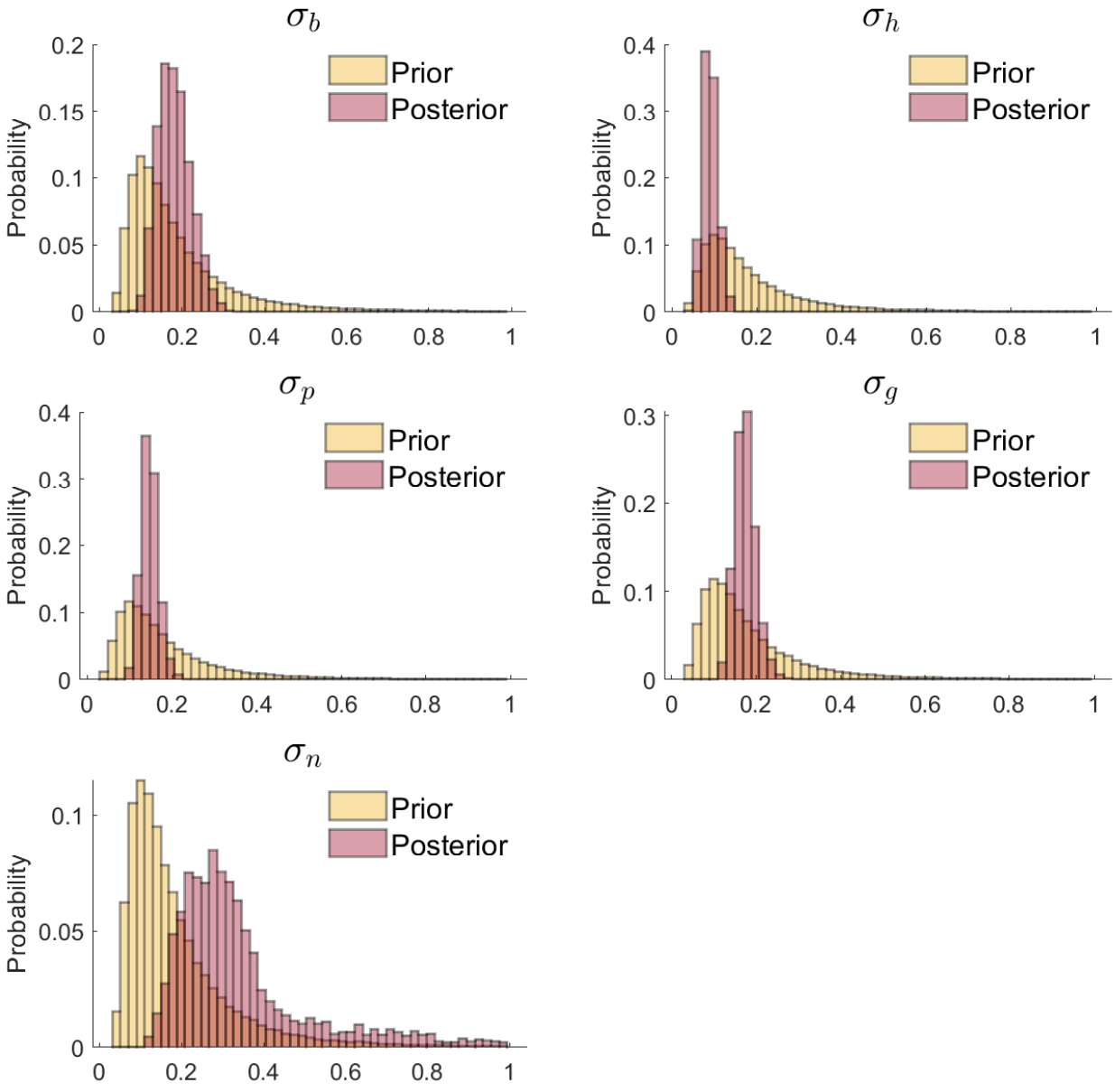


Figure A.3: Step-size Parameters. Table 1 presents definitions of these parameters. Prior histograms (yellow) are based on 100,000 independent draws from the prior distribution presented in Table 2. Posterior histograms (red) are based on the MCMC chain with 90,000 posterior draws obtained after a burn-in period of 10,000 draws.

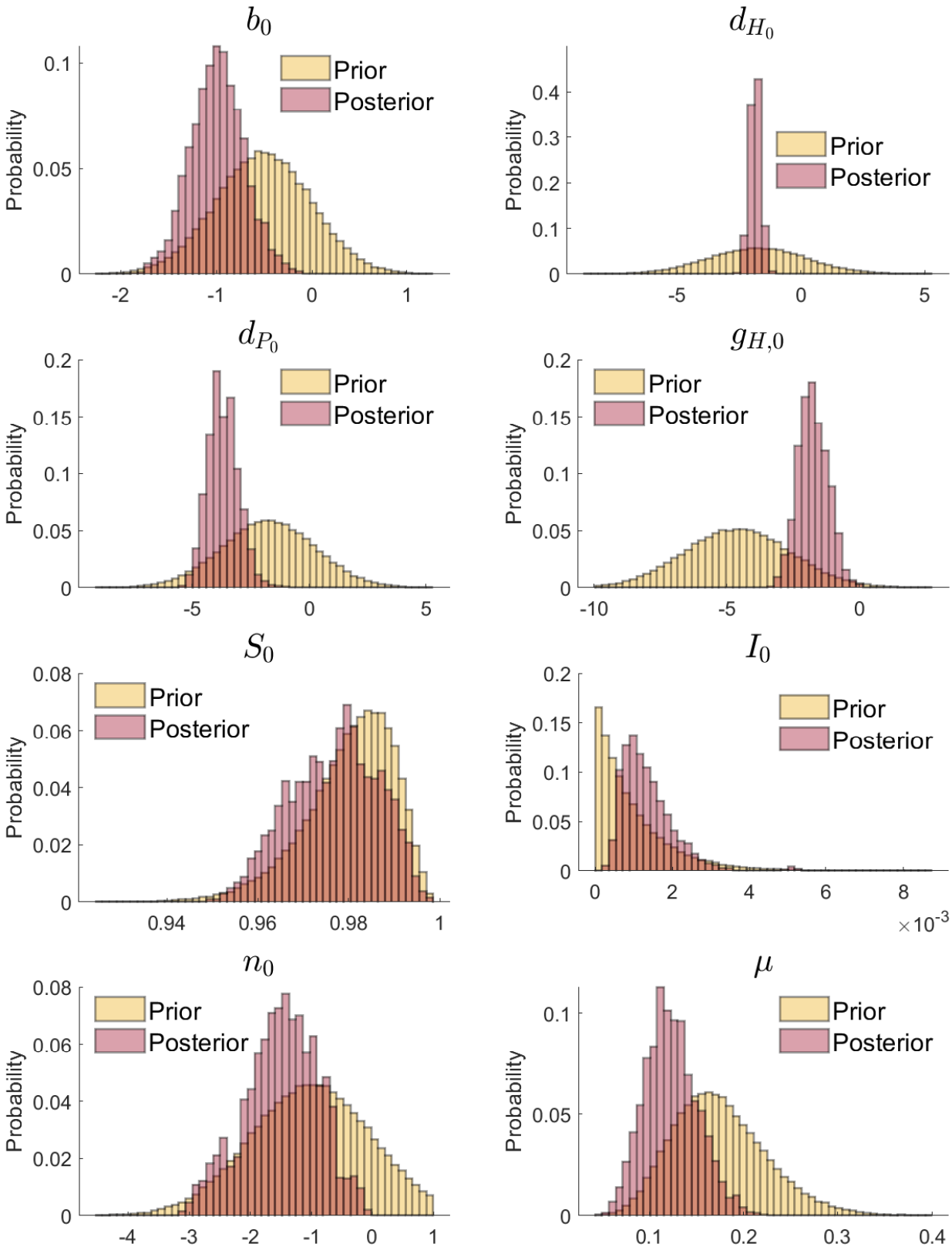


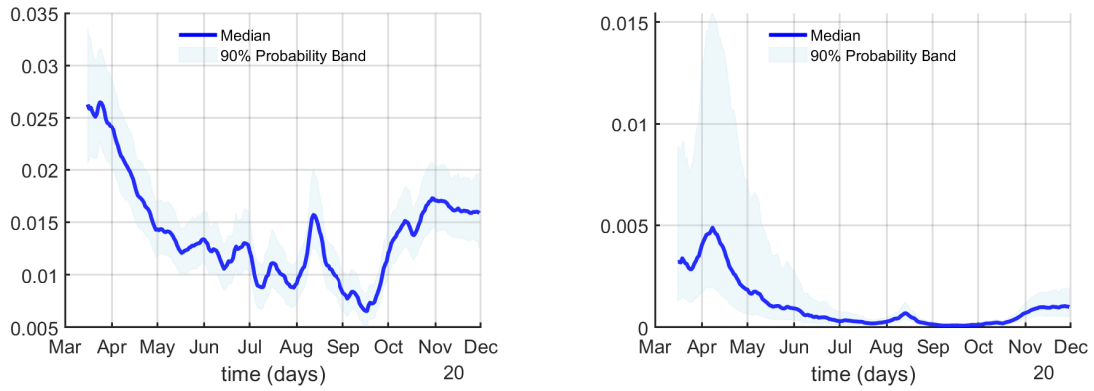
Figure A.4: Initial-value Parameters and Share of False Negatives Parameter. Table 1 presents definitions of those parameters. Prior histograms (yellow) are based on 100,000 independent draws from the prior distribution presented in Table 2. Posterior histograms (red) are based on the MCMC chain with 90,000 posterior draws obtained after a burn-in period of 10,000 draws.

A.3 Other States

Finally, we present the smoothed estimates of the time-varying death probabilities, the share of the population that is infectious, the share of the population that is recovering (in hospitals and at home), and the inflow of hospitalizations as a share of those that are no longer infectious.

Another state variable in our model is the time-varying death probabilities. Those probabilities can change for many reasons. We can enumerate a few. First, medical protocols vary. As health workers learn more about an infection, they can handle patients better, even in the absence of effective treatments. Second, hospitals experience different occupancy rates, with variations in the inflows and total capacity, as the supply of beds and ICU units responds to the crisis. Third, the demographics of patients can change, by varying either in terms of age or in terms of comorbidity levels.[27]

The left panel of Figure A.5 shows the in-hospital death probability (median smoothed, plus the 90% probability band), which went down from over 2.5% in March to less than 1% by early July. The sizeable third peak of COVID infections in the fall of 2020 increased that probability by only around 1.5%, suggesting a considerable degree of improvement in clinical outcomes.



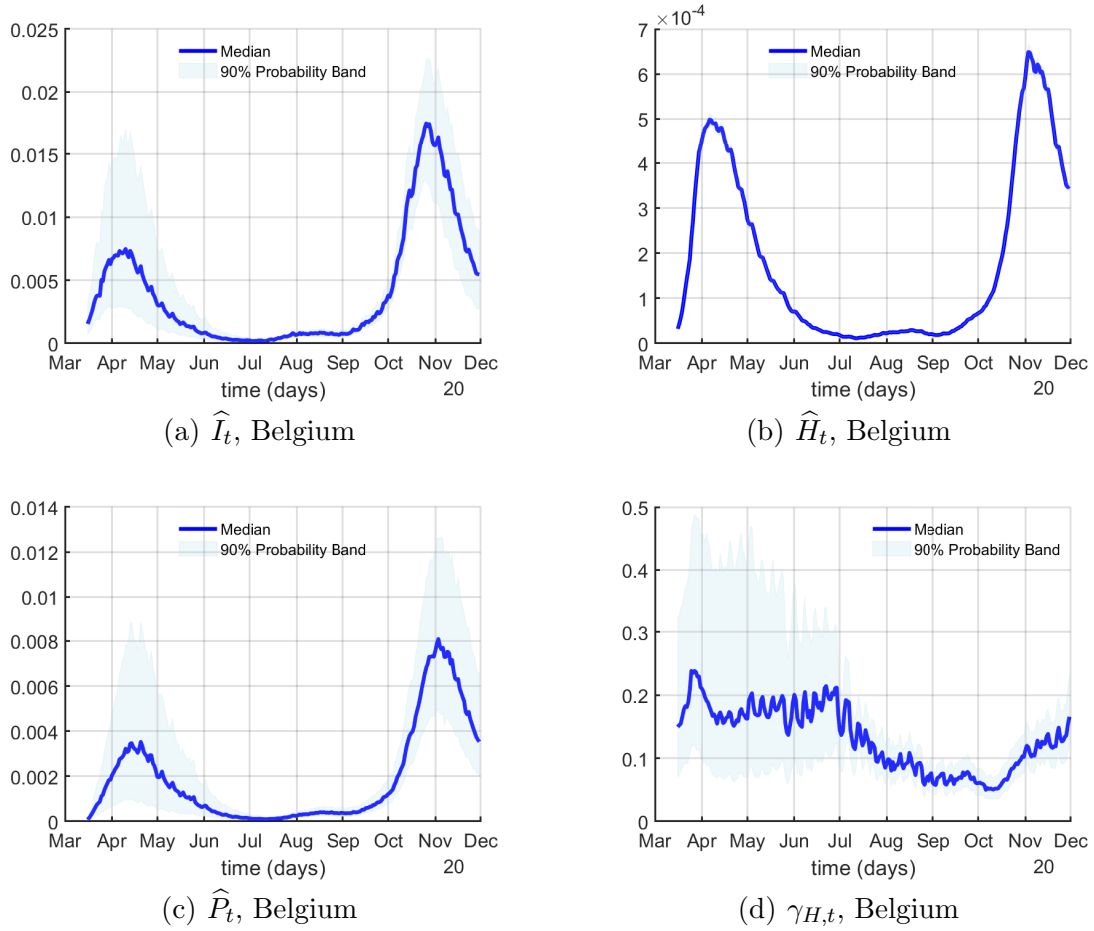
(a) In-hospital death probabilities, Belgium (b) At-home death probabilities, Belgium

Figure A.5: Death probabilities, Belgium

The right panel of Figure A.5 shows the at-home death probability, which fell from around 0.5% in March to less than 0.1% by early July. Here the changing conditions at retirement communities, which were unprepared for the virulence of COVID-19 in the late winter of 2020, are probably at the core of the estimated variation in death probabilities.

Figure A.6 shows the time series of these smoothed variables throughout our sample. The share of infectious increases to 1.8% by the end of October: more than twice as large as the

[27] Imagine, for example, that individuals with a high probability of infection (e.g., due to their social networks) and high fatality rate (e.g., smokers) got infected in the first wave. As there are fewer of these individuals in the population when the second wave arrives, the measured death rates will mechanically fall.



(a) $\widehat{I}_t$, Belgium

(b) $\widehat{H}_t$, Belgium

(c) $\widehat{P}_t$, Belgium

(d) $\gamma_{H,t}$, Belgium

Figure A.6: Other state variables, Belgium

March 15 estimate of 0.7%. This sharp increase in the share of infectious may provide a rationale for the first nationwide lockdown imposed by the Belgian government on March 18, 2020. These measures appear to have had an effect as the number of infectious dropped at the beginning of April and continued to decline, reaching a trough at the end of June 2020. After that, the share of infectious began to increase but at a moderate pace up until the first week of September, when we observe a second exponential increase in the share of infectious, leading to a reintroduction of lockdown measures on October 16, 2020.

The shares of the population recovering in hospitals and at home broadly track the contour of the share of infectious. For example, the first peak of infections and hospitalizations occurs in early April. A similar pattern emerges for the share of the population recovering outside of hospitals. Notice that while the peak of the third wave of infections and the peak of those recovering at home are more than twice as high as the peak of the first wave, during the second wave, hospitalizations peak at a level only marginally higher than the peak of the first wave. This is consistent with the decline in the share of those recovering from COVID-19 in hospitals during the third wave relative to the first one.

A.4 Sentiment and Real GDP Growth Units

Table A.1 shows the OLS coefficients, the 95% confidence intervals for the coefficient estimates (in brackets), and the R^2 of a regression of four-quarter real GDP growth one quarter ahead on a constant and ENS_q over the sample 2000Q1-2019Q4, i.e., $\text{RGDPq4q4}_{q+1} = \alpha + \gamma \text{ENS}_q + u_t$, where $q \in \{2000\text{Q1}, \ldots, 2019\text{Q3}\}$.[28]

Table A.1

Regressors	Coefficients
Constant	1.61 [1.36;1.87]
Economic News Sentiment	1.17 [0.84;1.51]
$R^2 = 0.39$	

Figure A.7 plots the quarterly average of the daily economic news sentiment index expressed in GDP units along with one-quarter-ahead four-quarter real GDP growth.

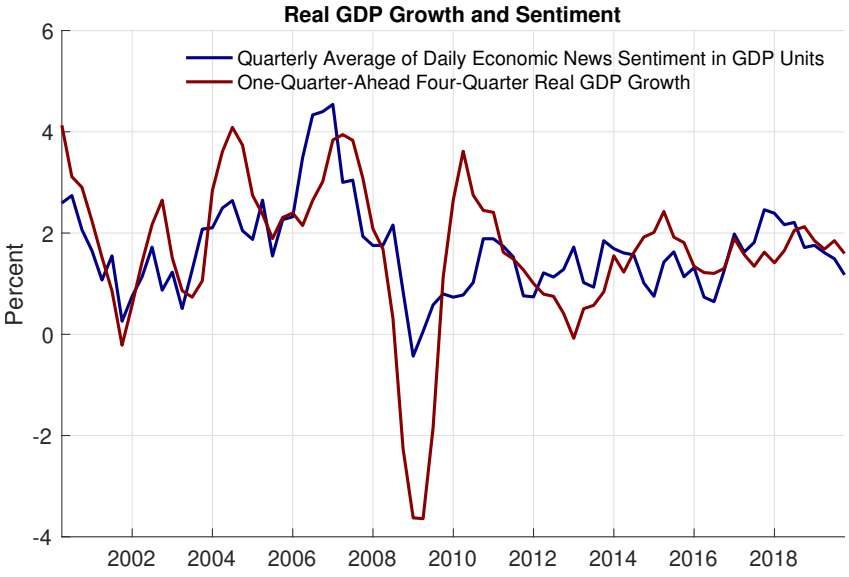


Figure A.7

[28]The real GDP data for Belgium were retrieved from the FRED database, Federal Reserve Bank of St. Louis.

A.5 Robustness to State Uncertainty

Figure A.8 gauges the robustness of our IRFs to state uncertainty. First, we draw a sequence of the state variables from the posterior distribution of the smoothed state variables. Second, we estimate an SVAR as in Section 5.1 and compute the point-wise posterior median IRFs. We repeat these two steps 100 times. For each SVAR, the posterior median is based on about 1,000 independent draws of the structural parameters.

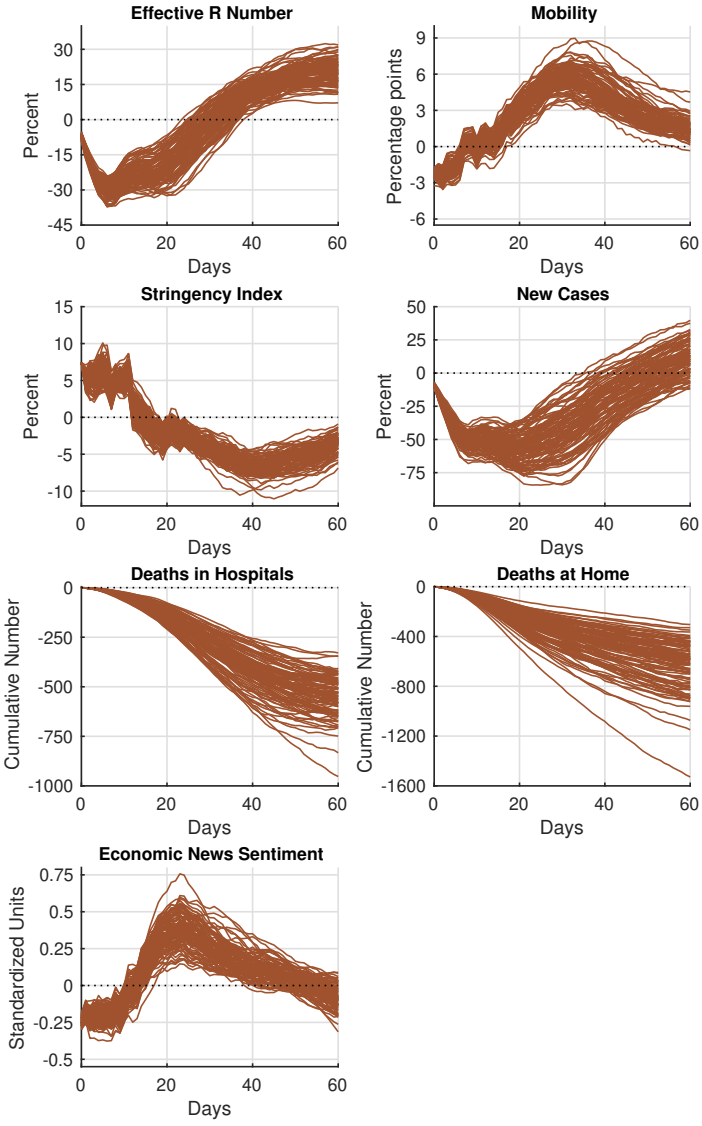


Figure A.8: IRFs to a stringency shock

A.6 Prior Robustness

Figure A.9 shows that the main conclusions from our SVAR analysis are robust to using the prior robust approach for SVARs proposed by Giacomini and Kitagawa (2018). The solid curves represent the point-wise posterior medians, the shaded areas represent the 68% equal-tailed point-wise probability bands, dotted curves represent the set of prior robust posterior means, and dashed-dotted curves depict the 68% robust credible regions. The figure is based on 1,000 independent draws of the reduced-form parameters and 100,000 orthogonal matrices draws for each reduced-form parameter.

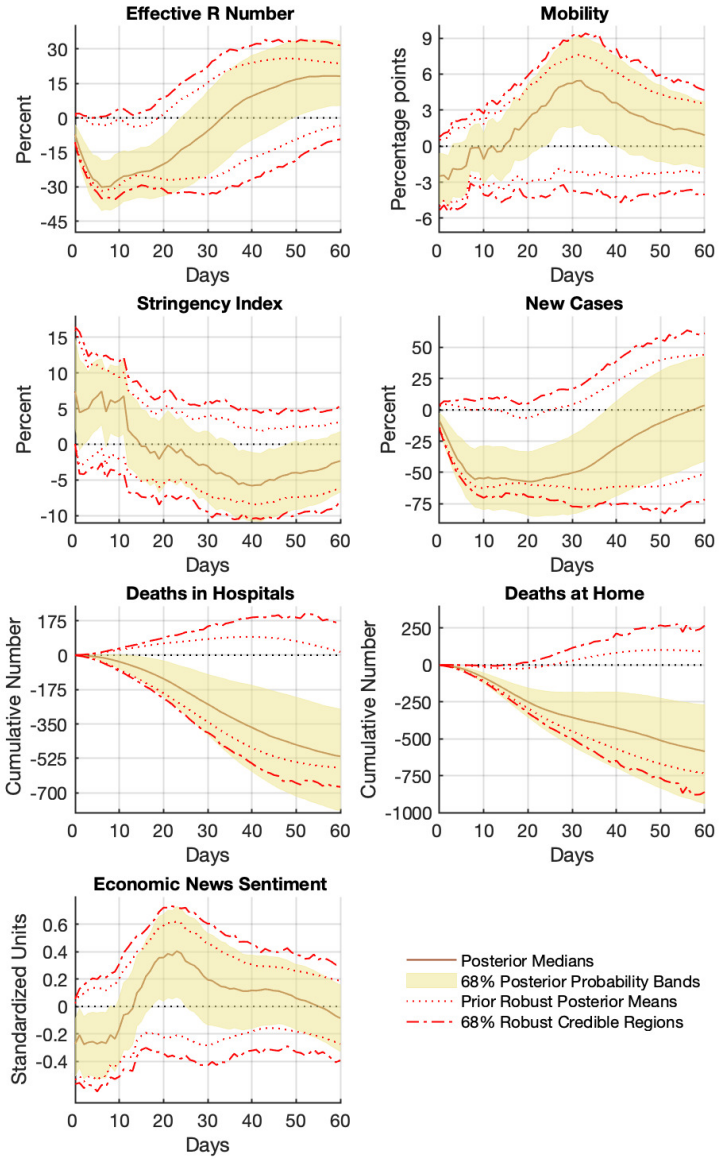


Figure A.9: IRFs to a stringency shock using a prior robust approach

A.7 Robustness to Higher Threshold

Figure A.10 shows that if the cutoff for high stringency is set higher, the main results do not change much. Notice that, in this case, the identified set for the IRF of the reproduction number in the high stringency regime is for the most part above the one in the low stringency regime. This is because, as the threshold for the stringency regime increases, the division between the early and the rest of the sample increases and hence the IRFs possibly reflect more contagious variants.

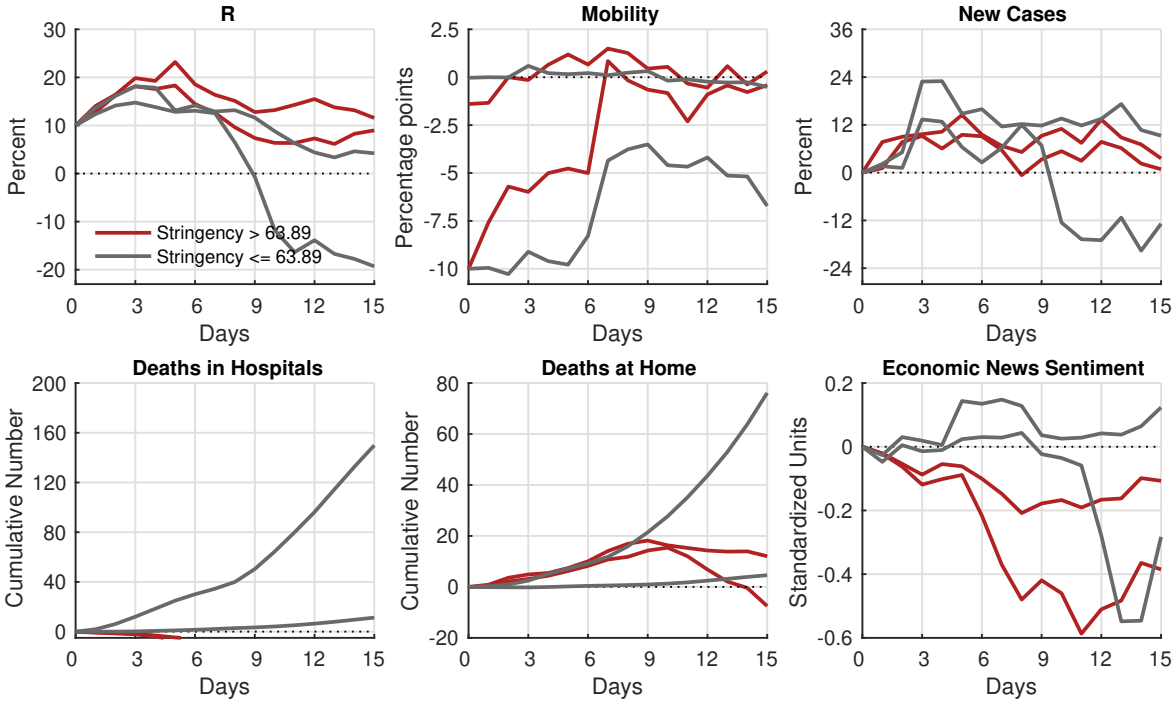


Figure A.10: IRFs to a reproduction shock